



Unfolding the link between multiple ecosystem services and bundles of functional traits to design multifunctional crop variety mixtures

Florence Dubs^{1,2,3} · Jerome Enjalbert² · Sebastien Barot³ · Emmanuelle Porcher¹ · Vincent Allard⁴ · Claude Pope⁵ · Arnaud Gauffreteau⁶ · Audrey Niboyet^{3,7} · Thomas Pommier⁸ · Sebastien Saint-Jean⁹ · Tiphaine Vidal^{5,9} · Xavier Le Roux⁸

Accepted: 14 September 2023
© The Author(s) 2023

Abstract

In face of the current environmental challenges, developing multifunctional cropping systems is increasingly needed, and crop variety mixtures are particularly interesting since they can deliver diverse services including grain production, yield stability, N₂O production regulation, disease control, and reduction of N-fertilizer losses. However, the relationships between intraspecific diversity and ecosystem multifunctionality are poorly understood so far, and practitioners lack science-based guidance to design mixtures. We used a pool of 16 bread wheat varieties classified into 4 functional groups based on 26 below- and aboveground functional traits, to conduct a field trial (88 large plots cultivated with single varieties or mixtures of 2, 4, or 8 varieties), quantifying 15 provisioning and regulating services for each plot. To assess yield stability between local conditions and years, the trial was replicated at 4 other locations and for 2 years, using 2 managements each time. We analyzed how variety number and functional groups predicted the variance in services, and applied in an innovative manner the RLQ co-inertia analysis to relate the (variety × traits) matrix Q to a (plot × services) matrix R , using a (plot × variety) composition matrix L as a link. Our results show that using variety mixtures allowed delivery of baskets of services not reachable when cultivating single varieties, and that mixtures mitigated tradeoffs between different pairs of services. Variety number or functional groups poorly predicted the variance in services, but the RLQ approach allowed the identification of groups of plots delivering consistent baskets of services. Moreover, we demonstrated for the first-time significant relationships between specific baskets of services and bundles of variety traits. We discuss how our results increase our understanding of intraspecific diversity–agroecosystem multifunctionality relationships, and propose the next steps using our new approach to support practitioners for designing variety mixtures that provide particular baskets of services.

Keywords Biodiversity–ecosystem functioning relationships · Ecosystem services · Plant functional traits · RLQ analysis · Tradeoffs · Wheat variety mixtures

Florence Dubs and Xavier Le Roux equally contributed to this work.

✉ Xavier Le Roux
xavier.le-roux@univ-lyon1.fr

- ¹ Centre d'Ecologie et des Sciences de la Conservation (Cesco), Sorbonne Université, Muséum national d'Histoire naturelle, CNRS, 75005 Paris, France
- ² Université Paris-Saclay, INRAE, CNRS, AgroParisTech, Unité GQE Le Moulon, 91190 Gif-sur-Yvette, France
- ³ Sorbonne Université, Université Paris Cité, UPEC, CNRS, INRAE, IRD, Institut d'Ecologie et des Sciences de l'Environnement de Paris (IEES-Paris), 75005 Paris, France
- ⁴ INRAE, Unité Génétique, Unité Diversité et Ecophysiologie des Céréales (GDEC), 63000 Clermont-Ferrand, France

- ⁵ Unité Biologie et Gestion des Risques en Agriculture (Bioger), INRAE, INA-PG, 78850 Thiverval-Grignon, France
- ⁶ Université Paris-Saclay, INRAE, AgroParisTech, Unité Agronomie, 78850 Thiverval-Grignon, France
- ⁷ AgroParisTech, 91120 Palaiseau, France
- ⁸ Laboratoire d'Ecologie Microbienne (LEM), Université Claude Bernard Lyon 1, INRAE, CNRS, VetAgroSup, 69622 Villeurbanne, France
- ⁹ Université Paris-Saclay, INRAE, AgroParisTech, Unité Écologie Fonctionnelle et Ecotoxicologie des Agroécosystèmes (Ecosys), 78850 Thiverval-Grignon, France

1 Introduction

Many agroecosystems are characterized by simplification of biological communities and high inputs of chemicals to fertilize soils and control crop pests, resulting in negative impacts on the environment and biodiversity (Pe'er et al. 2014; Tilman et al. 2002). These simplified systems might also be poorly adapted in the context of climate change (Brisson et al. 2010; Howden et al. 2007; van Etten et al. 2019). There is thus a need for an environmental transition in agriculture (Altieri 1999; Foley et al. 2011), which can be achieved by reinforcing ecological functions through better management of cultivated and wild biodiversity (Foley et al. 2011; Eggermont et al. 2015; Malézieux 2011). In this context, the multifunctionality of cropping systems, i.e., their capacity to provide diverse sets—or baskets—of services (Gaba et al. 2015; Renting et al. 2009) is of paramount importance. This transition, however, requires a shift in mindset and research paradigm since most studies focus on a single or very few ecosystem services, whereas agriculture and societies generally require a wide range of services. The latter include provisioning services such as grain yield, grain quality, and yield stability, along with regulating services such as disease control, biocontrol of weeds and crop pests, maintenance of soil fertility, regulation of climate change through reduced greenhouse gas emissions, and limitation of fertilizer-N losses from the soil-plant system. These services cannot always be maximized simultaneously, due to trade-offs between services (e.g., Shi et al. 2019), and increasing multifunctionality generally requires a higher level of biodiversity, as shown for grassland and forest ecosystems (Gamfeldt et al. 2013; Hector and Bagchi 2007; Zavaleta et al. 2010).

The genetic diversity of crop species is an important lever for designing more sustainable and resilient farming systems (Beillouin et al. 2021; Østergård et al. 2009; Sirami et al. 2019). Among the many ways to increase crop diversity in both space and time (from crop rotations to complex agroforestry systems), the use of crop variety mixtures is interesting because it can increase the in-field agrobiodiversity, promote a diversity of ecological processes, and ultimately enhance the delivery of multiple ecosystem services (Hajjar et al. 2008; Hughes et al. 2008). The practice of variety mixtures is not something recent, and was well developed, e.g., during the twentieth century until the 1970s in Western Europe and in Eastern European countries prior to the fall of the Iron Curtain (Finckh et al. 2000), although their use decreased at the end of the twentieth century (Bonnin et al. 2014). Recently, in Europe, new financial mechanisms allow the payment of farmers who decrease pesticide inputs through the use of variety mixtures (e.g., French Government 2018), which will promote the use of mixtures. Variety

mixtures have well-known positive effects on disease control (Finckh et al. 2000) and biocontrol of insects considered as crop pests (Tooker et al. 2012; Vidal et al. 2020), but can also benefit grain yield and yield stability (Vidal et al. 2020; Kiær et al. 2009; Sarandon and Sarandon 1995; Zeller et al. 2012), malting quality (Newton et al. 1998), and farmland biodiversity (Chateil et al. 2013; Johnson et al. 2006). With its directive 66/402, the European Commission has recently authorized the marketing of seed mixtures for cereals, allowing a more widespread use of mixtures. However, the choice of varieties to be grown together remains often empirical, and practitioners largely lack science-based guidance for selecting particular mixtures according to prioritized agronomic and environmental objectives (Barot et al. 2017). Important concepts and challenges associated with the design of variety mixtures have been identified by different authors (Barot et al. 2017; Borg et al. 2018; Litrico and Violle 2015; Wuest et al. 2021), but experimental trials have generally focused on yield and/or control of diseases (Lopez and Mundt 2000; Mille et al. 2006). We still need approaches that could guide the design of variety mixtures according to the multiple ecosystem services prioritized by practitioners and the screening of many possible mixtures.

One possible approach is to make better use of the insights gained from previous ecological studies of the relationships between biodiversity and ecosystem functioning (BEF), particularly on the processes underlying these relationships (Cardinale et al. 2003; Diaz and Cabido 2001; Loreau and Hector 2001; Weisser et al. 2017). For instance, the role of biodiversity is explained by complementary and selection effects (Loreau and Hector 2001), and the number of both species and functional groups are important for ecosystem functioning and the ensuing number and nature of provided services (Reich et al. 2004). The functional traits of plant species or genotypes are of major importance to understand and predict biodiversity effects on ecosystem functioning (Weisser et al. 2017; Lavorel and Garnier 2002; Tilman et al. 1997). Classifying crop varieties into functional groups based on the screening of many functional traits (Martin et al. 2015) therefore seems an appealing approach to design variety mixtures. However, it has been reported that *a priori* classification of plant species into functional groups is rarely effective to predict ecosystem functioning (Wright et al. 2006), due to the lack of knowledge about the impact of functional traits (their mean values and their diversity) on ecosystem functioning and the provision of services. Alternative approaches to *a priori* classification of crop varieties into functional groups should be tested to help farmers choose varieties to be mixed to reach specific agronomic and environmental goals.

The objective of this study was to analyze to what extent the functional diversity associated with varieties/mixtures within

Fig. 1 View of a part of the experimental trial in Versailles used to unfold the link between the variety (and functional trait) diversity within wheat plots and the baskets of services delivered by these plots. Each 10.5 m × 8.0 m plot was buffered from adjacent plots by 1.75-m-wide rows of triticale (darker green). The same field trial was replicated with smaller plots at four other locations in France (see Fig. S3).



wheat plots can predict a set of agroecosystem services provided by these varieties/mixtures. Here, we used 16 wheat varieties characterized for 26 below- and aboveground functional traits (Table S1, see Cantarel et al. 2021) and classified accordingly into 4 functional groups (Fig. S1, see Dubs et al. 2018). This pool of varieties was used to design a main field trial at Versailles with 88 wheat plots cultivated with single varieties or mixtures of 2, 4, or 8 varieties differing in terms of functional group composition (Fig. 1; Fig. S2 & S3). Thirteen provisioning and regulating services (listed in Table S2) were quantified for each plot over one year on this main field trial. In addition, the experimental design was replicated using smaller plots at 4 other locations and for 2 years, using 2 managements (conventional and low input) each time, to assess two other proxies of services, i.e., yield stability between years and yield stability between locations × managements. This led to the quantification of 15 (proxies of) services in total. We initially assumed that (i) the variety number would be a poor predictor of the variance of the different services studied, whereas functional group composition would better predict services as reported in previous BEF studies. Further, we hypothesized that (ii) a higher part of the variance in the different services could be explained using the RLQ method (Dray and Legendre 2008), which could also unfold the link between variety traits and services. This method initially designed to assess trait filtering by environmental variables was not applied in BEF studies before. RLQ is a co-inertia analysis that couples multiple data sets and identifies co-relationships between them. Normally, the RLQ analysis relates a (species × functional traits) matrix Q to a (sites × environmental variables) matrix R , using a (sites × species abundances) matrix L as a link. Here, sites are wheat plots, and we replaced species by varieties in the matrices Q and L , and environmental variables by ecosystem services in the matrix R . This corresponds to a novel approach for assessing to what extent particular baskets of services are associated to specific bundles of functional traits (mean values and diversity of variety traits within plots).

We discuss the implications of our results for the advancement of knowledge of BEF relationships focused on the role of intraspecific diversity. We also discuss the implications for the design of variety mixtures able to deliver baskets of services prioritized by farmers.

2 Materials and methods

The experimental design is based on the design of recent BEF experiments (e.g., Weisser et al. 2017) where both richness (here variety number) and functional group diversity are manipulated concurrently to minimize their correlation (Fig. S2). In this type of design, the objective is not to compare the performance of specific mixtures (hence no replicates of particular varieties or mixtures are used) but to analyze the role of variety diversity, in particular richness and functional group number. True replicates are used for each level of variety richness × functional group number using different variety compositions (Fig. S2).

2.1 Variety phenotyping, and clustering into functional groups

Initially, a set of 57 wheat lines widely used in France was phenotyped: (i) 32 elite varieties of bread wheat; (ii) 5 modern varieties widely used in organic or low-input farming systems; (iii) 9 landraces resulting from mass selection by farmers, cultivated in France in the early twentieth century; and (iv) 11 highly diverse lines derived from a MAGIC (multiparent advanced generation intercross) mapping population. We used the data collected on these 57 varieties for 26 agronomic (e.g., yield components, earliness and disease control) and ecophysiological (e.g., specific leaf area and root absorption capacity of mineral N forms) traits (Table S1; detailed in Cantarel et al. 2021).

Multivariate clustering analysis (Ward method) resulted in 6 functional groups, FGs, of varieties, two of them being

unstable (see details in Dubs et al. 2018). The four stable functional groups retained were characterized as follows: FG1 included varieties sensitive to fungal diseases and with root traits indicative of low potential for soil exploration/exploitation; FG2 was composed of varieties with intermediate values for most functional traits; FG3 included varieties with slow growth but high plant-plant aggressiveness; and FG4 gathered varieties resistant to fungal diseases and with high potential for soil exploration/exploitation. Four varieties were randomly selected within each functional group.

2.2 Study sites and experimental design

The 16 selected varieties were assembled into 72 different mixtures (plus the 16 varieties in pure stands, hence a total of 88 plots), allowing us to explore a gradient in both variety number (1, 2, 4, and 8) and functional diversity (1, 2, 3, or 4 functional groups) (Fig. S2).

The main experimental trial was located at the INRAE experimental station in Versailles, France (Fig. S3). The characteristics of the site in Versailles are presented in Table S3. The soil had a total N content of $1.0 \pm 0.1 \text{ g kg}^{-1}$, a total C content of $12.2 \pm 1.8 \text{ g kg}^{-1}$, and an organic matter content of $21.0 \pm 3.1 \text{ g kg}^{-1}$, with $204 \pm 21 \text{ g kg}^{-1}$ of clay, $534 \pm 74 \text{ g kg}^{-1}$ of silt and $260 \pm 57 \text{ g kg}^{-1}$ of sand for the 0 to 15 cm depth. Pure stands and mixtures were sown in November 2014 with $180 \text{ g of seeds m}^{-2}$. The total seed density was the same in the mixtures and monocultures, and the component varieties in the mixtures were all equi-proportional, i.e., 50:50 for two varieties mixed; 25:25:25:25 for four varieties mixed, etc. The plots ($10.5 \text{ m} \times 8.0 \text{ m}$) were randomly distributed over the experimental area and each plot was buffered from adjacent plots or site edge by a 1.75-m-wide row of triticale (\times Triticosecale) (Fig. 1). The plots were managed conventionally but with relatively low input levels, i.e., only one herbicide spray (Archipel® and Harmony Extra®) in mid-March, and sub-optimal levels of ammonium-nitrate applied (i.e., a total of 140 kg-N ha^{-1} representing 67% of the N input needed to reach yield potential at this site: 40 kg-N ha^{-1} on March 5, 60 kg-N ha^{-1} on April 16, 40 kg-N ha^{-1} on May 11). No insecticide/fungicide was used.

In addition, the same field trial (but with smaller plots) was replicated at four other locations (also INRAE experimental stations) in France during the same growing season (2014–2015) and the following one (2015–2016) using two managements (high and low fertilization) each time (i.e., 88 plots \times 4 locations \times 2 managements \times 2 years), in order to quantify the yield stability across managements/sites and across years. The main characteristics of the sites are presented in Table S3 (see also a map in Fig. S3). At each site, the 88 variety compositions were tested for two different N input levels (conventional vs.

low input). Each trial was based on a randomized block design with 2 blocks and all the (88 variety compositions \times 2 N levels) by block. Under conventional management, plots were managed following local agronomic practices (full pest management including herbicides, fungicides, and insecticides and the additional use of a growth regulator in Rennes). The required mineral N input for optimal yield was calculated based on previous culture mineral N leftover and maximum yield expectations for each site. The low-input management differed only by the amount of N-fertilizer applied with a pre-defined yield objective of ca. $-100 \text{ kg-N ha}^{-1}$ compared to the conventional management. In practice, given that the occurrence of high leftovers at some sites for some years, N applied at the four sites and two years were (all numbers in kg-N ha^{-1}): Clermont-Ferrand 2015 (conventional: 60; low input: 0); Clermont-Ferrand 2016 (140, 60); Dijon 2015 (130, 60); Dijon 2016 (170, 50); Rennes 2015 (110, 30); Rennes 2016 (155, 60); Toulouse 2015 (215, 120); and Toulouse 2016 (215, 115). Plot size was 7.5 m^2 in Clermont-Ferrand, 8.25 m^2 in Dijon, 6.34 m^2 in Rennes, and 8.78 m^2 in Toulouse. Plots were harvested at full physiological maturity during the summers 2015 and 2016, and grain yield was expressed on a 15% water content basis. These data were used to compute two yield stability indices, i.e., stability across sites and managements, and stability across years (see below).

2.3 Quantification of (proxies of) services in wheat plots

Fifteen (proxies of) services (4 proxies of provisioning services plus 2 indices of yield stability, and 9 proxies of regulating services) were quantified for all the 88 plots of the main field trial in Versailles.

Grain yield All plots were harvested during the last week of July and the first week of August, with a MB Hege 140 combine harvester (Hege Maschinen GmbH, Waldenburg, Germany) with a cut-width of 1.75 m, for the central 1.75 m section along the entire plot length. Grain yield was expressed as mean grain mass measured at 15% humidity, in t ha^{-1} .

Grain N content The percentage of N content of the grains is a crucial feature determining the economic value of a wheat crop and the functional quality of the flour. Grains were dried at 80°C and were ground. The N content was determined using a CN elemental analyzer (Vario Microcube, ELEMENTAR, Germany) according to the Dumas method, using on average 6.2 mg of milled grains.

Specific weight of grains Specific weight (i.e., weight of a volume of grain, sometimes also called “hectoliter mass”) is another important facet of the grain production service. It

is one of the oldest specifications used in wheat grading and an important indicator of the physical quality of wheat, with requirements defined for milling or export (Manley et al. 2009). Grain specific weight was measured using a Dickey John GAC 2000 grain analysis computer (Church Industries, Minneapolis, USA).

Lodging resistance It was derived from the quantification of the percentage of the plot surface which was lodged (%LS). This surface was quantified visually using a 5% step scale. Lodging resistance was expressed as $(100 - \%LS)$, with 0%—highly susceptible, to 100%—fully resistant.

Shoot biomass before harvest Shoot (straw and foliage) biomass was harvested from June 1 to June 6, 2015, at the onset of flowering in sub-plots of 50×52.5 cm centered on three rows by uprooting whole plants and by separating shoots from roots. Shoots were dried at 65°C for 72h and weighed. Data are expressed in g m^{-2} .

Recovery efficiency of the N-fertilizer by wheat A ^{15}N labeling experiment was conducted in the 88 experimental plots on the main field experiment in Versailles to quantify the capacity of the monocultures and variety mixtures to exploit the N-fertilizer applied. To do so, and in addition to the application of 40 kg-N ha^{-1} of a nitrogenous fertilizer (ammonium-nitrate $\text{NH}_4\text{-NO}_3$) in the 88 experimental plots on March 5, 2015, an application of ^{15}N , as $^{15}\text{NH}_4^{15}\text{NO}_3$ (i.e., the same form as the N-fertilizer used at the study site) at a rate of $36 \text{ mg } ^{15}\text{N m}^{-2}$ (labeled at 98%) was made by sprinkling slowly one liter of demineralized water containing the labeled N over an area of $90 \text{ cm} \times 90 \text{ cm}$ in each experimental plot (avoiding any lateral transfer to the nearby area) on March 11, 2015. Total shoot and root biomasses of wheat were then collected from June 1 to June 6, 2015, at flowering stage, in a $50 \text{ cm} \times 52.5 \text{ cm}$ area located in the middle of the same $90 \text{ cm} \times 90 \text{ cm}$ labeling area, which included 3 wheat rows. All shoot biomass was collected. Roots were collected from 0 to 15cm: main roots were collected when collecting the whole plants on the three rows and fine roots from 0 to 15cm using two soil (8 cm diameter) cores. Roots were washed and shoot and root biomasses were then dried at 65°C for 72h before weighing. Shoot and root biomasses were expressed in g m^{-2} . Shoot and root materials were ground and analyzed for %N and $\delta^{15}\text{N}$ using an elemental analyzer coupled to an isotope ratio mass spectrometer (EA-IRMS, Carlo-Erba NA-1500 NC Elemental Analyzer on line with a Fisons Optima Isotope Ratio Mass Spectrometer). Shoot and root ^{15}N amounts ($\text{mg } ^{15}\text{N m}^{-2}$) were calculated as the atom % excess ^{15}N concentration in shoot or root (measured atom % ^{15}N minus natural abundance ^{15}N) times the mass of N in the shoot or root biomass (g N m^{-2}). Shoot or root ^{15}N recovery efficiency was determined as the ratio

between shoot or root ^{15}N content and the amount of ^{15}N added, and expressed in %. Total ^{15}N recovery efficiency by wheat plants was determined as the sum of shoot and root ^{15}N recovery efficiencies, expressed in % of the added N.

Recovery efficiency of the N-fertilizer by the soil-plant system The same ^{15}N labeling was used to quantify the ability of the different varieties and mixtures to retain the fertilizer-N applied in the plant–soil system (i.e., ability to reduce total losses of the added N through denitrification, leaching). Shoot and root ^{15}N recovery were determined as described above. In addition, soil N and ^{15}N contents were measured using soil cores (8 cm diameter) retrieved from the 0- to 15-cm soil layer. The soil samples were oven dried for 24h at 105°C , ground and analyzed for %N and $\delta^{15}\text{N}$ using an elemental analyzer coupled to an isotope ratio mass spectrometer (EA-IRMS, Carlo-Erba NA-1500 NC Elemental Analyzer on line with a Fisons Optima Isotope Ratio Mass Spectrometer). Soil N content of the 0–15-cm soil layer was calculated as the product of the soil N concentration and total soil amount per square meter in the 0–15-cm layer, i.e., accounting for soil bulk density. It was expressed in g N m^{-2} . Soil ^{15}N mass (in $\text{mg } ^{15}\text{N m}^{-2}$) was calculated as the atom % excess ^{15}N concentration times the mass of N (g N m^{-2}) in the 0–15-cm layer. Total ^{15}N recovery efficiency by the plant–soil system was determined as the sum of shoot, root, and soil ^{15}N recovery, expressed in % of the added ^{15}N .

Soil supply of mineral N Time-integrated values of soil N-mineral made available for plants were quantified using resin bags that allow quantification of the release of mineral N by soil over time (Robson et al. 2007). Five resin bags composed of 5 g of Amberlite® IRN-150 ion exchange resin (VWR) were buried in each of the 88 plots at 10cm depth in May 2015 and let in the field for 10 days. After the resin bag recovery, the ions adsorbed on the resin were desorbed using a HCL 1M solution. For each resin bag, nitrate and ammonium levels from the desorbed solution were measured using a photometer (Smartchem 200, KPM Analytics).

Regulation of N_2O production by soil Soil was sampled in May 2015, at the beginning of the flowering phase. In each plot, ten soil samples were randomly taken using a corer (0–8 cm depth; 8cm diameter), i.e., a total of 880 cores (10 cores \times 88 plots), and a composite soil sample was obtained by pooling the ten samples per plot. For each of the 88 composite samples, fresh soil was sieved (2-mm mesh) and stored in plastic bags at $+4^\circ\text{C}$ a few days before measurements. The capacity of soil to produce N_2O was measured using fresh soil (10 g dw equivalent) according to Patra et al. (2005) for all the composite samples with a gas chromatograph (R3000 SRA, Marcy l'Etoile France). The N_2O production capacity was calculated from the linear rate

of production of N_2O during an 8-h incubation. A proxy of the service “regulation of N_2O production” was computed for each plot as: maximum production rate observed across plots—production rate for the given plot.

Predation for biocontrol of crop pests Predation rates were estimated during 24 h at the following dates: April 22, May 4, May 18, and July 2, 2015, according to the method of Östman (2004). In each plot, 40 identical baits (dried *Calanus* copepods) were used to measure the rate of prey removal (strictly speaking the rate of scavenging), which we used as a proxy for predation rates. These baits were glued on small pieces of adhesive tape (1 cm × 4 cm Tesa double-sided tape taped on white paper), with 10 baits × 4 pieces of adhesive tape per plot. The baited papers were anchored to the ground with pins and spread homogeneously in each plot. The rate of bait removal was then monitored through time. At each date, the experiment started around 10 am. At all dates, the number of baits removed from each paper was recorded after ca. 4 h of exposure to predators. We tested the relevance of using dead preys to assess predation rates by comparing the rate of removal of dead vs. live preys, which were provided in equal amounts in all plots on the first monitoring session. The rate of removal was slightly higher for live vs. dead preys, but the differences among plots was not significant, without any interaction between type of prey and wheat diversity treatments, such that the rate of dead prey removal could be used as a good proxy for rate of live prey removal. The level of predation for control of crop pests was expressed in each plot as the sum across the four measurement dates of the mean number of baits removed after 4 h per baited paper.

Weed control A floristic survey was carried out in each plot on four equal areas of 2 m × 1.75 m (the most abundant weed species being *Agrostis stolonifera*, *Equisetum arvense*, and *Galium aparine*). All weed individuals were counted. Weed abundance in plot p , W_p , was measured as the total number of individual plants N_p standardized as follows to vary between 0 and 100: $W_p = 100 \times N_p / \max_p(N_p)$, where $\max_p(N_p)$ is the maximum weed abundance across plots. The level of weed control was characterized as $\log[(100 - W_p) + 1]$. Log-transformation was used to improve normality (one outlier plot with a large weed abundance).

Control of yellow rust and Septoria leaf blotch Yellow rust disease levels were assessed three times during the cropping season for the three upper leaves of single stems to catch the dynamics of the epidemic. The date with largest contrast between plots (heading stage) was retained for further analysis. A semi-quantitative scale (10% steps) was used, based on the percentage of the total leaf area covered by sporulating lesions. The number of main stems scored per plot was

proportional to the number of cultivars included in the mixture (i.e., 8 stems per monoculture plot, but 16, 32, and 64 stems in 2-, 4-, and 8-variety mixture plots, respectively) to generate data with sampling efforts that account for mixture complexity. This led to a total of 2688 stems scored over the cropping season. The control of yellow rust was expressed as: (100 - the percentage of the leaf surface visibly infected by yellow rust). The same method was used to quantify control of *Septoria* leaf blotch, which was also expressed as (100 - the percentage of the leaf surface visibly infected by *Septoria*).

Yield stability Two indicators of yield stability were computed. The first one is a spatial stability indicator calculated on yields quantified in 2015 which measures the ability of the mixtures/varieties to keep their yield stable between Versailles and 8 other environments that are combinations of site (Clermont-Ferrand, Dijon, Rennes, Toulouse) and management (conventional, or low input). This is thus a proxy of the ability of each mixture to face different sites (soil/climate) and management conditions on a given year. The yield of mixture/variety i in environment j and block k is expressed as follows:

$$Y_{ijk} = \mu + V_i + E_j + EB_{jk} + VE_{ij} + \varepsilon_{ijk}$$

with μ the average yield of the mixtures/varieties in Versailles, V_i the yield of mixture/variety i in Versailles (meaning the difference between the yield of mixture/variety i and the average yield μ), E_j the mean response of the mixtures/varieties to environment j , EB_{jk} the effect of block k in environment j , VE_{ij} the interaction between mixture/variety i and environment j , and ε_{ijk} the residue for mixture/variety i in environment j and block k .

The yield variability across sites/managements was calculated for each mixture/variety as follows:

$$\sqrt{\frac{1}{8} \sum_{j=1}^8 (VE_{ij})^2}$$

Those yield variabilities (in t/ha) were then converted to stability indicators by calculating their absolute distance to the maximum value across all the estimates of performance variability.

The second indicator measures the inter-annual stability of the mixtures/varieties and corresponds to the capacity of each mixture/variety to limit yield variations between 2015 and 2016 over the 8 environments studied (site × management combinations). The yield of the mixture/variety i in environment j , in year m and in block k is expressed as follows:

$$Y_{ijmk} = \mu + V_i + E_j + A_m + AE_{mj} + VE_{ij} + AEB_{mjk} + VAE_{imj} + \varepsilon_{ijmk}$$

with μ the average yield of mixtures/varieties in 2015, V_i the average performance of mixture/variety i in 2015 (i.e., difference between the average yield of mixture/variety i

in 2015 and μ), E_j the average response of the mixtures/varieties to the environment j in 2015, VE_{ij} the interaction between mixture/variety i and environment j , A_m the average change in yield between 2015 and 2016, $A_m + AE_{mj}$ the average change in yield between 2015 and 2016 in environment j , $A_m + AE_{mj} + VAE_{imj}$ the average change in yield between 2015 and 2016 for mixture/variety i in environment j , AEB_{mjk} the effect of block k in year m in environment j , and ε_{ijmk} the residue for mixture/variety i in environment j , year m and block k . The inter-annual yield variability was calculated for each mixture/variety i as follows:

$$\sqrt{\frac{1}{8} \sum_{j=1}^8 (A_m + AE_{mj} + VAE_{imj})^2}$$

The inter-annual yield variabilities for each mixture/variety (in t/ha) were then converted to stability indicators by calculating the difference to the maximum value.

2.4 Data analysis

For each ecosystem service, the spatial coordinates (longitude X and latitude Y) of plots, and their squared value, were used to test for possible spatial gradients or border effects. Spatial gradients were detected for 4 services (grain quality, biocontrol of crop pests, regulation of N_2O production, and shoot biomass) and corrected by retaining the residuals of a linear regression of these services on the 4 spatial variables X , Y , X^2 , and Y^2 .

To assess tradeoffs and synergies between ecosystem services, Spearman correlations were computed between all pairs of services, either for the 16 pure stands or 72 mixtures, using the PerformanceAnalytics package CRAN 2.0.4 (Peterson et al. 2020).

Single or multiple linear models and analyses of variance were used to predict each service according to mixture composition in terms of (i) variety number, (ii) functional group number, and (iii) percentage of each functional group. The significance of effects was calculated using type II sums of squares for unbalanced designs (Bolker et al. 2009). The goodness of fit of each model was calculated as an adjusted R^2 for linear models.

To describe the relationships between baskets of services and bundles of variety traits, we performed a RLQ analysis linking a table R (here plots \times ecosystem services) and a table Q (varieties \times functional traits) through a table L (plots \times variety compositions). The RLQ analysis consists in analyzing the joint structure of these three tables to decompose the variance of each component of the cross-matrix, and it provides the common ordination axes onto which variety traits and ecosystem services are projected (Dolédéc et al. 1996). First, each table was separately analyzed by a specific multivariate

analysis, allowing the determination of the proportion of the total variance of each table represented by the RLQ. The variety composition table L was analyzed by a correspondence analysis, while principal component analysis was applied to the quantitative trait table Q and service table R . The significance of the relationship between services and traits was tested using random permutations (Dray and Legendre 2008; Ter Braak et al. 2012). Finally, a Ward's hierarchical classification based on Euclidian distance along the first two RLQ axes allowed an a posteriori clustering of plots and a more synthetic description of plot properties. The criteria used to define the clusters were (i) to generate between 4 and 8 groups, in order to identify sufficiently different baskets of services (gathering plots into too large groups would indeed dilute group specificities) while avoiding to generate too many groups that would include a too low number of varieties/mixtures (which would restrict statistical tests between groups), and (ii) to make sure that the number of groups was stable with a $\pm 10\%$ relative change in the value of the distance threshold used for the clustering. Each of the plots was further described in terms of baskets of services and bundles of variety traits (sown community-weighted mean values of traits, CWM, and Rao's Q diversity index, RaoQ, for each trait, that are frequently used to measure functional diversity; Botta-Dukát 2005). The CWM and RaoQ values were computed at sowing, i.e., considering the initial and balanced proportions of variety mixture components. Realized sown values would have been interesting to measure, but quantifying the (mass) proportions of the individual mixture components is hardly tractable for variety mixtures. For each variable (each service, and each trait in term of sown CWM or RaoQ), the difference between groups of plots was calculated using type II sums of squares for unbalanced designs in linear models. Effects were tested using multiple comparisons of means (Tukey's honestly significant difference). We finally tested whether the baskets of services associated to the groups of plots defined from the RLQ results were related to the bundles of traits characterizing these groups.

All statistical analyses were performed using R (version 4.0.5), including ADE-4, and—for unbalanced design—the car Package (Fox and Weisberg 2019).

3 Results and discussion

3.1 Variety mixtures weaken tradeoffs and synergies normally observed between services when cultivating single varieties

The range of grain yield values observed across the 88 plots from the main field trial was 4.49 to 8.27 t ha⁻¹. Analyses

of mixture effects on specific services with absolute values of services can be found in recent publications of the Wheatamix consortium (in particular Vidal et al. 2020 for yield and disease severity). When considering the 16 wheat varieties cultivated alone (not the mixtures), strong positive correlations were observed between some services (Fig. S4), in particular (i) between rust control and either grain yield, weed control, or resistance to lodging (Spearman correlation coefficients ρ from 0.44 to 0.84); (ii) between grain yield and shoot production ($\rho = 0.52$); and (iii) between the recovery efficiency of fertilizer-N by the soil-plant system and grain yield or shoot production ($\rho = 0.54$ to 0.73). In contrast, tradeoffs (i.e., negative correlations) among services were observed in particular (i) between grain N content and either grain yield, yellow rust control, weed control, or resistance to lodging ($\rho = -0.48$ to -0.81), and (ii) between the level of predation for biocontrol of crop pests and recovery efficiency of the fertilizer-N by the soil-plant system ($\rho = -0.49$) (Fig. S4).

The three positive correlations observed with yellow rust control can be explained by the fact that this trait has been selected along with resistance to lodging and variety capacity to produce shoots and grains by farmers and seed companies (Ellis et al. 2014; Shah et al. 2019). The positive link between fertilizer-N recovery efficiency and shoot production makes sense as the higher the plant biomass production, the higher its demand for N, and likely the higher its immobilization of fertilizer-N. The negative correlation observed between grain yield and grain N content is well-known and discussed in the literature (e.g., Bogard et al. 2010). In contrast, understanding some positive correlations such as the one observed between yellow rust control and weed control,

or a negative correlation such as the one observed between crop pest biocontrol and recovery efficiency of the fertilizer-N by the soil-plant system, is not straightforward. This type of correlation may be due to either synergies between plant traits in relation to genomic features and/or to hidden variety selection effects.

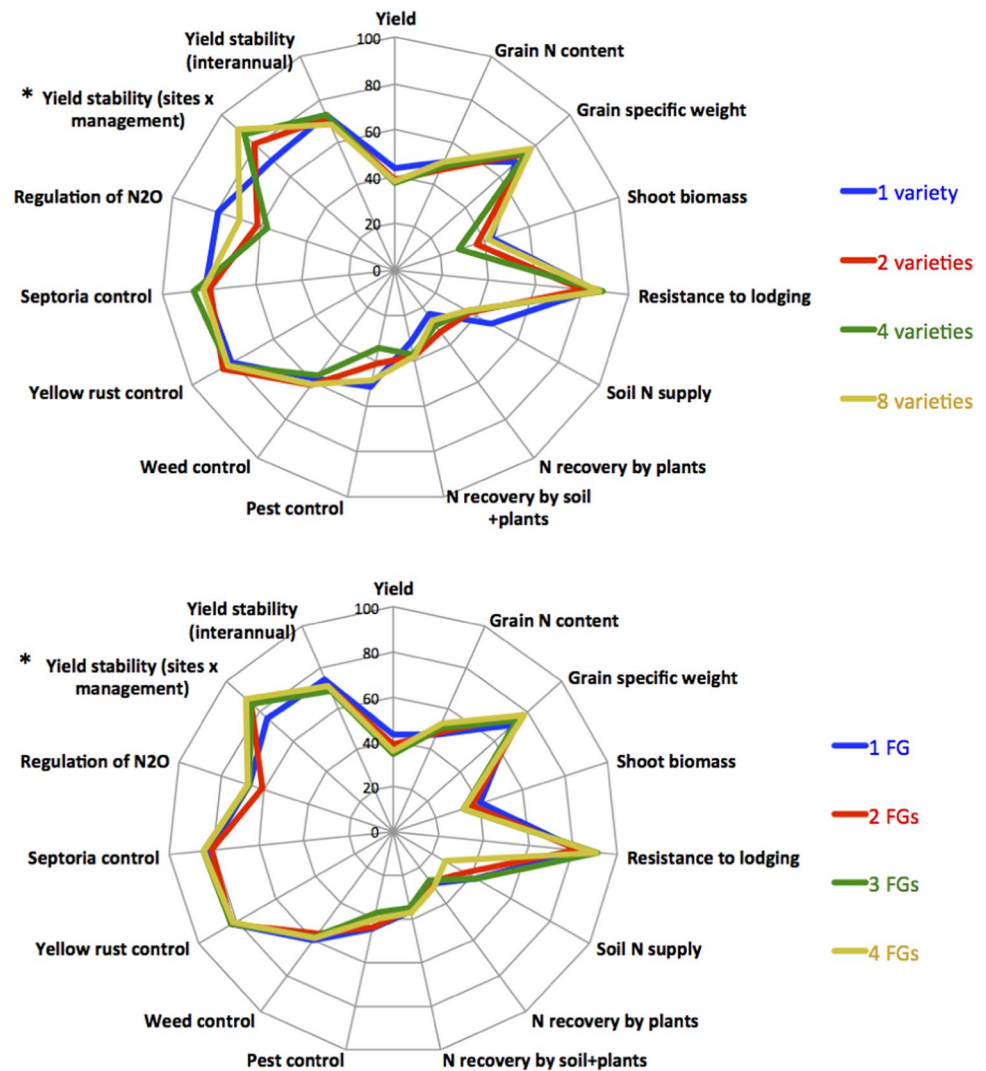
When considering the varieties in pure stands only, in total 13 significant correlations were observed between services, 12 with a ρ absolute value ≥ 0.49 . But when considering the wheat mixtures rather than the 16 varieties in monocultures, only 2 correlations were observed with a ρ absolute value ≥ 0.49 (0.49 and -0.61 ; Fig. S5). This weakening of the strength of the correlations between services (especially negatives ones, i.e., tradeoffs) thanks to the use of mixtures is interesting because it paves the way to obtaining baskets of services not reachable when cultivating only a single variety in each plot.

The—on average—lower strength of the correlations between services observed when considering the mixtures (variety number > 1) than when considering the pure stands are consistent with a “Jack-of-all-trades” effect, which formalizes the notion that species—or here genotypes—have a certain degree of specialization so that a single species/genotype can maximize particular functions favorable to some services but at the expense of other functions and services, due to tradeoffs between functional traits at the plant individual level (Futuyma and Moreno 1988). Furthermore, it has been reported that the strength of plant diversity effects differs between different categories of (agro)ecosystem processes (Allan et al. 2013), which can also explain modified relationships between the services provided by mixtures compared to those observed for monocultures.

Table 1 Fraction of variance in ecosystem services explained by the number of varieties (var #), the number of functional groups *a priori* defined from variety functional traits (fg #), the percentages of these 4 functional groups (% fg), and the groups identified from the RLQ analysis (RLQg). *, $0.01 < p < 0.05$; **, $0.01 < p < 0.001$; ***, $p < 0.001$; ns, not significantly different.

(Proxy of) service	var # R ² adj	P	fg # R ² adj	P	% fg R ² adj	P	RLQg R ² adj	P
Grain yield		ns		ns	0.381	***	0.644	***
Grain N content		ns		ns	0.507	***	0.638	***
Grain specific weight		ns		ns	0.559	***	0.731	***
Shoot biomass		ns		ns		ns		ns
Yield stability across sites	0.122	**	0.066	*		ns	0.435	***
Yield inter-annual stability		ns		ns		ns	0.183	*
Predation for pest control		ns		ns		ns	0.105	*
Regulation of weeds		ns		ns		ns		ns
Regulation of N ₂ O		ns		ns		ns	0.255	**
Lodging resistance		ns		ns	0.361	***	0.524	***
Yellow rust control		ns		ns	0.313	***	0.586	***
Septoria control		ns		ns	0.198	***	0.413	***
Soil N supply to plants		ns		ns		ns		ns
N recovery efficiency by plants		ns		ns		ns	0.169	*
N recovery efficiency by the soil-plant system		ns		ns		ns		ns

Fig. 2 Radar charts presenting the baskets of the 15 ecosystem services characterizing the 88 wheat plots classified according to the within-field number of (Top) varieties or (Bottom) variety functional groups (FGs) defined based on 26 variety functional traits. For each service, scores of 0 and 100 correspond respectively to the lowest and highest level of service observed across all plots. No significant effect of variety number or functional group number was observed on any service, except for the yield stability across sites x managements (variety number effect: $p=0.0006$; FG number effect: $p=0.027$).



3.2 Variety number and *a priori* defined functional group composition poorly predict services

We analyzed the 15 (proxies of) ecosystem services documented for the 88 wheat plots based on their composition in terms of variety number and functional groups defined from the previous quantification of 26 functional traits for each variety (Table 1). Variety number did not influence any of the 15 ecosystem services, except—though weakly—yield stability across sites/managements that increased significantly with variety number (Table 1; Fig. 2). Similarly, the number of functional groups did not influence any ecosystem service, except marginally yield stability between sites/managements (Fig. 2). Linear modeling showed that information on the percentages of the 4 functional groups of varieties present in each mixture was useful to predict only 6 of the 15 services, and generally with a weak predictive power (2 services with decent R^2 : 0.51 and 0.56 for grain N content and specific weight, respectively; and lower R^2 ranging from

0.20 to 0.38 for the 4 other services). Overall, despite the huge effort devoted here to the categorization of the varieties into functional groups using 26 below- and aboveground functional traits, information on variety number and functional groups was not sufficient to predict most ecosystem services well (Table 1) and to infer which baskets of services were associated to which types of mixtures (Fig. 2).

The lack of predictive capacity of variety number for most of the studied services is consistent with the conclusions of many BEF studies indicating that plant functional composition is much more important than richness for ecosystem functioning and services (Tilman et al. 1997; Le Roux et al. 2013; Weisser et al. 2017). Interestingly, wheat variety number only influenced yield stability between sites/managements. This is consistent with many reports showing that increasing species richness (in particular in grasslands) decreases the temporal variations of whole-community biomass (Gross et al. 2014). The limitation of the *a priori* classification of plants into functional groups has also already

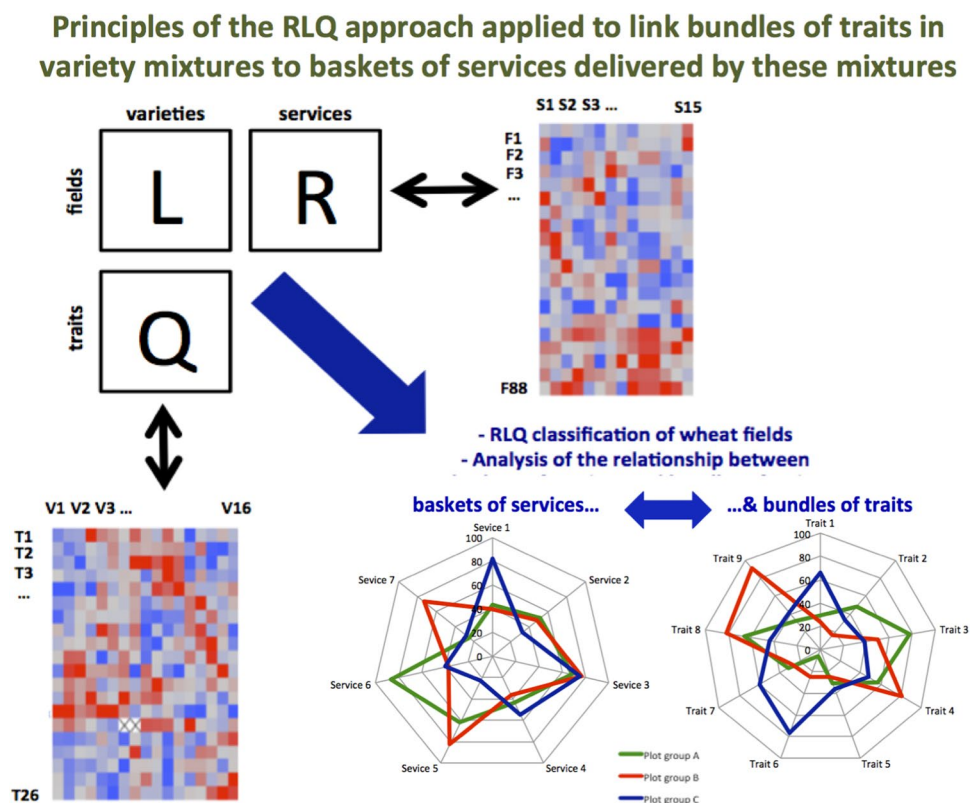
been observed in BEF studies. For instance, studies linking plant functional diversity to ecosystem functioning typically employ *a priori* classifications of species into hypothetically complementary groups like grass/forb/legume. However, Wright et al. (2006) reported that the predictive capacity of such classifications was seldom significantly higher than that of random classifications, and that optimal post hoc classifications of species had a higher predictive power of ecosystem functions. Several authors acknowledged that alternative classifications based explicitly on species ecophysiological and morphological traits might be more useful (Reich et al. 2004, Craine et al. 2002; Petchey and Gaston 2002) and capture more of the functional variation that leads to diversity effects than traditional classifications (Petchey 2004). Nevertheless, our results using this kind of trait-based approach show that the sole large-scale phenotyping of crop varieties might be a cul-de-sac for designing variety mixtures relevant to tackle specific agronomic and environmental objectives.

3.3 The RLQ method can determine which varieties and variety mixtures, associated to which bundles of traits, deliver a type of baskets of services

As an alternative to *a priori* classification of varieties, we applied the RLQ method which links three matrices (Fig. 3): here, the (plot \times variety composition) table; the

(varieties \times functional traits) table; and the (plot \times ecosystem services) table, by providing ordination scores to summarize the joint structure among the three tables. This allowed the analysis of the (services \times functional traits) relationships as an output (Fig. 3 and 4). The RLQ analysis revealed a significant relationship between the services and traits ($p < 0.0001$). The first two axes of the RLQ plan extracted 88% of the total variance (75.2 and 12.7% for axes 1 and 2, respectively; Table S3). In particular, the first two axes of the RLQ accounted for 83% of variability of the ordination achieved on services and 94% of the variability of the ordination on the trait table, and Fig. 4A–D is thus a very good representation of the variance in services and traits. In term of services, axis 1 was strongly linked to resistance to lodging and yield (positively), and negatively to grain N content and—to a lesser extent—*Septoria* control and yield stability across sites/managements (Fig. 4D). Axis 2 was mostly linked to crop pest biocontrol and recovery efficiency of the fertilizer-N by the soil-plant system (positively) as well as the specific weight of grains (negatively) (Fig. 4D). Co-inertia analysis showed a significant correlation ($p = 0.001$) between the service matrix and the sown CWM values of traits. In contrast the correlation between the service matrix and trait diversity (sown RaoQ values) was lower. This result is consistent with the mass ratio hypothesis (Grime 1998), well supported by both theory

Fig. 3 Principles of the RLQ analysis relating the composition of wheat fields in term of mixtures of varieties, V , and functional traits, T , to the baskets of ecosystem services, S , they provide. The objective is to analyze the information contained in three tables named respectively R (fields \times services), which corresponds to the measurements of ecosystem services made in the plots; L (fields \times varieties), which provides the variety composition of the plots; and Q (traits \times varieties), which links the varieties to their measured functional traits. The outcomes of the analysis are (i) a test of the significance of the link between the baskets of services and the bundles of traits, and (ii) a RLQ-based classification of the wheat fields/variety mixtures that allows an *a posteriori* analysis of the relationship between bundles of traits and baskets of services.



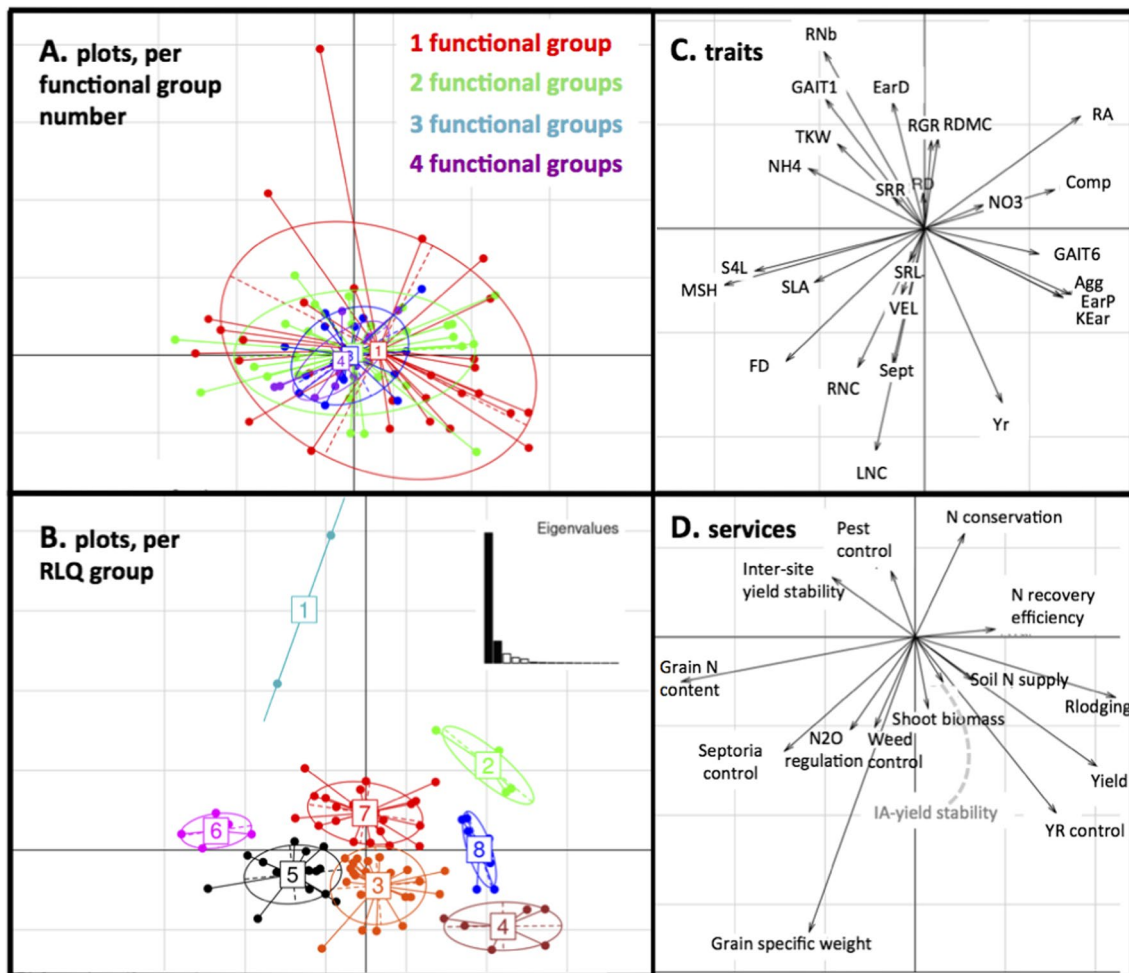


Fig. 4 Results of the RLQ analysis. The panels A and B present the location of the plots (represented by dots) in the plot defined by the first two axes of the RLQ. Panel A distinguishes plots according to the number of functional groups present in each plot, and panel B according to the groups of plots defined by a Ward's hierarchical classification using the RLQ results. Panels C and D present the projections of the vectors of the variety functional traits and the ecosystem services delivered by plots, respectively. The grey dotted line in panel D locates one service vector very close to the intercept. The insert in panel B shows eigenvalues, with first two axes shown in black. The different panels were built with the same two axes. Acronyms for traits are as follows: SRR, shoot:root ratio; RGR, relative growth rate; LNC, flag leaf nitrogen concentration; RNC, root nitrogen concentration; NO_3 , NO_3 -absorption capacity; NH_4 , NH_4 + absorption capacity; RD, mean

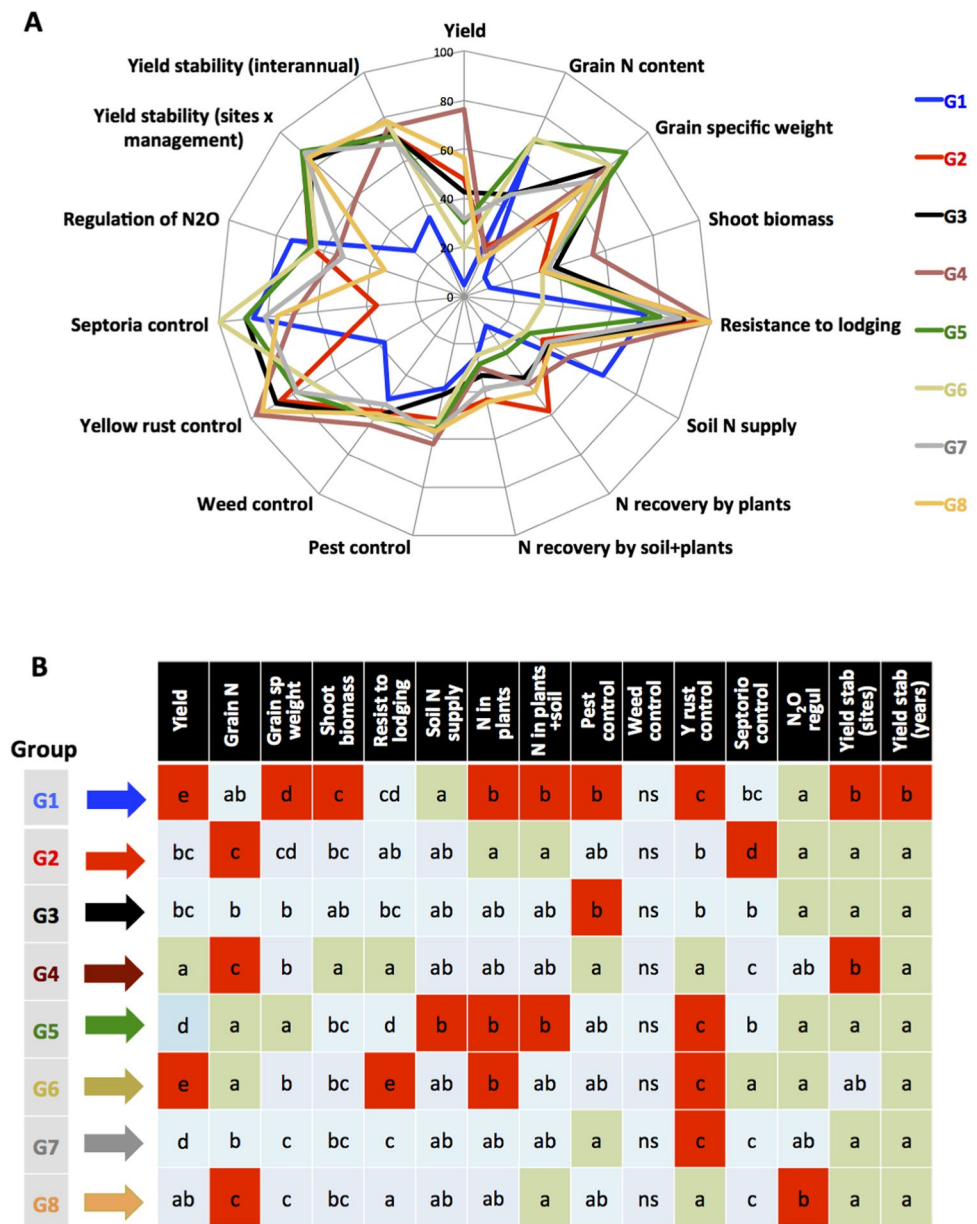
root diameter; RNB, mean root number; SRL, specific root length; RA, mean root angle; RDMC, root dry matter content; LNC, flag leaf nitrogen concentration; S4L, surface of the four superior leaves; VEL, vertical coefficient of extinction of light; MSH, mean height of the main stem shoot; GAIT1, green area index in December; GAIT6, green area index in April; Comp, compensation capacity between 2 seeding densities (ratio of ear density between sowing at 36 and 170 plants m^{-2}); EarD, ear density; Agg, aggressiveness index (ratio between tillering under low density and high N and under high density and low N); Yr, sensitivity to yellow rust; Septo, sensitivity to septoria; FD, flowering date; EarP, mean number of ears per plant; TKW, thousand kernels weight; KEar, mean number of kernels per ear.

and empirical evidence (Sonkoly et al. 2019), which states that ecosystem functions or services are chiefly determined by CWM trait values rather than trait diversity.

Based on a Ward's hierarchical classification using Euclidian distance between mixtures along the first two RLQ axes, the RLQ analysis allowed the identification of 8 groups of wheat plots, including pure stands or variety mixtures (Fig. 4B and Fig. S6). These groups significantly differed from one another for 14 of the 15 services

studied (Fig. 4D and 5), and this classification explained much more variance in the services than the composition of plots in *a priori* defined functional groups ($R^2 > 0.4$ for 7 services; Table 1). For instance, plots from group 4 were associated with high values for several services (Fig. 5), in particular grain yield, shoot production, resistance to lodging, yellow rust control, and inter-annual yield stability. But the plots from this group had the lowest value for grain N concentration and yield stability across sites/

Fig. 5 A Radar chart of the baskets of ecosystem services delivered by the 8 groups of wheat plots (i.e., pure stands and variety mixtures) identified from the RLQ analysis results (one color per RLQ group). For each service in the radar chart, values are normalized from 0 to 100 for the minimum and maximum values observed on individual plots, respectively. **B** For each service (i.e., each column), different letters identify significant difference between plot groups, *G* (ns, non-significant). Cells in green and red indicate the groups of mixtures/monocultures delivering the highest and lowest level of the service, respectively.



managements (Fig. 5). These monocultures and variety mixtures are thus interesting if the main objective is to maximize provisioning services at the expense of other services such as grain N content.

In contrast, group 2 corresponded to plots that maximized the recovery efficiency of fertilizer-N by the agroecosystem, the recovery efficiency of the fertilizer-N by wheat, regulation of soil N₂O production, as well as yield stability (across both years and sites/managements). Plots from group 2 also had poor performance for grain N content and control of Septoria leaf blotch. This group could thus be adequate for farmers accepting moderate decrease in provisioning services

to favor yield stability as well as recovery efficiency of fertilizer-N by the soil-plant system, and hence water quality regulation by their croplands. Group 5 had the highest value for grain quality (both N content and specific weight), regulation of N₂O production, and yield stability (across both years and sites/managements), and the lowest value for soil N supply and recovery efficiency of fertilizer-N by the soil-plant system (Fig. 5). Group 3 corresponded to plots with high value for inter-annual yield stability, high performance for yield stability across sites/managements, and for regulation of N₂O production, with high to intermediate scores for all services except pest control. This group could

thus be recommended to increase the multifunctionality of the wheat fields to farmers ready to accept a slight decrease in grain yield. Group 7 was the only one leading simultaneously to high values for the biocontrol of crop pests and both aspects of yield stability (i.e., across both years and sites/managements), with low values for yellow rust control and intermediate scores for the other services. Finally, plots of group 1 had intermediate or—often—low values for all services, except high but variable values for soil supply of mineral N and regulation of N_2O production (Fig. 5). This group would thus be difficult to recommend to practitioners.

Noticeably, some tradeoffs observed here between baskets of ecosystem services are consistent with tradeoffs between services reported in previous studies. For instance, grain yield is often negatively correlated to grain quality (Bogard et al. 2010), and this tradeoff will likely increase with increased atmospheric CO_2 concentrations (Broberg et al. 2017). Similarly, as observed here, it was reported that grain yield and shoot production were negatively correlated to the regulation of N_2O emissions in managed grasslands (Shi et al. 2019). This shows that at least some tradeoffs observed in our experiment are general and that an approach allowing practitioners to select mixtures delivering different baskets of services is needed. Remarkably, while one group was only composed of varieties cultivated alone (group 1 with only two plots), some groups were composed almost exclusively of variety mixtures. This demonstrates the interest of mixing varieties for providing new baskets of services compared with those delivered by single varieties. More generally, our results show that, even when mixing up to 8 varieties within fields, no group of plots was able to maximize many ecosystem services. For instance, would farmers want to maximize grain yield and shoot production, they should use varieties or mixtures from group 4 that also allow high control of crop pests and yellow rust and high resistance to lodging, but have low grain N content, low Septoria control, and low yield stability between sites and managements. This is consistent with a “Jack-of-all-trades” effect, with species or genotypes characterized by a certain degree of specialization that cannot maximize all functions. In the same vein, results from the study of van der Plas et al. (2016) show that tree species diversity is positively related with multifunctionality only when moderate levels of ecosystem functioning are required, but negatively when very high function levels are desired. Hence, even if variety mixtures offer novel possibilities for the provision of new baskets of services, when farmers have to choose varieties or mixtures, they have to make informed choices in term of a few services to be prioritized at the expense of others, or a broader range of services to be provided at a good—even if not maximal—level.

The RLQ-based classification of wheat plots also allowed the identification of the sown CWM values of functional traits and/or trait diversity values that characterized each group of plots, i.e., each type of basket of services (Fig. 6; Fig. S7). For instance, group 4 that maximized production services was characterized by low relative growth rate, low root number, low ammonium uptake capacity, low leaf surface, high compensation capacity, high ear density, and high sensitivity to yellow rust (Fig. 6). The fact that ammonium uptake capacity by roots was low for this RLQ group is likely due to an overlooked effect of the selection of wheat varieties to maximize provisioning services as already reported by Cantarel et al. (2021). This root trait can also explain why plots from group 4 did not maximize fertilizer-N recovery by wheat plants and by the soil-plant system. It is indeed likely that varieties or mixtures with low ammonium uptake capacity promoted a nitrate-based N cycling rather than ammonium-based N cycling in soil, hence increasing N losses since nitrate is more prone to leaching and denitrification than ammonium (Subbarao and Searchinger 2021). In parallel, the plots from group 4 had low values for trait diversity (Fig. S7). This shows that maximizing provisioning services was best achieved by elite varieties in pure stand, and that very few mixtures can perform as well as these varieties in this perspective. In contrast, the plots of group 2 corresponding to mixtures minimizing N losses from the agroecosystem had intermediate sown CWM values for many traits studied but had relatively high trait diversity values and in particular for some root traits such as root angle and ammonium uptake (Fig. S6). The good performance of plots from group 2 for reducing fertilizer-N losses could be explained by complementarity effects between varieties with diverse root architectures and N uptake capacities. These complementarity effects likely maximized the ability of the mixtures to access a larger soil volume and efficiently uptake diverse soil N forms, thus minimizing N losses as previously observed for grassland (Bessler et al. 2012; Kahmen et al. 2006) and seaweed species (Bracken and Stachowicz 2006). Similarly, the plots from group 3 that had intermediate values for many services had intermediate average and diversity values for most traits. In addition, the plots from group 1 that had the lowest level of yield stability across both years and sites/managements also had the lowest values of trait diversity (Fig. S7). This supports the insurance hypothesis stating that diversity insures ecosystems against declines in their functioning in a fluctuating environment because diverse species or genotypes provide greater guarantees that some will maintain functioning even if others fail (Yachi and Loreau 1999).

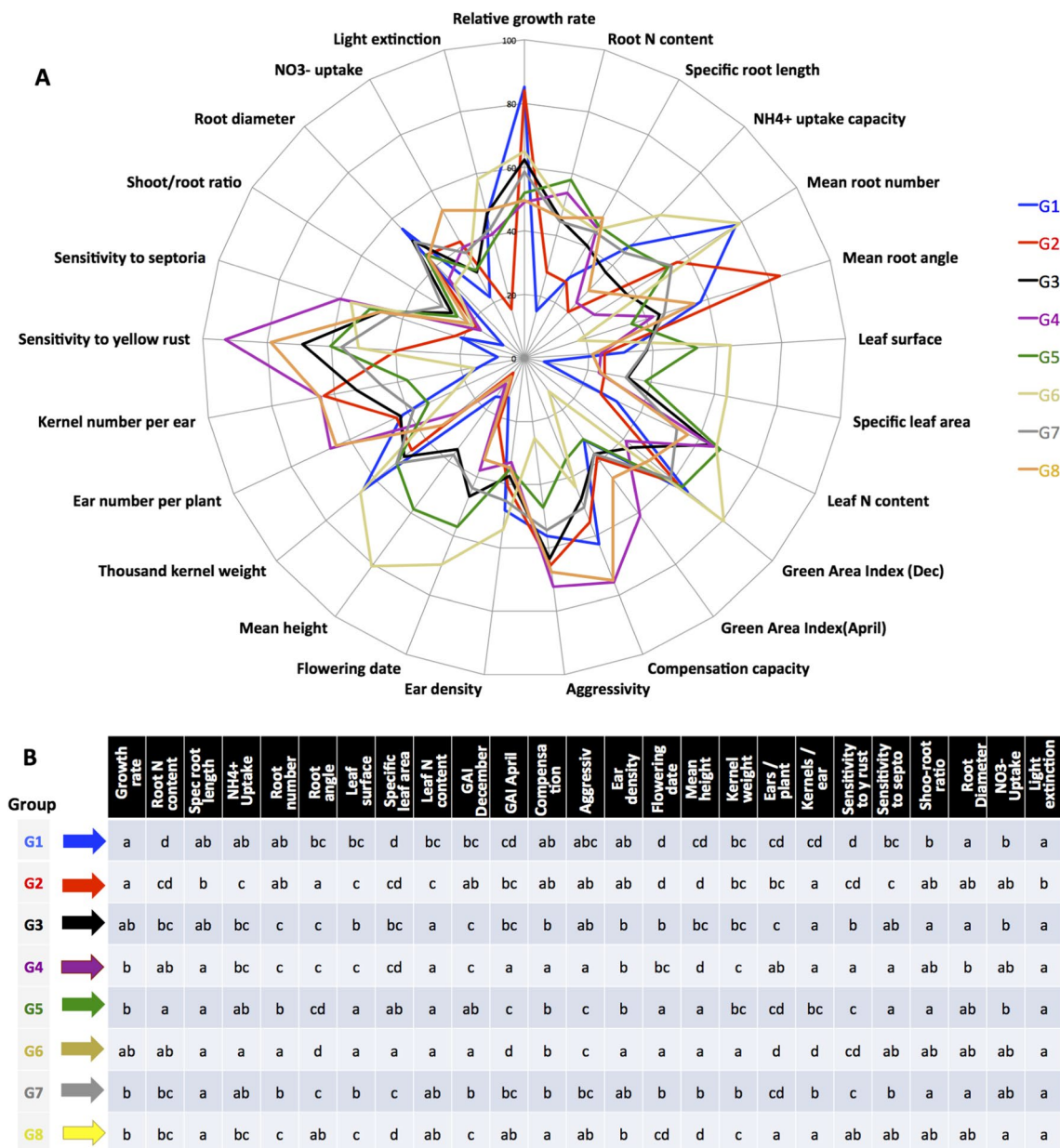


Fig. 6 **A** Radar chart of the bundles of functional traits (here, community-weighted mean values) associated to the 8 groups of wheat plots identified from the RLQ analysis results (one color per RLQ group). For each trait in the radar chart, values are normalized from 0 to 100 for the minimum and maximum values observed on individual

plots, respectively. Results of the same analysis for trait diversity are presented in Fig. S7. **B** For each trait (table column), different letters identify significant difference between plot groups. Group effect on root dry matter content, RDMC, was not significant and is not displayed.

4 Conclusion

We assessed the relationships between intraspecific diversity—in term of variety functional traits—and agroecosystem multifunctionality—in term of 15 (proxies of) provisioning and regulating services—in a field trial with 88 plots exploring a gradient of diversity (bread wheat varieties either in pure stand or in 72 mixtures of 2, 4, or 8 components). Taken together, our results demonstrate that the number of wheat varieties and the

classification of varieties into functional groups (defined on the basis of 26 functional traits) predict only poorly the provision of multiple services and can hardly guide the design of mixtures of varieties. For the first time, we applied the RLQ method to unfold the link between intra-field variety diversity and multifunctionality, and showed that this method allows relating particular baskets of services to specific bundles of variety traits (in terms of mean values—and to a lesser extent variance—of traits in the wheat variety mixtures). For instance, our results

show that farmers can decide to (1) maximize grain yield and shoot production by using varieties/variety mixtures that also maximize resistance to lodging, but at the expense of, e.g., grain N content, recovery efficiency of the fertilizer-N by soil+plants, and yield stability across sites/managements. In contrast, farmers may decide to (2) use mixtures with less optimal—but still high—yield values to optimize yield stability across years and sites/managements while also optimizing the recovery efficiency of the fertilizer-N by the soil+plant system; or to (3) maximize grain quality (both N content and specific weight) as well as yield stability at the expense of yield and shoot production. Moreover, our results point to the specific suite of functional traits that is associated to each type of basket of services.

The approach presented here might generate actionable knowledge fitting expectations from practitioners since it provides guidance for the design of particular variety mixtures according to the baskets of services to be delivered. It could be argued that there would be a hierarchy between some services, with some regulating services being advantageous and valued mostly through their contribution to yield. Actually, although this hierarchy exists, we advocate that each service may have a value on its own. For instance, regulating services such as the limitation of the spreading of diseases and the bio-control of crop pests and weeds tackle the goal of decreasing pesticide inputs (even accepting a slight decrease of yield). New financial mechanisms recently proposed in Europe that allow to pay farmers who decrease their use of pesticides by using variety mixtures (French Government 2018) are a good example of how “regulating” services can be valued on their own. In the future, it may be envisaged that other services such as the regulation of greenhouse gas production by agroecosystems may also be valued. Our approach allows a comprehensive analysis of the baskets of services, considering that each service may be valued to some extent on its own.

We used here the same weight for all the 15 services, but this approach is flexible because the RLQ analysis can be run after selecting only some services particularly targeted by users and/or after weighing differently the services considered depending on perceptions, economics, or agri-political conditions (e.g., giving higher importance to some services such as yield, yield stability and reduction of losses from fertilizer-N, while minimizing the importance of other services). Obviously, each set of service weights will lead to specific RLQ outcomes. The choice of the set(s) of service weights to be used could be informed by participatory approaches with farmers, allowing them to propose different scenarios for selecting and ranking services, according to farm specificities and to current or future environmental and socio-economic conditions as well as local/cultural specificities.

A limitation of the RLQ approach is that the clustering of plots/baskets of services is made *ex post*, and that new experiments could be needed if a new set of varieties is to be tested. This can be a problem since the type of experiment presented

here is laborious and expensive. But the RLQ approach does not only establish relationships between groups of varieties/variety mixtures and baskets of services: it also establishes links between the baskets of services delivered and the bundles of traits that characterize the varieties/mixtures. The predictive power of the approach and usefulness for practice lies in the latter link between baskets of services and bundles of traits. Would these links be robust enough (which remains to be tested), users could select bundles of traits in their mixtures independently from the very varieties employed.

To fully exploit our approach to help farmers and actors from the seed sector designing mixtures of varieties, we envision the following steps: (1) New experiments on small plots in experimental farms should be implemented to test the robustness of our RLQ results (both the mitigation of tradeoffs between services by variety mixtures and the links between baskets of services and bundles of traits). The experiments should also test the interactions with crop management (e.g., comparing organic vs. non-organic agriculture). (2) Workshops with farmers, extension services, breeders, and actors from the seed sector should be organized to determine priorities in the services to be delivered, related to scenarios of, e.g., climate change, costs of chemical inputs used in agriculture, taxes linked to disservices, and payments for regulating services. The data could then be analyzed using the RLQ approach to determine the sensitivity of the traits-services relationships to the selection and weighing of the services (our dataset and RLQ code, made freely available, can allow readers to do so). (3) These new results should be used to design variety mixtures according to farmers’ objectives, and these mixtures and the services they provide should be tested on farm.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s13593-023-00924-6>.

Acknowledgements We acknowledge the help and inputs of all the members of the Wheatamix consortium (the Wheatamix consortium is composed of: V. Allard, B. Andrieu, S. Barot, J. Borg, A. Brunet, A.A.M. Cantarel, C. Cervek, F. Coleno, Y. De Olivera, C. De Vallavieille-Pope, D. Descoureaux, F. Dubs, J. Enjalbert, E. Forst, N. Galic, A. Gauffreteau, J. Gervais, I. Goldringer, M. Hannachi, G. Houivet, M. Hugoni, M.H. Jeuffroy, S. Jouanne, C. Kerbirou, P. Labarthe, J.-C. Lata, X. Le Roux, C. Lecarpentier, M. Leconte, S. Lemarié, L. Lejars, B. Lemain, F. Leny, P. Lusley, C. Montagnier, A. Niboyet, B. Omon, R. Perronne, F. Poly, S. Piaud, T. Pommier, E. Porcher, S. Saint-Jean, S. Salmon, D. Tropée, T. Vidal). We thank the AME research platform in Lyon (UMR INRAE 1418) for resin bag analyses, and the IRD-Sorbonne Université *Allysés* analytical platform for N and ¹⁵N analyses.

Authors' contributions Conceptualization, X.L.R. and F.D., with S.B., J.E., and E.P.; methodology and investigation, F.D., X.L.R., J.E., S.B., E.P., V.A., C.P., A.G., A.N., T.P., S.S.-J., T.V.; data analysis, F.D. and X.L.R. with help from all authors; writing—original draft, X.L.R. and F.D.; writing—review and editing, all authors.

Funding This work was funded by ANR through the Agrobiosphere program (project Wheatamix ANR-13-AGRO-0008) and by the INRAE meta-program ECOSERV (project SolFaMi).

Data availability Details on the field trial and methodology employed can be found in the method paper on HAL archive (<https://hal.archives-ouvertes.fr/hal-01843564>). The dataset on agroecosystem services generated and analyzed during the current study is available on the INRAE folder of French data-verse: <https://entrepot.recherche.data.gouv.fr/dataverse/inrae>. The datasets on variety traits are available through Cantarel et al. (2021; <https://academic.oup.com/jxb/article/72/4/1166/5932789?login=true>).

Code availability (software application or custom code) The code of the RLQ model and data used for this study are available on the INRAE folder of French data-verse: <https://entrepot.recherche.data.gouv.fr/dataverse/inrae> (<https://doi.org/10.57745/FDNQ3V>)

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allan E, Weisser WW, Fischer M, Schulze ED, Weigelt A, Roscher C, Baade J, Barnard R, Beßler H, Buchmann N, Buscot F, Ebeling A, Eisenhauer N, Engels C, Fergus A, Gleixner G, Gubsch M, Habekost M, Halle S, Klein AM, König S, Kowalski E, Kreutziger Y, Kertscher I, Kummer V, Lange M, Lauterbach D, Le Roux X, Marquard E, Migunova VD, Milcu A, Mwangi P, Niklaus P, Oelmann Y, Peterman J, Poly F, Rottstock T, Rosenkranz S, Sabais A, Schercher C, Scherer-Lorezen M, Scheu S, Schmitz M, Schumacher J, Soussana JF, Steinbeiss S, Temperton V, Tschantke T, Voigt W, Wilcke W, Wirth C, Schmid B (2013) A comparison of the strength of biodiversity effects across multiple functions. *Oecologia* 173:223–237. <https://doi.org/10.1007/s00442-012-2589-0>
- Altieri MA (1999) The ecological role of biodiversity in agroecosystems. *Agric Ecosyst Environ* 74:19–31. [https://doi.org/10.1016/S0167-8809\(99\)00028-6](https://doi.org/10.1016/S0167-8809(99)00028-6)
- Barot S, Allard V, Cantarel AAM, Enjalbert J, Gauffreteau A, Goldringer I, Lata JC, Le Roux X, Niboyet A, Porcher E (2017) Designing mixtures of varieties for multifunctional agriculture with the help of ecology: a review. *Agron Sustain Develop* 37:13. <https://doi.org/10.1007/s13593-017-0418->
- Beillouin D, Ben-Ari T, Malézieux E, Seufert V, Makowski D (2021) Positive but variable effects of crop diversification on biodiversity and ecosystem services. *Global Change Biol* 27:4697–4710. <https://doi.org/10.1111/2020.09.30.320309>
- Bessler H, Oelmann Y, Roscher C, Buchmann N, Scherer-Lorezen M, Schulze ED, Temperton VM, Wilcke W, Engels C (2012) Nitrogen uptake by grassland communities: contribution of N₂ fixation, facilitation, complementarity, and species dominance. *Plant Soil* 358:301–322. <https://doi.org/10.1007/s11104-012-1181-z>
- Bogard M, Allard V, Brancourt-Hulmel M, Heumez E, Machet JM, Jeuffroy MH, Gate P, Martre P, Le Gouis J (2010) Deviation from the grain protein concentration–grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. *J Exp Bot* 61:4303–4312. <https://doi.org/10.1093/jxb/erq238>
- Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Henry M, Stevens H, Jada-Simone S (2009) White generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol* 24:127–135. <https://doi.org/10.1016/j.tree.2008.10.008>
- Bonnin I, Bonneuil C, Goffaux R, Goldringer I (2014) Explaining the decrease in the genetic diversity of wheat in France over the 20th century. *Agric Ecosyst Environ* 195:183–192. <https://doi.org/10.1016/j.agee.2014.06.003>
- Borg J, Kiaer LP, Lecarpentier C, Goldringer I, Gauffreteau A, Saint Jean S, Barot S, Enjalbert J (2018) Unfolding the potential of wheat cultivar mixtures: a meta-analysis perspective and identification of knowledge gaps. *Field Crops Res* 221:298–313. <https://doi.org/10.1016/j.fcr.2017.09.006>
- Botta-Dukát Z (2005) Rao's quadratic entropy as a measure of functional diversity based on multiple traits. *J Veget Sci* 16:533–540. <https://doi.org/10.1111/j.1654-1103.2005.tb02393.x>
- Bracken MES, Stachowicz JJ (2006) Seaweed diversity enhances nitrogen uptake via complementary use of nitrate and ammonium. *Ecology* 87:2397–2403. [https://doi.org/10.1890/0012-9658\(2006\)87\[2397:SDENUV\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[2397:SDENUV]2.0.CO;2)
- Brisson N, Gate P, Gouache D, Charmet G, Oury FX, Huard F (2010) Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. *Field Crop Res* 119:201–212. <https://doi.org/10.1016/j.fcr.2010.07.012>
- Broberg CM, Högy P, Pleijel H (2017) CO₂-induced changes in wheat grain composition: meta-analysis and response functions. *Agronomy* 7:32. <https://doi.org/10.3390/agronomy7020032>
- Cantarel AAM, Allard V, Andrieu B, Barot S, Enjalbert J, Gervais J, Goldringer I, Pommier T, Saint Jean S, de Vallavielle-Pope C, Le Roux X (2021) Plant functional trait variability and trait syndromes among wheat (*Triticum aestivum*) varieties: the footprint of artificial selection. *J Exp Bot* 72:1166–1180. <https://doi.org/10.1093/jxb/eraa491>
- Cardinale BJ, Palmer MA, Collins SL (2003) Species diversity enhances ecosystem functioning through interspecific facilitation. *Nature* 415:426–429. <https://doi.org/10.1038/415426a>
- Chateil C, Goldringer I, Tarallo L, Kerbirou C, Le Viol I, Ponge JF, Salmon S, Gachet S, Porcher E (2013) Crop genetic diversity benefits farmland biodiversity in cultivated fields. *Agric Ecosyst Environ* 171:25–32. <https://doi.org/10.1016/j.agee.2013.03.004>
- Craine JM, Tilman D, Wedin D, Reich P, Tjolkner M, Knops J (2002) Functional traits, productivity and effects on nitrogen cycling of 33 grassland species. *Funct Ecol* 16:563–574. <https://doi.org/10.1046/j.1365-2435.2002.00660.x>
- Diaz S, Cabido M (2001) Vive la différence: plant functional diversity matters to ecosystem process. *Trends Ecol Evol* 16:646–655. [https://doi.org/10.1016/S0169-5347\(01\)02283-2](https://doi.org/10.1016/S0169-5347(01)02283-2)
- Dolédéc S, Chessel D, ter Braak CJF, Champely S (1996) Matching species traits to environmental variables: a new three-table ordination method. *Environ Ecol Stat* 3:143–166. <https://doi.org/10.1007/BF02427859>
- Dray S, Legendre P (2008) Testing the species traits–environment relationships: the fourth-corner problem revisited. *Ecology* 89:3400–3412. <https://doi.org/10.1890/08-0349.1>

- Dubs F, Le Roux X, Allard V, Andrieu B, Barot S, Cantarel A, De Vallavieille-Pope C, Gauffreteau A, Goldringer I, Montagnier C, Pommier T, Porcher E, Saint-Jean S, Borg J, Bourdet-Massein S, Carmignac D, Duclouet A, Forst E, Galic N, Gerard L, Hugoni M, Hure A, Larue A, Lata JC, Lecarpentier C, Leconte M, Le Saux E, Le Viol I, L'Hôte P, Lusley P, Mouchet M, Niboyet A, Perronne R, Pichot E, Pin S, Salmon S, Tropic D, Vergnes A, Vidal T, Enjalbert J (2018) An experimental design to test the effect of wheat variety mixtures on biodiversity and ecosystem services. <https://hal.archives-ouvertes.fr/hal-01843564>
- Eggermont H, Balian E, Azevedo J, Beumer V, Brodin T, Claudet J, Fady B, Grube M, Keune H, Lamarque P, Reuter K, Smith M, van Ham C, Weisser WW, Le Roux X (2015) Nature-based solutions: new influence for environmental management and research in Europe. *Gaia* 24:243–248. <https://doi.org/10.14512/gaia.24.4.9>
- Ellis JG, Lagudah ES, Spielmeier W, Dodds PN (2014) The past, present and future of breeding rust resistant wheat. *Front Plant Sci* 5: <https://doi.org/10.3389/fpls.2014.00641>
- Finckh MR, Gacek ES, Goyeau H, Lannou C, Merz U, Mundt CC, Munk L, Nadziak J, Newton AC, de Vallavieille-Pope C, Wolf MS (2000) Cereal variety and species mixtures in practice, with emphasis on disease resistance. *Agronomie* 20:813–837. <https://doi.org/10.1051/agro:2000177>
- Foley JA, Ramankutty N, Brauman KA, Cassidy ES, Gerber JS, Johnston M, Mueller ND, O'Connell C, Ray DK, West PC, Balzer C, Bennett EM, Carpenter SR, Hill J, Monfreda C, Polasky S, Rockstrom J, Sheehan J, Siebert S, Tilman D, Zaks DPM (2011) Solutions for a cultivated planet. *Nature* 478:337–342. <https://doi.org/10.1038/nature10452>
- Fox J, Weisberg S (2019) An R companion to applied regression, Third edition. Sage, Thousand Oaks CA. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>
- French Government (2018) Ecophyto II plan. French Government Report, 68 pp. https://food.ec.europa.eu/system/files/2020-02/pesticides_sup_nap_fra-ecophyto-2plus_plan_en.pdf
- Futuyma DJ, Moreno G (1988) The evolution of ecological specialization. *Annu Rev Ecol Syst* 19:207–233
- Gaba S, Lescourret F, Boudsocq S, Enjalbert J, Hinsinger P, Journet EP, Navas ML, Wery J, Louarn G, Malézieux E, Pelzer E, Prudent M, Ozier-Lafontaine H (2015) Multiple cropping systems as drivers for providing multiple ecosystem services: from concepts to design. *Agro Sust Dev* 35:607–623. <https://doi.org/10.1007/s13593-014-0272-z>
- Gamfeldt L, Snäll T, Bagchi R, Jonsson M, Gustafsson L, Kjellander P, Ruiz-Jaen MC, Froberg M, Stendahl J, Philipson CD, Mikusinski G, Andersson E, Westerlund B, Andren H, Moberg F, Moen J, Bengtsson J (2013) Higher levels of multiple ecosystem services are found in forests with more tree species. *Nat Commun* 4:1340. <https://doi.org/10.1038/ncomms2328>
- Grime JP (1998) Benefits of plant diversity to ecosystems: immediate, filter and founder effects. *J Ecol* 86:902–910. <https://doi.org/10.1046/j.1365-2745.1998.00306.x>
- Gross K, Cardinale BJ, Fox JW, Gonzalez A, Loreau M, Polley HW, Reich PB, van Ruijven J (2014) Species richness and the temporal stability of biomass production: a new analysis of recent biodiversity experiments. *Am Nat* 183:1–12. <https://doi.org/10.1086/673915>
- Hajjar R, Jarvis DI, Gemmill-Herren B (2008) The utility of crop genetic diversity in maintaining ecosystem services. *Agric Ecosyst Environ* 123:261–270. <https://doi.org/10.1016/j.agee.2007.08.003>
- Hector A, Bagchi R (2007) Biodiversity and ecosystem multifunctionality. *Nature* 448:188–190. <https://doi.org/10.1038/nature05947>
- Howden SM, Soussana JF, Tubiello FN, Chhetri N, Dunlop M, Meinke H (2007) Adapting agriculture to climate change. *Proc Natl Acad Sci USA* 104:19691–19696. <https://doi.org/10.1073/pnas.0701890104>
- Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M (2008) Ecological consequences of genetic diversity. *Ecol Lett* 11:609–623. <https://doi.org/10.1111/j.1461-0248.2008.01179.x>
- Johnson MTJ, Lajeunesse MJ, Agrawal AA (2006) Additive and interactive effects of plant genotypic diversity on arthropod communities and plant fitness. *Ecol Lett* 9:24–34. <https://doi.org/10.1111/j.1461-0248.2005.00833.x>
- Kahmen A, Renker C, Unsicker SB, Buchmann N (2006) Niche complementarity for nitrogen: an explanation for the biodiversity and ecosystem functioning relationship? *Ecology* 87:1244–1255. [https://doi.org/10.1890/0012-9658\(2006\)87](https://doi.org/10.1890/0012-9658(2006)87)
- Kiær LP, Skovgaard IM, Østergård H (2009) Grain yield increase in cereal variety mixtures: a meta-analysis of field trials. *Field Crop Res* 114:361–373. <https://doi.org/10.1016/j.fcr.2009.09.006>
- Lavorel S, Garnier E (2002) Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Func Ecol* 16:545–556. <https://doi.org/10.1046/j.1365-2435.2002.00664.x>
- Le Roux X, Schmid B, Poly F, Barnard RL, Niklaus PA, Guillaumaud N, Habekost M, Oelmann Y, Philippot L, Salles J, Schloter M, Steinbeiss S, Weigelt A (2013) Soil environmental conditions and buildup of microbial communities mediate the effect of grassland plant diversity on nitrifying and denitrifying enzyme activities. *PLOS One* 8:e61069. <https://doi.org/10.1371/journal.pone.0061069>
- Litrico I, Violle C (2015) Diversity in plant breeding: a new conceptual framework. *Trends Plant Sci* 20:604–613. <https://doi.org/10.1016/j.tplants.2015.07.007>
- Lopez CG, Mundt CC (2000) Using mixing ability analysis from two-way cultivar mixtures to predict the performance of cultivars in complex mixtures. *Field Crop Res* 68:121–132. [https://doi.org/10.1016/S0378-4290\(00\)00114-3](https://doi.org/10.1016/S0378-4290(00)00114-3)
- Loreau M, Hector A (2001) Partitioning selection and complementarity in biodiversity experiments. *Nature* 412:72–76. <https://doi.org/10.1038/35083573>
- Malézieux E (2011) Designing cropping systems from nature. *Agron Sustain Dev* 32:15–29. <https://doi.org/10.1007/s13593-011-0027-z>
- Manley M, Engelbrecht ML, Williams PC, Kidd M (2009) Assessment of variance in the measurement of hectolitre mass of wheat, using equipment from different grain producing and exporting countries. *Biosyst Engin* 103:176–186. <https://doi.org/10.1016/j.biosystemseng.2009.02.018>
- Martin AR, Isaac ME, Manning P (2015) Plant functional traits in agroecosystems: a blueprint for research. *J Appl Ecol* 52:1425–1435. <https://doi.org/10.1111/1365-2664.12526>
- Mille B, Belhaj Fraj B, Monod H, de Vallavieille-Pope C (2006) Assessing four-way mixtures of winter wheat cultivars from the performances of their two-way and individual components. *Eur J Plant Pathol* 114:163–173. <https://doi.org/10.1007/s10658-005-4036-0>
- Newton AC, Swanston JS, Guy DC, Ellis RP (1998) The effect of cultivar mixtures on malting quality in winter and spring barley. *J Inst Brew* 104:41–45. <https://doi.org/10.1002/j.2050-0416.1998.tb00973.x>
- Østergård H, Finckh MR, Fontaine L, Goldringer I, Hoad SP, Kristensen K, Lammerts van Bueren ET, Mascher F, Munk L, Wolfe MS (2009) Time for a shift in crop production: embracing complexity through diversity at all levels. *J Sci Food Agric* 89:1439–1445. <https://doi.org/10.1002/jsfa.3615>
- Östman O (2004) The relative effects of natural enemy abundance and alternative prey abundance on aphid predation rates. *Biol Control* 30:281–287. <https://doi.org/10.1016/j.biocontrol.2004.01.015>

- Patra AK, Abbadié L, Clays-Josserand A, Degrange V, Grayston S, Loiseau P, Louault F, Mahmood S, Nazaret S, Philippot L, Poly F, Prosser JI, Richaume A, Le Roux X (2005) Effects of grazing on microbial functional groups involved in soil N dynamics. *Ecol Monogr* 75:65–80. <https://doi.org/10.1890/03-0837>
- Pe'er G, Dicks LV, Visconti P, Arlettaz R, Báldi A, Benton TG, Collins S, Dieterich M, Gregory RD, Hartig F, Henle K, Hobson PR, Kleijn D, Neumann RK, Robijns T, Schmidt J, Schwartz A, Sutherland WJ, Turbé A, Wulf F, Scott AV (2014) EU agricultural reform fails on biodiversity. *Science* 344(1090):1092. <https://doi.org/10.1126/science.1253425>
- Petchey OL (2004) On the statistical significance of functional diversity effects. *Funct Ecol* 18:297–303. <https://doi.org/10.1111/j.0269-8463.2004.00852.x>
- Petchey OL, Gaston KJ (2002) Functional diversity (FD), species richness and community composition. *Ecol Lett* 5:402–411. <https://doi.org/10.1046/j.1461-0248.2002.00339.x>
- Peterson BG, Carl P, Boudt K, Bennett R, Ulrich J, Zivot E, Cornilly D, Hung E, Lestel M, Balkissoon K, Wuertz D, Christidis AA, Martin RD, Zhou ZZ, Shea JM (2020) PerformanceAnalytics R package, CRAN 2.0.4 <https://cran.r-project.org/web/packages/PerformanceAnalytics/PerformanceAnalytics.pdf>
- Reich PB, Tilman D, Naeem S, Ellsworth DS, Knops J, Craine J, Wedin D, Trost J (2004) Species and functional group diversity independently influence biomass accumulation and its response to CO₂ and N. *Proc Natl Acad Sci USA* 101:10101–10106. <https://doi.org/10.1073/pnas.0306602101>
- Renting H, Rossing WA, Groot JC, Van der Ploeg JD, Laurent C, Perraud D, Stobbelaar DJ, Van Ittersum MK (2009) Exploring multifunctional agriculture. A review of conceptual approaches and prospects for an integrative transitional framework. *J Environ Manage* 90:112–123. <https://doi.org/10.1016/j.jenvman.2008.11.014>
- Robson TM, Lavorel S, Clement JC, Le Roux X (2007) Neglect of mowing and fertilisation leads to slower nitrogen cycling in sub-alpine grasslands. *Soil Biol Biochem* 39:930–941. <https://doi.org/10.1016/j.soilbio.2006.11.004>
- Sarandon SJ, Sarandon R (1995) Mixture of cultivars: pilot field trial of an ecological alternative to improve production or quality of wheat (*Triticum aestivum*). *J Appl Ecol* 32:288–294. <https://doi.org/10.2307/2405096>
- Shah L, Yahya M, Shah SMA, Nadeem M, Ali A, Ali A, Wang J, Waheed Riaz M, Rehman S, Wu W, Muhammad Khan R, Abbas A, Riaz A, Bakr Anis G, Si H, Jiang H, Ma C (2019) Improving lodging resistance: using wheat and rice as classical examples. *Int J Mol Sci* 20:4211. <https://doi.org/10.3390/ijms20174211>
- Shi Y, Wang J, Le Roux X, Ao Y, Gao S, Zhang J, Knops JMH, Mu C (2019) Tradeoffs and synergies between seed yield, forage yield and N-related disservices for a semi-arid perennial grassland under different nitrogen fertilization strategies. *Biol Fertil Soil* 55:497–509. <https://doi.org/10.1007/s00374-019-01367-6>
- Sirami C, Gross N, Borem Baillod A, Bertrand C, Carrié R, Hass A, Henckel L, Miguet P, Vuillot C, Alignier A, Girard J, Batáry P, Clough Y, Violle C, Giral D, Bota G, Badenhauer I, Lefebvre G, Gauffre B, Vialatte A, Calatayud F, Gil-Tena A, Tischendorf L, Mitchell S, Lindsay K, Georges R, Hilaire S, Recasens J, Oriol Solé-Senan X, Robleño I, Bosch J, Antonio Barrientos J, Ricarte A, Marcos-García MA, Miñano J, Mathevet R, Gibon A, Baudry J, Balent G, Poulin B, Burel F, Tschardt T, Bretagnolle V, Siriwardena G, Ouin A, Brotons L, Martin JL, Fahrig L (2019) Increasing crop heterogeneity enhances multitrophic diversity across agricultural regions. *Proc Natl Acad Sci USA* 116:16442–16447. <https://doi.org/10.1073/pnas.1906419116>
- Sonkoly J, Kelemen A, Valkó O, Deák B, Kiss R, Tóth K, Migléc T, Tóthmérész B, Török P (2019) Both mass ratio effects and community diversity drive biomass production in a grassland experiment. *Sci Rep* 9:1848. <https://doi.org/10.1038/s41598-018-37190-6>
- Subbarao GV, Searchinger TD (2021) A “more ammonium solution” to mitigate nitrogen pollution and boost crop yields. *Proc Natl Acad Sci USA* 118:e2107576118. <https://doi.org/10.1073/pnas.2107576118>
- Ter Braak CJF, Cormont A, Dray S (2012) Improved testing of species traits–environment relationships in the fourth-corner problem. *Ecology* 93:1525–1526. <https://doi.org/10.1890/12-0126.1>
- Tilman D, Knops J, Wedin D, Reich P, Ritchie M, Siemann E (1997) The influence of functional diversity and composition on ecosystem processes. *Science* 277:1300–1302. <https://doi.org/10.1126/science.277.5330.1300>
- Tilman D, Cassman KG, Matson PA, Naylor R, Polasky S (2002) Agricultural sustainability and intensive production practices. *Nature* 418:671–677. <https://doi.org/10.1038/nature01014>
- Tooker JF, Frank SD, Steffan-Dewenter I (2012) Genotypically diverse cultivar mixtures for insect pest management and increased crop yields. *J Appl Ecol* 49:974–985. <https://doi.org/10.1111/j.1365-2664.2012.02173.x>
- van der Plas F, Manning P, Allan E, Scherer-Lorenzen M, Verheyen K, Wirth C, Zavala MA, Hector A, Ampoorter E, Baeten L, Barbaro L, Bauhus J, Benavides R, Benneter A, Berthold F, Bonal D, Bouriaud O, Bruehlheide H, Bussotti F, Carnol M, Castagneyrol B, Charbonnier Y, Coomes D, Coppi A, Bastias CC, Dawud SM, De Wandler H, Domisch T, Finer L, Gessler A, Granier A, Grossiord C, Guyot V, Hattenschwiler S, Jactel H, Jaroszewicz B, Joly FX, Jucker T, Koricheva J, Milligan H, Muller S, Muys B, Nguyen D, Pollastrini M, Raulund-Rasmussen K, Selvi F, Stenlid J, Valladares F, Vesterdal L, Zielinski D, Fischer M (2016) Jack-of-all-trades effects drive biodiversity–ecosystem multifunctionality relationships in European forests. *Nat Commun* 7:11109. <https://doi.org/10.1038/ncomms11109>
- van Etten J, de Sousa K, Aguilar A, Barrios M, Coto A, Dell'Acqua M, Fadda C, Gebrehawaryat Y, van de Gevel J, Gupta A, Kiros AY, Madriz B MP, Mengistu DK, Mercado L, Nurhisn Mohammed J, Paliwal A, Pè ME, Quirós CF, Rosas JC, Sharma N, Singh SS, Solanki IS, Steinke J (2019) Crop variety management for climate adaptation supported by citizen science. *Proc Natl Acad Sci USA* 116:4194–4199. <https://doi.org/10.1073/pnas.1813720116>
- Vidal T, Saint-Jean S, Lusley P, Leconte M, Kréma SB, Boixel AL, de Valaveille-Pope C (2020) Cultivar mixture effects on disease and yield remain despite diversity in wheat height and earliness. *Plant Pathol* 69:1148–1160. <https://doi.org/10.1111/ppa.13200>
- Weisser WW, Roscher C, Meyer ST, Ebeling A, Luo G, Beßler H, Barnard R, Buchmann N, Buscot F, Engels C, Fischer C, Fischer M, Gessler A, Gleixner G, Halle S, Hildebrandt A, Hillebrand H, de Kroon H, Lange M, Leimer S, Le Roux X, Milcu A, Mommer L, Niklaus N, Oelmann Y, Proulx R, Scherber C, Scherer-Lorenzen M, Scheu S, Tschardt T, Wachendorf M, Wagg C, Weigelt A, Wilcke W, Schulze ED, Schmid B, Eisenhauer N (2017) Biodiversity effects on ecosystem functioning in a 15-year grassland experiment: patterns, mechanisms, and open questions. *Basic Appl Ecol* 23:1–73. <https://doi.org/10.1016/j.baec.2017.06.002>
- Wright JP, Naeem S, Hector A, Lehman C, Reich PB, Schmid B, Tilman D (2006) Conventional functional classification schemes underestimate the relationship with ecosystem functioning. *Ecol Lett* 9:11–120. <https://doi.org/10.1111/j.1461-0248.2005.00850.x>
- Wuest SE, Peter R, Niklaus PA (2021) Ecological and evolutionary approaches to improving crop variety mixtures. *Nat Ecol Evol* 5:1068–1077. <https://doi.org/10.1038/s41559-021-01497-x>
- Yachi S, Loreau M (1999) Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc Natl Acad Sci USA* 96:1463–1468. <https://doi.org/10.1073/pnas.96.4.1463>
- Zavaleta ES, Pasari JR, Hulvey KB, Tilman D (2010) Sustaining multiple ecosystem functions in grassland communities requires higher biodiversity. *Proc Natl Acad Sci USA* 107:1443–1446. <https://doi.org/10.1073/pnas.0906829107>
- Zeller SL, Kalinina O, Flynn DFB, Schmid B (2012) Mixtures of genetically modified wheat lines outperform monocultures. *Ecol Appl* 22:1817–1826. <https://doi.org/10.2307/41722896>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.