

# **Competition, Cooperation and Complementarity in Soybean**

---

**Dissertation**

**zur**

**Erlangung der naturwissenschaftlichen Doktorwürde**

**(Dr. sc. nat.)**

**vorgelegt der**

**Mathematisch-naturwissenschaftlichen Fakultät**

**der**

**Universität Zürich**

**von**

Emanuel Balthasar Kopp

**von**

Luzern LU

## **Promotionskommission**

Prof. Dr. Pascal Niklaus (Vorsitz)

Dr. Samuel Wüst (Leitung der Dissertation)

Prof. Dr. Niels Anten

Prof. Dr. Jordi Bascompte

**Zürich, 2025**

**La fantasia è un posto dove ci piove dentro.**

Italo Calvino, *Lezioni Americane*

# Contents

1. Summary .....	7
2. General introduction .....	8
2.1. Agriculture at a crossroads .....	8
2.2. The ecological perspective .....	10
2.3. Soybean: past, present and future importance .....	15
2.4. Thesis outlook .....	16
3. Simultaneous above- and belowground competition leads to a tragedy of the commons in soybean .....	18
3.1. Introduction .....	20
3.2. Materials and methods .....	22
3.3. Results .....	25
3.4. Discussion .....	29
3.5. Supplementary materials .....	32
4. Cooperative soybean genotypes mitigate drought stress through water sharing .....	33
4.1. Introduction .....	35
4.2. Materials and methods .....	37
4.3. Results .....	42
4.4. Discussion .....	44
4.5. Supplementary materials .....	50
5. Ecological principles to guide the development of crop variety mixtures .....	56
5.1. Introduction .....	58
5.2. From monocultures to mixtures: known benefits and future challenges .....	58
5.3. Box: Pathogen suppression – the ultimate ecosystem service? .....	60
5.4. Challenges in breeding for variety mixtures .....	62
5.5. Ecological theory related to biodiversity benefits .....	63
5.6. Box: Trait-based methods in ecology .....	64
5.7. Ecological principles applied to agriculture .....	67
5.8. Untapped potential for complementarity .....	68
5.9. Novel ideas to tackle an old problem .....	69
5.10. Conclusions .....	73
6. Novel approaches to identify functionally complementary genotypes promoting overyielding in soybean variety mixtures .....	75
6.1. Introduction .....	77
6.2. Materials and methods .....	78
6.3. Results .....	84
6.4. Discussion .....	86
6.5. Supplementary materials .....	92
6.6. Appendix: Error propagation .....	93
7. General discussion .....	98
7.1. From individual plants to genotypes .....	99
7.2. Genotypic interactions: cooperation and competition .....	100
7.3. From genotypes to communities .....	101
7.4. Ideotypes and ideotype breeding .....	102
7.5. Limitations .....	104
7.6. Concluding remarks and future prospects .....	105

8. Acknowledgments .....	107
8.1. Funding .....	108
9. Bibliography .....	109

## List of Figures

Figure 2.1: Agriculture is a multivariate optimization problem .....	10
Figure 2.2: Density dependency of hybrid maize breeding progress .....	13
Figure 3.1: Experimental design .....	22
Figure 3.2: Root biomass and shoot biomass to height ratio .....	25
Figure 3.3: Root to shoot ratio .....	28
Figure 3.4: Biomass and height .....	28
Figure 3.5: Supplementary material: picture of experiment .....	32
Figure 3.6: Supplementary material: growth and relative chlorophyll content .....	32
Figure 4.1: Design of group-versus-individual trade-off experiment .....	38
Figure 4.2: Group-vs-individual trade-off and genotypes selected for further experiments ...	42
Figure 4.3: Water-use efficiency and predicted drought probability .....	43
Figure 4.4: Manhattan plot and local score results .....	44
Figure 4.5: Effect of allele variant at 12:9258726 on GI trade-off .....	44
Figure 4.6: Overlap between USDA core collection and genotypes used in this thesis .....	53
Figure 4.7: Countries of origin of the genotypes used in this thesis .....	54
Figure 4.8: Analysis of accidental drought experiment .....	54
Figure 4.9: Predicted probability of drought symptoms of focal plant .....	55
Figure 5.1: Niches and traits. ....	66
Figure 5.2: Ecological specialization of varieties through diverse strategies. ....	70
Figure 5.3: Alternative designs in mixture experiments. ....	73
Figure 6.1: Mixture experiment .....	79
Figure 6.2: Trait experiments and calculation of mean trait and reaction norm differences .	81
Figure 6.3: Comparison of cross-validation strategies .....	82
Figure 6.4: Biomass overyielding distribution .....	84
Figure 6.5: Comparison of cross-validation strategies .....	85
Figure 6.6: Simplified data leakage example .....	86
Figure 6.7: Predictive ability of models after shuffling genotypes .....	86
Figure 6.8: Density distribution of ( $R^2$ ) depending on cross-validation strategy .....	89

## List of Tables

Table 2.1: Example of a payoff matrix .....	12
Table 3.1: Within-genotype competition experiment results overview .....	26
Table 4.1: List of genotypes used throughout the thesis .....	50
Table 6.1: Details of soils used in greenhouse experiment .....	92

# 1. Summary

In the face of a growing world population and the rapid pace of anthropogenic climate change, there is an urgent need for innovation in agriculture. To this end, one promising approach is to reframe agricultural challenges in ecological and evolutionary terms and concepts. Here, evolutionary ecology and agronomy are bridged to investigate fundamental ecological questions within crop systems and to explore novel approaches to address pressing agricultural challenges. Using soybean as a model system, an interdisciplinary approach involving large-scale experiments, innovative experimental designs, advanced statistical tools, genetic analyses and machine learning is presented. The first study examines how plants allocate resources to foraging structures in the presence of neighbors. Using pot experiments and manipulating above- and below-ground interaction as well as inter-plant distances, it shows that competition cues are processed hierarchically, with aboveground cues having precedence over belowground cues in determining resource allocation. Allocation also followed patterns interpreted to be “optimal” for individuals but sub-optimal for the group. This is expected to lead to a “tragedy of the commons” in crop stands, i.e. group-optimal resource foraging strategies are likely evolutionarily unstable. The second study describes the use of “competitive games” to identify more cooperative or more competitive genotypes. Competitive genotypes exhibited reduced drought tolerance, highlighting a trade-off between competitiveness and stress resistance. Genetic association studies identified a candidate gene associated with genotypic competitiveness and know to play a role in growth regulation and stress tolerance in other plants. These findings have significant implications for breeding strategies: selecting for cooperative genotypes with higher drought tolerance could considerably enhance agricultural sustainability. The last two studies investigate soybean variety mixtures. First, ecological mechanisms underlying mixture diversity effects are reviewed, and their application in agriculture discussed. Then, theoretical considerations about functional trait differences and niche complementarity are paired with empirical evaluations of two-way soybean genotype mixtures, and major methodological pitfalls in predicting mixture performance are identified. This work on one hand proposes advanced statistical and machine learning approaches to improve predictions of mixture diversity effects. At the same time, it highlights the need for rigorous protocols to ensure the validity of the obtained predictions. Overall, the work presented here highlights the potential of evolutionary agroecology as an interdisciplinary field. By integrating ecological and evolutionary principles into agricultural research to address critical challenges in global food security, this thesis provides new insights for sustainable crop production and innovative breeding strategies in soybean and beyond.

## **2. General introduction**

The First Agricultural Revolution marked the shift from hunting and gathering to agriculture, leading to plant domestication and the development of landraces adapted to local conditions, which enabled stable food production and the rise of complex societies (Qaim 2016). However, early methods and breeding techniques were rudimentary compared to modern practices.

Agriculture developed slowly until the Second Agricultural Revolution, involving profound social, economical and agricultural changes. These changes included the establishment of improved crop rotations, more efficient machinery like the heavy iron plough or a conceptual shift of land property rights, away from commons towards exclusive ownership (Overton 1996). The increased agricultural efficiency led to the liberation of a high labor force formerly occupied in agriculture: a pre-requisite for the industrial revolution. The industrial revolution then set the ground for the production of modern agricultural machinery or mineral fertilizers, catapulting both agricultural productivity and human growth to never-seen levels.

Since the industrial revolution, and especially since the middle of the 20th century, there has been a significant increase in yield production of the most important food crops worldwide, in particular wheat, rice and maize (Calderini and Slafer 1998, Kucharik and Ramankutty 2005, Xu et al. 2020). This increase in yield, referred to as the Green Revolution, is attributed to multiple factors, among which the development of improved genotypes such as so-called dwarf varieties in cereal crops, as well as the large-scale diffusion of mineral fertilizers, pesticides and irrigation (Jain 2010, Cassman and Grassini 2013, Lee et al. 2015). The shorter stature of dwarf varieties makes them suitable for high-density planting and much less susceptible to lodging, considerably decreasing yield losses. The intensive use of mineral fertilizers allowed to decouple crop production from animal production (once central for manure production) and to reap the entire potential of the newly developed varieties.

### **2.1. Agriculture at a crossroads**

Over the past century, the global population has increased at an unprecedented rate, and projections by the United Nations indicate that this growth will persist in the coming decades (United Nations 2024). The ever-increasing number of people living on the planet poses modern agriculture in front of the challenge to produce enough calories to feed all of them. Moreover, the fast advancement of climate change and global warming requires agriculture to reconsider its means

of production, as modern high intensity agriculture is one of the main drivers of anthropogenic climate change (Lynch et al. 2021, IPCC 2022). This means that the agriculture of tomorrow not only is required to be highly productive, but also to be less detrimental to the environment. Additionally, agriculture will not only have to face the challenge of a warming climate, but also of a climate becoming more and more variable (Salinger 2005, Pendergrass et al. 2017). For example, climate scenarios predict more intense and more frequent drought periods (Cook et al. 2018). Overall, this indicates that modern agriculture is in urgent need of innovations, and novel approaches have to be developed to tackle these pressing problems.

### **2.1.1. Plant breeding**

Selective breeding by manually crossing chosen parental plants in a directed way began around the 18th century, but made substantial progress after Mendel's principles of inheritance were rediscovered around a century ago (Lee et al. 2015). Afterwards however, plant breeding dramatically improved the yields of many major food crops in only a few decades. The conventional breeding methods applied in this era, and still in use today, comprise hybridization as well as X-ray and chemical mutagenesis (Lee et al. 2015).

Since genome sequencing became affordable and easily accessible, it has become an essential component of modern plant breeding. If plant breeding in earlier days mostly focused on the selection, crossing and descendent-evaluation of breeding lines showing certain desired characteristics, plant breeding today is a highly specialized and methodologically advanced discipline. Indeed, the availability of genomic information for many breeding lines, combined with progress in the fields of marker assisted selection, high throughput phenotyping systems, genomic prediction or also machine learning applied to plant genomes allowed modern plant breeding to become highly efficient by combining a large amount of phenotypic and genotypic data (Tester and Langridge 2010, Kim et al. 2020).

At the same time, agronomy and breeding remain a multivariate optimization problem. Indeed, agricultural research has been compared to the exploration of a *fitness landscape* of peaks and valleys with respect to many possible variables (Figure 2.1, adapted from Weiner 2003). Many agricultural practices (e.g. purple in Figure 2.1) are already close to local optima (yellow in Figure 2.1), and modern technologies have become highly efficient at climbing the remaining part of the local peak. However, many modern technologies lack the ability to find and climb global optima (green in Figure 2.1). For example, modern genomic prediction methods are becoming increasingly better at making predictions within the populations for which they were trained.

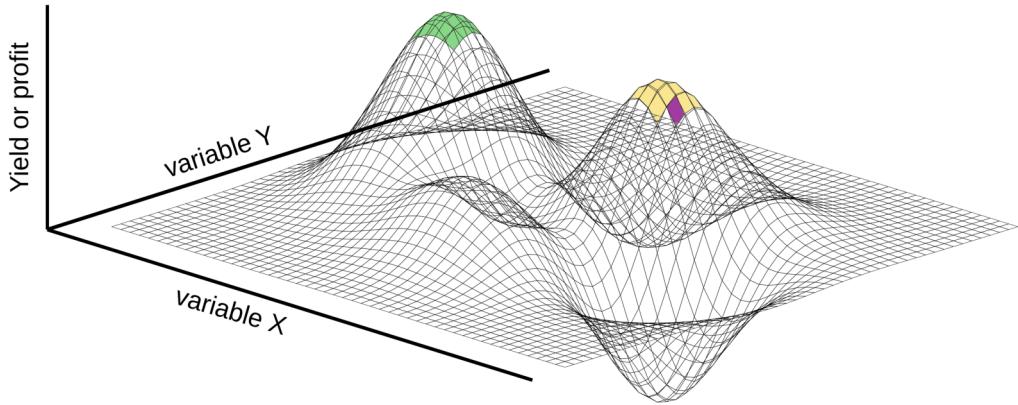


Figure 2.1: Agriculture and breeding can be seen as multivariate optimization problem. Many of today's agricultural practices (exemplified by the purple rectangle) are close to local optima (yellow) and modern technologies have become highly efficient to climb the remaining part of the local peak. However, they are not apt to find the global optimum (green) or to infer a way to reach it. Adapted from Weiner (2003).

However, they often struggle to extrapolate effectively to other populations, particularly when these populations differ genetically or environmentally from the training set.

Eventually, even the most astonishing methodological and technical advancement cannot answer the underlying question: *what are we breeding for?* To address this question, not only agricultural, but also evolutionary and ecological considerations are fundamental. Indeed, in a recent review Weiner (2019) wrote: “Evolutionary biology gives us compelling reasons to predict that many of the most ambitious objectives plant breeders and genetic engineers have proposed to increase crop yields are not likely to be successful”. In the present work, agricultural challenges will thus be addressed through an eco-evolutionary lens.

## 2.2. The ecological perspective

Agriculture can be considered as applied ecology, and agriculture should focus on communities or ecosystems rather than on processes at the level of the individual plant or lower (Weiner 2003). Therefore, progress in other fields like genetics or molecular plant breeding have to be interpreted

at the correct (community focused) contextual level if we want to fully understand their potential to address agricultural problems (Weiner 2017).

Surprisingly, agriculture is still often described as a system with inputs (land, water, fertilizers,...) and outputs (typically yield), where sustainability or environmental effects are considered to be “externalities” rather than intrinsic components of the system (Weiner 2017). Ecologists, on the other hand and in contrast to agronomists, typically do not describe their study systems as linear “input → output” systems. Rather, ecologists are used to think in terms of feedback loops and trade-offs. The consistent consideration of agronomy as applied ecology allows to reconsider agricultural systems from this perspective: consisting of trade-offs as opposed to a series of linear improvements (Weiner 2017). Combining ecology and evolutionary theory to address urgent agronomical challenges results in a discipline called *evolutionary agroecology* (Weiner et al. 2010), and the concept of *trade-off* lies at its heart.

### **2.2.1. Game theory and the tragedy of the commons**

Evolutionary game theory is the discipline studying how the frequency of one evolutionary strategy (for example a specific animal behavior, or an allocation pattern to competitive plant structures) changes given the presence of other strategies in the population (Alexander 2023). Within evolutionary game theory, the “tragedy of the commons” is a key principle, wherein individuals sharing and competing for a common resource with neighbors have a lower fitness than if the same resource were to be distributed fairly amongst them (Hardin 1968, Rankin et al. 2007).

In plants, and especially in crops, several studies have shown that tragedies of the common can occur even in simple two-player competition games (although there is some disagreement about the methodological consistency of some of these studies, as will be further explained in Chapter 3). When visualizing a tragedy of the commons through a so-called payoff matrix (Table 2.1), the arising dilemma becomes immediately clear: the group as a whole (here the two competing plants) would perform best if both plants cooperated (i.e. would not invest in costly competitive structures but rather in biomass or yield). This situation is called Pareto optimum. However, the single plant does best when it competes and its neighbor plant cooperates (i.e. it invests in competitive structures while the neighbor doesn’t, allowing the competitive plant to access more than its share of resources). This then leads to the situation where both plants investing in competitive structures is the evolutionary stable strategy (ESS, also Nash equilibrium), to the detriment of group performance.

Table 2.1: Possible payoff matrix for two competing plants. The numbers represent the payoff for the two players in each of the four possible scenarios. To maximize the total payoff, mutual cooperation between both plants is necessary (Pareto efficiency, light green). However, in the absence of knowledge about the opponent's strategy, competition always yields the greatest individual benefit for each plant (white cells). This leads to both players competing being a Nash equilibrium, i.e. a situation where no player can increase it's payoff by changing only it's own strategy (yellow cell).

		Plant B	
		Cooperate	Compete
Plant A	Coop.	3, 3	1, 4
	Comp.	4, 1	2, 2

While the payoff matrix is useful to visualize and better understand the concept of tragedy of the commons, Pareto optimality and Nash equilibrium, it also highlights a key difference between natural selection and what should be the goal of artificial selection. Indeed, natural selection acts primarily on the level of individuals (the underlying reason for why the tragedy of the commons appears). However, in an agricultural context, the focus is on group performance, as farmers are typically more interested in yield per area as compared to yield per plant. Therefore, emphasis should be placed on traits that can be optimized through breeding to approach Pareto efficiency. Indeed, the trade-off between individual performance and group-level performance could present substantial potential for breeding improvements (Weiner 2019). In this sense, some of the breeding successes of the Green Revolution can be directly attributed to mitigating a tragedy of the commons, where plants were made less competitive, thus freeing up resources for grain production (Jennings and Herrera 1968, Donald 1968, Zhu et al. 2019a).

### 2.2.2. Cooperation in crops: individuals vs. populations

In the last section it was stated that a part of the breeding success of the last century can be ascribed to a rapprochement of plant growth strategies towards a more group-optimal strategy. If this is true, then the breeding progress over the last century should be higher when observed

at group level then when observed at individual level. And indeed, common-garden experiments in hybrid maize demonstrated that the historical breeding progress is most evident under high planting densities, while the yield increase under low planting densities was comparatively modest (Figure 2.2, Duvick et al. 2004). In other words, modern elite maize varieties show a greater yield improvement over older varieties primarily in situations of strong plant-plant interactions, likely because the newer varieties are less competitive. This insight is of crucial importance as it highlights the divergence between individual performance and group performance: individual plants of modern varieties do not exhibit significantly higher yields when compared to individual plants of older varieties. However, modern varieties can be cultivated at much higher densities, thereby increasing the overall yield at the group level.

While in maize the reduction in competitiveness is attributed more to adaptations like a smaller root system (Rinehart et al. 2024) and more erect leaves (Elli et al. 2023) than to a smaller stature (Chen et al. 2019), the success of dwarfing genes in rice and wheat yield can at least partially be explained in the same way. Indeed, in a plant stand, it is advantageous for an individual to be taller than its neighbors, as predicted by the evolutionarily stable strategy. However, this investment in height incurs at a cost for the plant. If all plants in the group grow taller, the overall cost to the group also increases. The cost for the group will be greater the denser

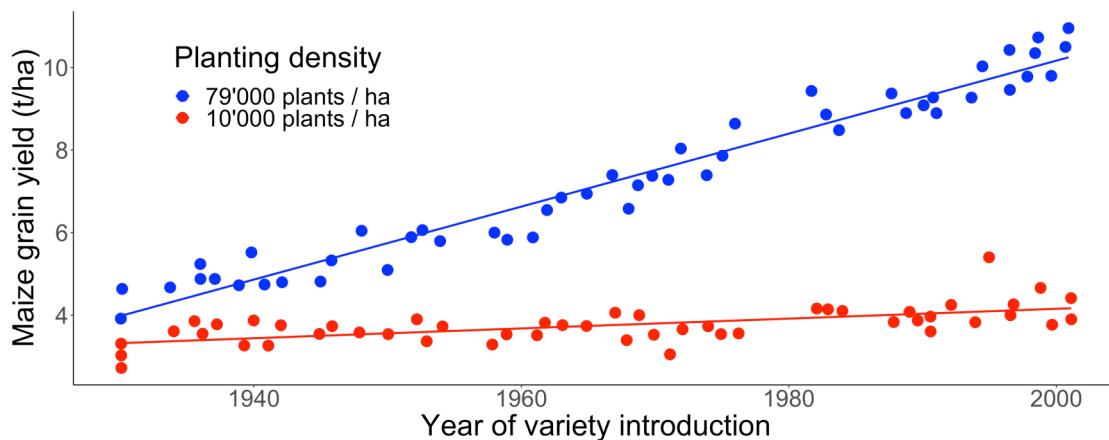


Figure 2.2: Modern maize varieties perform better than older varieties when planted at high densities (blue), while they perform similarly when planted at low densities (red). This suggests that the physiological improvement of the individual plants has been small compared to the adaptation of the plants to grow in dense stands without excessively allocating resources in competitive behaviors.  
Adapted from Duvick et al. (2004).

the stand. Conversely, if all plants in the stand remain smaller, more resources are available for investment in yield, bringing the stand closer to Pareto efficiency (Weiner 2019).

Of course, to be able to use these insights for practical applications, it is crucial to obtain an in depth understanding of plant-plant competition. Especially, it would be of great value to better understand the genetics underlying plant competition, as this could enable the targeted selection of less competitive plants within modern breeding programs. Moreover, another aspect that deserves attention is the investigation of the trade-offs associated with competition: as competition is typically costly, there might be situations in which for an individual plant it is more efficient to have a reduced competitive ability. For example, the “growth-defense” trade-off describes how plants can reduce their growth rate in favor of a higher defense against herbivores (Huot et al. 2014, He et al. 2022). Typically, a reduction in growth rate equals a reduction in competitive ability, but when the pathogen pressure is high the reduced competitiveness might be a price worth to pay. Albeit sounding trivial, this observation has profound implications. Indeed, it provides an evolutionary explanation for how and why cooperative traits in crops could be maintained in a population. Consequently, this consideration provides the scientific fundament for why searching for cooperative traits in crops might be a successful endeavour.

### **2.2.3. Biodiversity and ecosystem functioning**

Besides leading to a novel conceptual understanding of agriculture and the challenges it faces, the ecological (or eco-evolutionary) lens on agriculture also prepares the ground for thinking about non-standard agricultural systems. One of these systems is the mixed cultivation of multiple varieties of the same crop species on the same field. Variety mixtures can provide multiple advantages, the most well-known being pathogen suppression (Wolfe 1985, Smithson and Lenné 1996, Kristoffersen et al. 2020a, Vestergaard and Jørgensen 2024), but also including higher yield (Kiær et al. 2009, Borg et al. 2018) or improved productivity under drought (Kong et al. 2023). However, while the potential of variety mixtures is well known, their adoption is still low (Labarthe et al. 2021). One reason for this might be the renowned difficulty to predict mixture performance based only on data of the mixture components, a problem exacerbated by the exponential increase of possible mixtures with increasing considered varieties (Wuest et al. 2021). This impedes the development of high-performing mixtures, whose perceived disadvantages are more than outweighed by the demonstrable benefits.

Many studies investigating biodiversity-ecosystem functioning (BEF) relationships in plant communities have shown that the productivity of an ecosystems tends to increase with its diver-

sity (Hooper et al. 2005, Cardinale et al. 2011, Tilman et al. 2012, Huang et al. 2018). This effect is typically explained by so-called *niche complementarity* among the interacting species (Salles et al. 2009, Turnbull et al. 2016). However, identifying the exact mechanisms that drive such complementarity has proven to be a considerable challenge. Here again, agronomy and ecology might synergistically profit from each other. Agronomy would profit from a better understanding of diversity-productivity effects if this led to the ability to better assemble high performing variety or crop mixtures. Similarly, ecology can profit from agronomy as highly idealized study ecosystem wherein to investigate these relationships without being distracted by all the confounders present in natural ecosystems.

## 2.3. Soybean: past, present and future importance

Modern-day soybean (*Glycine max*) has its origins in East Asia, where the wild ancestor *Glycine soja* is believed to have been domesticated 6000-9000 years ago (Sedivy et al. 2017). This domestication led to a high diversity of *G. max* landraces, locally adapted to different climatic regions in East Asia (Haupt and Schmid 2022), from which modern-day varieties are descendants through selective breeding. Soybean is an annual plant, and most varieties are adapted to regions with warm summers and a long vegetation period. However, successful breeding programs in regions with moderately warm summers and shorter vegetation periods show that soybeans can also be bred to be adapted to these conditions (Schori et al. 2003).

### 2.3.1. Soybean today and tomorrow

Legumes, which can fix atmospheric nitrogen through a symbiotic relationship with rhizobia, are the second most significant group of plants for human use after grasses (Graham and Vance 2003). Their importance is expected to grow in the coming decades, as they provide a high-quality protein source with lower water usage and greenhouse gas emissions compared to animal-based products. Among legumes, soybean is by far the most prominent crop globally, although only 6% of its production is directly consumed as food, with the majority being used as livestock feed (Semba et al. 2021). Various studies have demonstrated genetic improvements in soybean achieved through breeding over the last century (Cober et al. 2005, Suhre et al. 2014). However, these advancements have been modest compared to those in maize, wheat, and rice. Consequently, global soybean yields remain lower than those of major cereal crops (Liu et al. 2020). This makes

soybean an ideal candidate on which to investigate the potential of ecological approaches for agronomical and breeding purposes.

## 2.4. Thesis outlook

This work investigates competition, cooperation and complementarity in soybean. To do so, approaches based on evolutionary agroecology, combined with a broad range of large experiments and modern statistical and computational procedures are used. The focus lies first on individual plants, then on genotypes and eventually on genotypic communities. The thesis addresses important agronomic and ecological questions: it investigates how cooperation and competition trade-offs can manifest in soybeans, it explores how cooperative genotypes and genes involved in cooperation can be identified and finally it reviews the current state of knowledge regarding ecological principles to develop variety mixtures and explores a novel approach to predict variety mixture performance.

To reconcile contrasting results from previous studies and investigate the presence and importance of tragedies of the commons in intra-genotypic soybean competition, a novel experimental design approach was developed. This approach allows to overcome the methodological limitations criticized in the past. To do so, the distance between the competing plants was varied within competition treatments. In this way, the factor of interest was no longer the competition treatment alone, but its interaction with the inter-plant distance: the same allocation is expected when plants are close but not competing and if the plants are competing but very far apart from each others. This allowed (1) to solve a long-lasting methodological debate about how to control for experimentally confounding factors in plant-plant competition pot experiments, (2) to show that tragedies of the commons can occur in soybeans and (3) to show that aboveground and belowground responses to competition are not independent (Chapter 3).

To identify cooperative soybean genotypes and explore them, a large partial-diallel experiment containing 90 highly different soybean genotypes was performed. Using a method recently presented for *Arabidopsis thaliana* (Wuest et al. 2022), this allowed to quantify competitiveness for all genotypes in the experimental population. A subsequent genome-wide association study allowed to scan the soybean genome for candidate genes involved in this behavior. The serendipitous observation of a correlation between a genotype's competitiveness and its drought susceptibility resulted in a fascinating additional level of complexity, giving insights in the costs of competition. This work showed that (1) there is variation in competitiveness between soybean genotypes, (2) this phenotypic variation can be associated with genetic variation at an identified locus and

(3) competitive genotypes have a higher water use resulting in increased drought susceptibility under limiting conditions, confirming that competitiveness should be considered as a trade-off comprising benefits but also costs (Chapter 4).

To better understand the challenges that are hindering a faster development of variety mixtures to be used in more sustainable agro-ecosystems, a literature review was conducted. The current state of knowledge was summarized and the current strategies applied in mixture assembly investigated. Then, novel approaches based on ecological theory were presented and their potential to guide the development of variety mixtures explained. The main messages of this review are that (1) current used metrics to approximate niche complementarity between potential mixture components are not precise enough, (2) novel and more precise niche complementarity approximators are needed and (3) the relationship between measured variety traits and conceptualized variety niches must be better understood (Chapter 5).

After some concrete ideas to improve the development of variety mixtures were proposed in Chapter 5, these ideas were then tested experimentally. Experiments conducted at two different sites allowed to compute a novel category of niche complementarity approximators, called *environmental reaction norms*. Moreover, non-linearities and discontinuities between functional trait differences and mixture diversity effects were accounted for. Two different cross validation strategies were compared to validate the ability of the models to predict mixture diversity effects from monoculture measurements. The comparison of the cross validation methods showed that (1) community level traits (i.e. mixture overyielding) are confounded with community components identities through data leakage and error propagation (Chapter 6). These results shed light on methodological issues that are presumably widespread in biodiversity-ecosystem functioning studies.

The results presented in this dissertation have profound ecological and agricultural implications. First, they demonstrate the potential of ecological and evolutionary approaches to tackle pressing agricultural questions. Second, they delineate exiting new paths to follow in agriculture and breeding to obtain more sustainable agro-ecosystems: from ways to reap the potential of latent cooperativeness in breeding lines to methodological challenges in the development of better variety mixtures. And third, they use simplified (agro-)ecosystems to investigate fundamental ecological questions like the integration of above- and belowground signals in plants upon competition or the costs of such competition.

### **3. Simultaneous above- and belowground competition leads to a tragedy of the commons in soybean**

Emanuel B. Kopp<sup>1, 2</sup>, Niels P.R. Anten<sup>3</sup>, Pascal A. Niklaus<sup>2, 4</sup>, Samuel E. Wuest<sup>1</sup>

1: Agroscope, Group Breeding Research, Müller-Thurgau-Strasse 29, 8820 Wädenswil, Switzerland

2: Department of Evolutionary Biology and Environmental Studies and Zurich-Basel Plant Science Center, University of Zurich, Winterthurerstrasse 190, 8057 Zurich

3: Centre for Crop Systems Analysis, Wageningen University, Wageningen, The Netherlands

4: University of Zurich Research Priority Program on Global Change and Biodiversity, University of Zurich, 8057 Zurich

## **Abstract**

1. Evolutionary game theory posits that plants competing with neighbors divert resources away from seed production and thereby reduce stand-level seed production. This phenomenon is known as the tragedy of the commons (TOC), and is undesirable in grain crops. Specifically, changes in light quality due to neighbor plants often induce costly responses like petiole or stem elongation, increasing individual competitiveness at the expense of stand-level yield. Similarly, it has been postulated that belowground competition can induce increased root production in the vicinity of neighbors, which can also reduce seed yield.
2. There has been a long-standing debate about the extent to which such a belowground TOC actually occurs. Experimental setups used to measure it in pot systems have been questioned because of the difficulty of simultaneously controlling pot volume, nutrient amount, and nutrient concentration. Moreover, observations of root growth under competition have been inconsistent, ranging from increased root production to spatial root system segregation.
3. Here, using soybean as a model system, we conducted a competition experiment to disentangle above- and belowground drivers of this tragedy of the commons. Specifically, we used above- and belowground physical dividers and varied inter-plant distances (between 3 and 24 cm) to avoid a confounding of divider treatments with plant-available soil volume.
4. Plants closer together had higher root-to-shoot ratios, indicating that an increased fraction of assimilated was diverted to root growth, but this only happened when they competed both above- and belowground (no dividers). Aboveground competition triggered shade avoidance responses, resulting in slender plant habit and reduced total root and shoot biomass.
5. Synthesis: Competition among soybean plants caused a tragedy of the commons. Interestingly, effects of belowground competition became only evident when plants were allowed to also compete aboveground. More generally, our findings show that varying the distance between competing plants allows to avoid the confounding of divider treatments with plant-available soil volume.

### 3.1. Introduction

Competition among plant individuals is a major driver of evolution and an important determinant of the functioning of ecosystems or communities (Harper 1977, Aschehoug et al. 2016). Despite the interest competition elicits in population ecological and evolutionary studies, competition has received relatively less attention in agronomy and breeding. At the same time, it has long been postulated and experimentally shown that population-level productivity in crop plants can be increased by minimizing allocation to competitive structures (Jennings and Herrera 1968, Donald 1968, Denison 2012, Weiner 2019). This notion is in line with evolutionary game theory that posits that natural or artificial selection in competitive environments often favors resource acquisition strategies that benefit the individual but are detrimental at the level of the population (Zhang et al. 1999, Gersani et al. 2001, Anten and Vermeulen 2016). In other words, as the resources acquired by competing plants are typically shared, but the gains through acquisition by each plant are private, the outcome of resource competition can lead to a tragedy of the commons (TOC), whereby populations that exhibit group-optimal resource foraging strategies can be invaded by individuals that invest more into acquisition at the expense of others (Hardin 1968, Anten and Vermeulen 2016).

In crops, there exist some well-described traits which enhance individual competitiveness to the detriment of population-level performance, particularly when plants “overinvest” into them. Minimizing this effect by breeding can increase stand yield. For example, varieties with reduced plant height and increased compactness contributed to substantially increased yield in crops such as maize, rice and wheat (Jennings and Herrera 1968, Hedden 2003, Duvick et al. 2004). Other traits may be equally relevant for crop improvement but are more difficult to target. A well-known example is “shade avoidance” (SA), whereby sensing neighboring plants through changes in light quality can trigger responses which increase individual competitive ability for light but reduce overall yield (Weiner et al. 2010). In soybean (*Glycine max*), shade often increases petiole and internode elongation but reduces branching (Lyu et al. 2023), and ultimately leads to higher disease susceptibility and lower yield per plant (Green-Tracewicz et al. 2011, Lyu et al. 2023). From an agricultural perspective, shade avoidance is typically undesired within a field because it reallocates resources to shade avoidance responses that would otherwise be available for seed production (Weiner et al. 2010, Pantazopoulou et al. 2021). At the same time, it is evident that some aspects of light sensing and the resulting responses can also have beneficial effects on the organization of the canopy and light interception in crop stands (López Pereira et al. 2017, Zhou

et al. 2024). Therefore, breeding progress by specifically targeting shade avoidance responses may be hampered by the existence of trade-offs.

Belowground competitive responses and their consequences for population-level functioning are comparably less well understood. Over two decades ago, two game theoretical modelling studies showed that the adaptation of individual plants to root competition can lead to a reallocation of resources towards increased root production, at the expense of overall group productivity (Zhang et al. 1999, Gersani et al. 2001). Gersani et al. found that soybean plants that competed belowground increased allocation to roots and had a lower seed yield than plants that were separated belowground by a physical soil barrier. This confirms game-theoretical considerations, and the notion of a root-competition induced tragedy if the commons. However, this interpretation was later criticized: although the two owners did have the same overall soil nutrients available as the sharers, each individual had only access to half the soil volume, which itself could be regarded a resource (Hess and De Kroon 2007, Semchenko et al. 2007, Poorter and Sack 2012, McNickle 2020). Several studies have since tried to correct for this bias (Maina et al. 2002, Chen et al. 2015), but controlling all three critical variables pot volume, total amount of nutrients per plant and soil nutrient concentration simultaneously has proven to be challenging (see McNickle 2020 for an in depth discussion). More recently, one solution to this problem has been to use nutrient-permeable but root-impermeable mesh dividers between competing plants, compared to using completely impermeable dividers for non-competing plants (Zhu et al. 2019b, Chen et al. 2021a). However, here the assumption is that root responses are caused by cues that pass the mesh divider but not the solid divider (Zhu et al. 2019b, McNickle 2020, Chen et al. 2021a, Cabal 2022). Overall, these studies indicate that the presence of neighbor plants can stimulate root growth in the zone in which plants compete. However, the opposite also has been observed: plants sometimes avoid root competition by producing spatially segregated root systems (Schmid et al. 2015, Lepik et al. 2021). Indeed, a recent meta-analysis showed that root responses in pea are inconsistent throughout published studies (Mobley et al. 2022).

Recently, Cabal et al. (2020) suggested that these apparently contradictory neighbor-effects on root growth occur because they are distance-dependent. They proposed that the costs of growing and maintaining roots increases with distance from the stem. Under these assumptions, a game-theoretical model predicted that plants respond to competition by producing more roots close to the plant's stem where the costs of maintaining roots are low, but producing less roots further away where costs are high. Therefore, the phenomenon emerging from such a model, termed “exploitative segregation of plant roots”, is one where the very same resource foraging

strategy leads to divergent neighbor responses depending on inter-plant distance (Cabal et al. 2020, Cabal 2022).

Here, we investigated the occurrence of a tragedy of the commons by studying the influence of competition on allocation patterns and resulting stand-level productivity in terms of biomass. We isolated contributions from shoot- and root competition and their distance-dependency. For this, we conducted an experiment that combined the traditionally-used separation treatments (but both separately above and belowground) with a range of inter-plant distances (Figure 3.1). With this design, we can separate effects of available soil volume from effects of interactions with the neighbor, because the competitive interaction decreases with distance, while the total soil volume remains constant. Hence, we focus on the statistical interaction of the separation treatment with the distance to the neighbor. For both the belowground and the aboveground treatments, we expect that neighbor-induced shifts in allocation only occur when plants are not separated and when the neighbor is close (c.f. Cabal et al. 2020, Cabal 2022). Specifically, (1) when allowing belowground competition only, we expect that the relative allocation to roots will increase with increasing competition, i.e. root-to-shoot ratios will increase (Figure 3.1a); (2) when allowing aboveground competition only, we expect that the relative allocation to shoots will increase with increasing competition, i.e. root-to-shoot ratios will decrease (Figure 3.1b); (3). In the absence of any divider, plants will compete both above- and belowground, and these two components of competition will likely not be independent. We therefore expect that effects of belowground competition will differ depending on whether plants also compete aboveground or not (Figure 3.1 c).

### 3.2. Materials and methods

To study competition among soybean plants, we conducted an outdoor pot experiment in Wädenswil, Switzerland (47.222°N, 8.669°E, 509 m a.s.l.) in summer 2023. All pots contained two plants of the same genotype that were either separated aboveground by an 80 x 20 cm white polystyrene sheet to prevent light competition, separated belowground by a plastic sheet to provide the plants with isolated soil compartments, separated above- and belowground, or neither separated above- nor belowground. These divider combinations were replicated with inter-plant distance ranging from 3 to 24 cm in 3 cm steps (Figure 3.1). We used four soybean genotypes obtained from the USDA soybean germplasm collection (Weber - PI 548524, Natsu Kurakake – PI 417187, Kanegawawase - PI 229330, and Kanagawa Wase - PI 506832), which form a gradient of competitive strategies ranging from “cooperative” (Weber) to “competitive” (Natsu

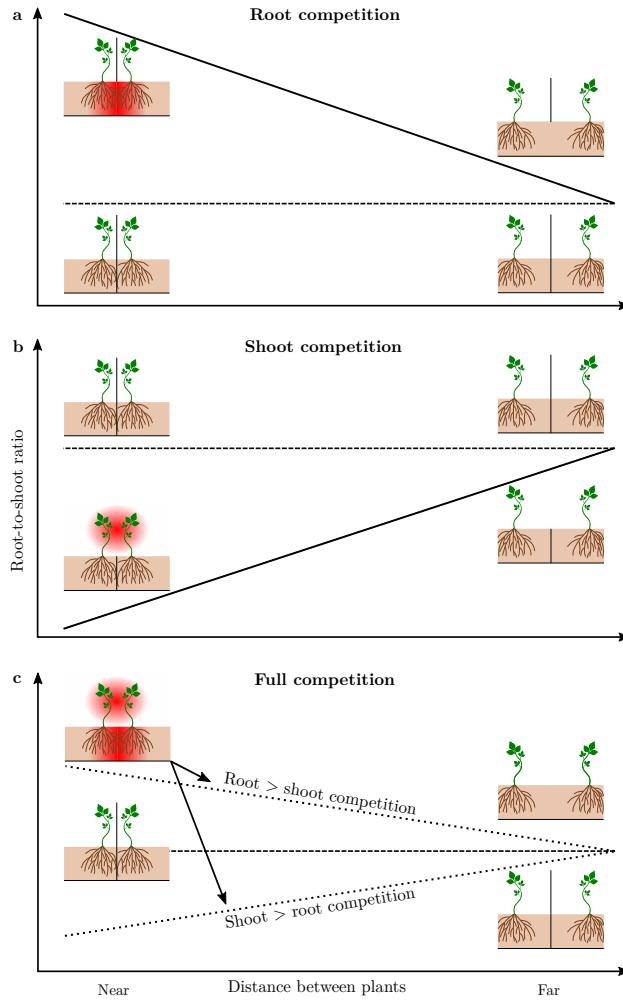


Figure 3.1: Expected outcome for root-to-shoot ratio under a tragedy of the commons. (a) Plants competing belowground and close together are expected to allocate relatively more towards their roots and to have a higher root-to-shoot ratio than non-competing plants. When the two plants are distant, both treatments are expected to show the same root-to-shoot ratio. (b) Plants competing aboveground and close together are expected to allocate relatively more towards their shoots and to have a lower root-to-shoot ratio than non-competing plants. When the two plants are distant, both treatments are expected to show the same root-to-shoot ratio. (c) Plants competing both below- and aboveground are expected to shift their root-to-shoot ratio according to the strength of the competition in the respective domain. If root competition is stronger than shoot competition, root-to-shoot ratio is expected to increase when the plants are closer together. If shoot competition is stronger than root competition, root-to-shoot ratio is expected to decrease when the plants are closer together. If competition above- and belowground is equally strong, the root-to-shoot ratio is not expected to deviate from the no-competition scenario. When the two plants are distant, both treatments are expected to show the same root-to-shoot ratio.

Kurakake). These strategies were determined in a separate experiment (see Chapter 4) using the methodology described in Wuest et al. (2022). Each separation treatment ( $n = 4$ ) x distance ( $n = 8$ ) x genotype ( $n = 4$ ) combination was replicated three times, resulting in a total of 384 pots. Pots were protected from hail by a net and irrigated ad libitum (Figure SM 3.5). Plants were germinated in propagator trays in the greenhouse for 14 days, and seedlings then transplanted to 7.5 L rectangular pots filled with lawn soil (Ökohum Art. 2633200) and moved outside.

Plant height was measured once a week, for a total of six measurements. Relative chlorophyll content was quantified one day before the end of the experiment with a MultispeQ V 2.0 device (Kuhlgert et al. 2016) by comparing the absorption of red light (absorbed by chlorophyll) to infrared light (reference). The plants were grown for 55 days until the majority of them had started to flower, then shoots were clipped at ground level, and roots washed. All biomass samples were dried at 60°C for 48 hours and weighed. Given that the two plants in each pot were in symmetric positions (i.e. the same distance from the opposite edges of the pots – see Figure 3.1), and that not all roots could unequivocally be assigned to a particular plant individual, we analyzed all data at the pot level.

Total plant biomass did not correlate significantly with root-to-shoot ratio ( $R^2 = 0.001, p > 0.5$ ), and we therefore analyzed root-to-shoot ratio without correcting for potential allometric effects. From the plant height measurements, we estimated linear growth rate by fitting the model  $H = H_0 + r \cdot t$  ( $H$ : height of the plants,  $t$  days since sowing of the height measurement) for each plant and extracting the linear growth rate  $r$ , which we averaged at the pot level. As we harvested the plants at flowering, height growth had not yet slowed, and a linear model fitted height better than a logistic one. Moreover, we calculated the shoot-biomass-to-height ratio as proxy of how much biomass were used per unit height growth.

We analyzed shoot biomass, root biomass, total biomass, root-to-shoot ratio, height growth rate, relative chlorophyll content and shoot-biomass-to-height ratio in R with the function **aov** (R 4.3.1; R Core Team 2021). In the model, we first corrected for spatial effects in the hail protection tunnels by fitting row and column terms (the pots were arranged in a rectangular grid pattern), and then added a full-factorial combination of genotype, distance between the pair of plants, aboveground divider treatment and belowground divider treatment. We were mostly interested in the dependence of the effect of the competition treatments on distance (two-way interaction of treatment with distance) and in the dependence of this relationship on genotype (three-way interactions of competition treatments, distance and genotype).

### 3.3. Results

Genotype had a statistically significant effect on all our response metrics ( $p < 0.001$  for main effect of genotype, Table 3.1, Figure 3.2a). None of the other effects we tested depended on genotype, i.e. there was no statistical interaction of genotype with any of the other model terms. The focus of our analysis was on effects of distance interactions, since distance moderates competition when dividers are absent. We found a significant three-way interaction between aboveground separation, belowground separation and inter-plant distance for final plant height, linear height growth rate, relative chlorophyll content, shoot biomass, and shoot-to-root ratio ( $p$ -values  $< 0.05$ ; Table 3.1). For shoot-biomass-to-height ratio, root biomass and total plant biomass, there was a significant two-way interaction between aboveground separation and inter-plant distance.

The shoot-biomass-to-height ratio, which is a proxy of the shade avoidance response, was lower in plants that competed aboveground, i.e. were not separated aboveground and grew closely together ( $F_{1,347} = 4.03, p < 0.05$ ; Figure 3.2b). Thus, plants competing more strongly

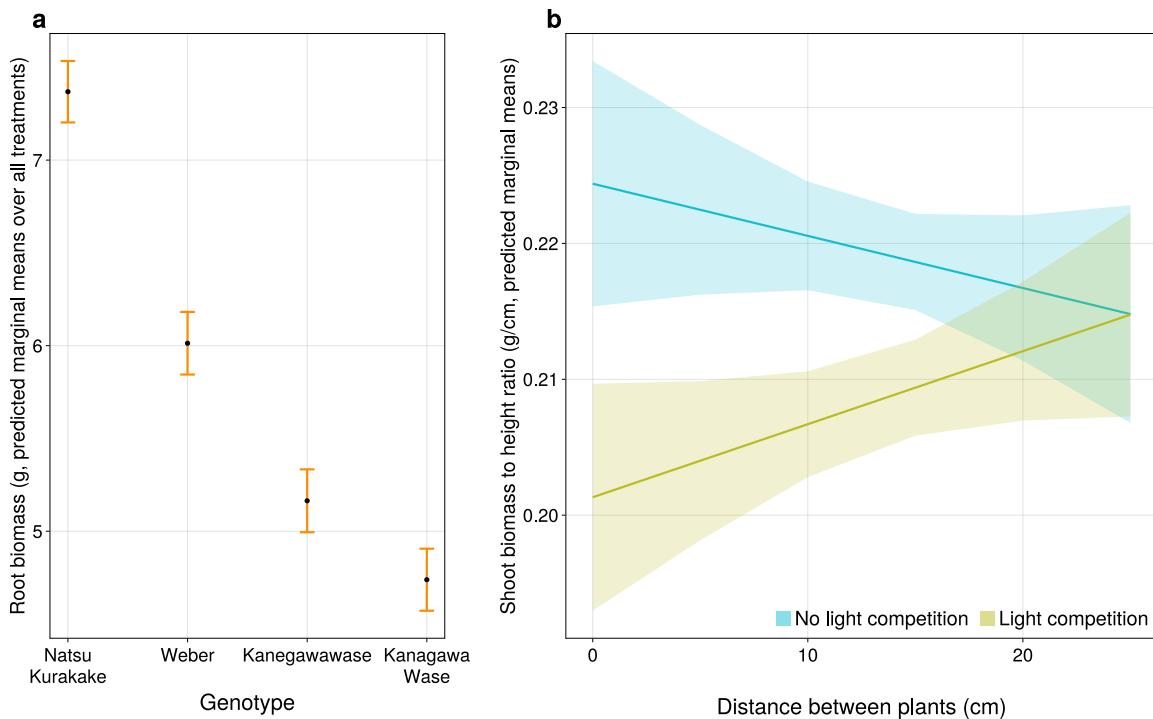


Figure 3.2: (a) The four genotypes strongly differed in all measured traits, as shown here for root biomass. (b) Plants which were competing aboveground and growing closer together invested less biomass per cm of height growth, in accordance with expectations based on known shade avoidance responses.

aboveground had a slenderer growth habit, in agreement with expectations. We next assessed the extent to which plants competing with their neighbors (i.e. no dividers and close distance) increased relative allocation to roots. We observed a consistent increase in the root-to-shoot ratio of soybean plants that grew closer together, but only when there was both above- belowground competition, i.e. when there were no dividers ( $F_{1,347} = 4.16, p < 0.05$ ; Figure 3.3).

Table 3.1: F statistic and p-value for linear models.  $Y$  was modelled in dependence of  $G$  (genotype),  $D$  (the distance between the pair of plants),  $C_A$  (the aboveground treatment) and  $C_B$  (the belowground treatment) and including all two- and three-way interactions between  $G$ ,  $D$ ,  $C_A$  and  $C_B$ .

Response	Explanatory variable	F (df <sub>n</sub> , df <sub>d</sub> )	p-value
Total biomass	G	161.03 (3, 347)	< 0.001 *
	D	0.001 (1, 347)	0.99
	D x C <sub>A</sub>	6.845 (1, 347)	<b>0.009 *</b>
	D x C <sub>B</sub>	0.02 (1, 347)	0.9
	G x D x C <sub>A</sub>	1.784 (3, 347)	0.15
	G x D x C <sub>A</sub>	0.122 (3, 347)	0.95
	D x C <sub>A</sub> x C <sub>B</sub>	2.62 (1, 347)	0.106
Shoot biomass	G	132.46 (3, 347)	< 0.001 *
	D	0.14 (1, 347)	0.71
	D x C <sub>A</sub>	5.71 (1, 347)	<b>0.017 *</b>
	D x C <sub>B</sub>	0.15 (1, 347)	0.7
	G x D x C <sub>A</sub>	1.04 (3, 347)	0.37
	G x D x C <sub>A</sub>	0.2 (3, 347)	0.9
	D x C <sub>A</sub> x C <sub>B</sub>	5.4 (1, 347)	<b>0.021 *</b>
Root biomass	G	187.11 (3, 347)	< 0.001 *
	D	2.66 (1, 347)	0.6
	D x C <sub>A</sub>	5.49 (1, 347)	<b>0.02 *</b>
	D x C <sub>B</sub>	0.69 (1, 347)	0.79
	G x D x C <sub>A</sub>	2.2 (3, 347)	0.088
	G x D x C <sub>A</sub>	0.05 (3, 347)	0.99
	D x C <sub>A</sub> x C <sub>B</sub>	0.09 (1, 347)	0.76
Root-to-shoot ratio	G	164.88 (3, 347)	< 0.001 *
	D	0.793 (1, 347)	0.37
	D x C <sub>A</sub>	0.006 (1, 347)	0.94
	D x C <sub>B</sub>	0.58 (1, 347)	0.45
	G x D x C <sub>A</sub>	0.786 (3, 347)	0.5

Response	Explanatory variable	F (df <sub>n</sub> , df <sub>d</sub> )	p-value
	G x D x C <sub>A</sub>	0.172 (3, 347)	0.92
	D x C <sub>A</sub> x C <sub>B</sub>	4.163 (1, 347)	<b>0.042 *</b>
Linear height growth rate	G	70.32 (3, 347)	< <b>0.001 *</b>
	D	0.001 (1, 347)	0.93
	D x C <sub>A</sub>	2.692 (1, 347)	0.1
	D x C <sub>B</sub>	0.912 (1, 347)	0.34
	G x D x C <sub>A</sub>	0.32 (3, 347)	0.811
	G x D x C <sub>A</sub>	0.323 (3, 347)	0.81
	D x C <sub>A</sub> x C <sub>B</sub>	6.993 (1, 347)	<b>0.009 *</b>
Final height	G	136.85 (3, 347)	< <b>0.001 *</b>
	D	0.001 (1, 347)	0.98
	D x C <sub>A</sub>	2.24 (1, 347)	0.135
	D x C <sub>B</sub>	0.611 (1, 347)	0.435
	G x D x C <sub>A</sub>	0.113 (3, 347)	0.952
	G x D x C <sub>A</sub>	0.3 (3, 347)	0.825
	D x C <sub>A</sub> x C <sub>B</sub>	4.774 (1, 347)	<b>0.03 *</b>
Shoot-biomass-to-height ratio	G	67.55 (3, 347)	< <b>0.001 *</b>
	D	0.11 (1, 347)	0.74
	D x C <sub>A</sub>	4.031 (1, 347)	<b>0.045 *</b>
	D x C <sub>B</sub>	0.01 (1, 347)	0.908
	G x D x C <sub>A</sub>	1.101 (3, 347)	0.349
	G x D x C <sub>A</sub>	0.992 (3, 347)	0.397
	D x C <sub>A</sub> x C <sub>B</sub>	2.176 (1, 347)	0.141
Relative chlorophyll content	G	6.545 (3, 347)	< <b>0.001 *</b>
	D	2.98 (1, 347)	0.085
	D x C <sub>A</sub>	1.761 (1, 347)	0.185
	D x C <sub>B</sub>	0.244 (1, 347)	0.621
	G x D x C <sub>A</sub>	1.282 (3, 347)	0.28
	G x D x C <sub>A</sub>	0.696 (3, 347)	0.555
	D x C <sub>A</sub> x C <sub>B</sub>	4.06 (1, 347)	<b>0.045 *</b>

Shifts in root-to-shoot ratio may be caused by changes in root biomass, in shoot biomass, or in both. We found that plants competing aboveground had lower root biomass ( $F_{1,347} = 5.49, p < 0.05$ ; Figure 3.4a). Shoot biomass was reduced when plants competed both above- and belowground ( $F_{1,347} = 5.4, p < 0.05$ ; Figure 3.4b). On the other hand, plants competing

belowground only had a higher shoot biomass (purple lines in Figure 3.4b). Height growth and relative chlorophyll contents responded similarly ( $F_{1,347} = 6.7, p < 0.01$  and  $F_{1,347} = 4.06, p < 0.05$ , respectively; Figure SM 3.6): plants grew faster and had higher relative chlorophyll content when competing belowground but not aboveground. This was also evident in the final height of the plants, where plants competing above- and belowground were smaller than plants competing belowground only ( $F_{1,347} = 4.77, p < 0.05$ ; Figure 3.4c).

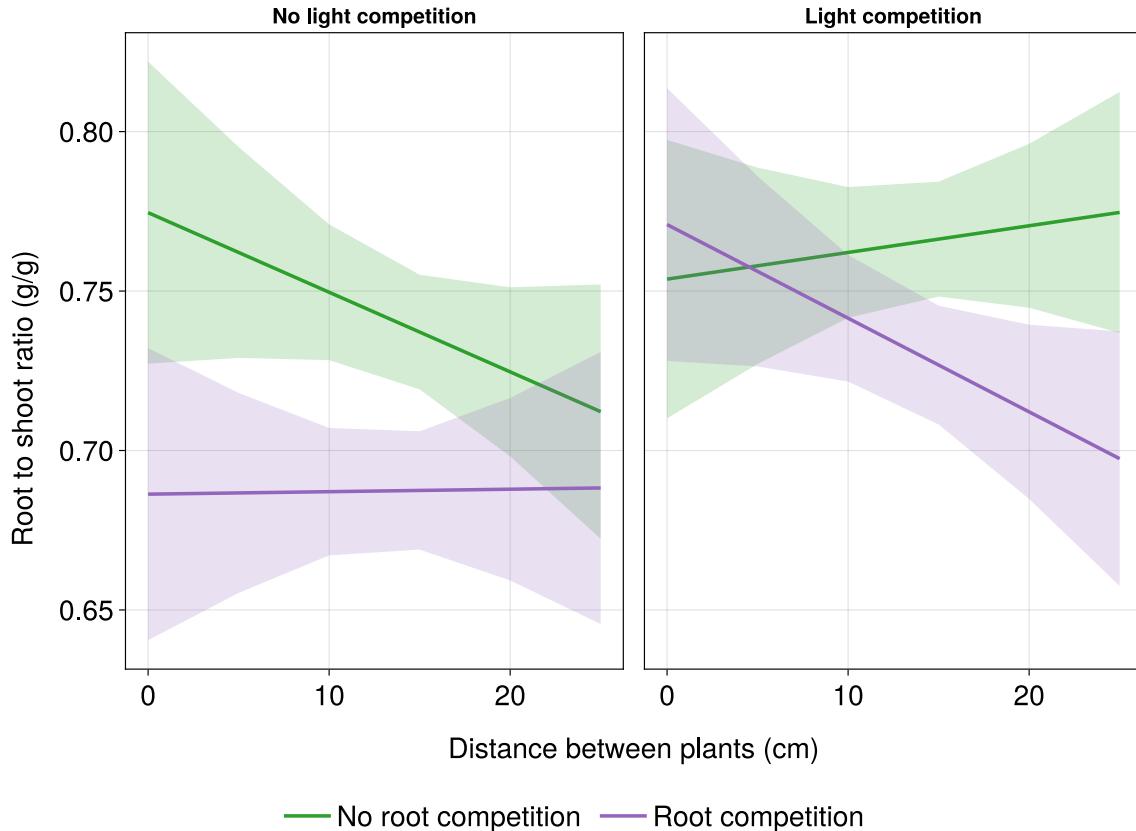


Figure 3.3: Predicted marginal means for root-to-shoot ratio. When facing no aboveground competition, the plants didn't increase their relative root production when exposed to belowground competition (left). When facing aboveground competition, plants increased their relative root production when simultaneously exposed to belowground competition and close together (right).

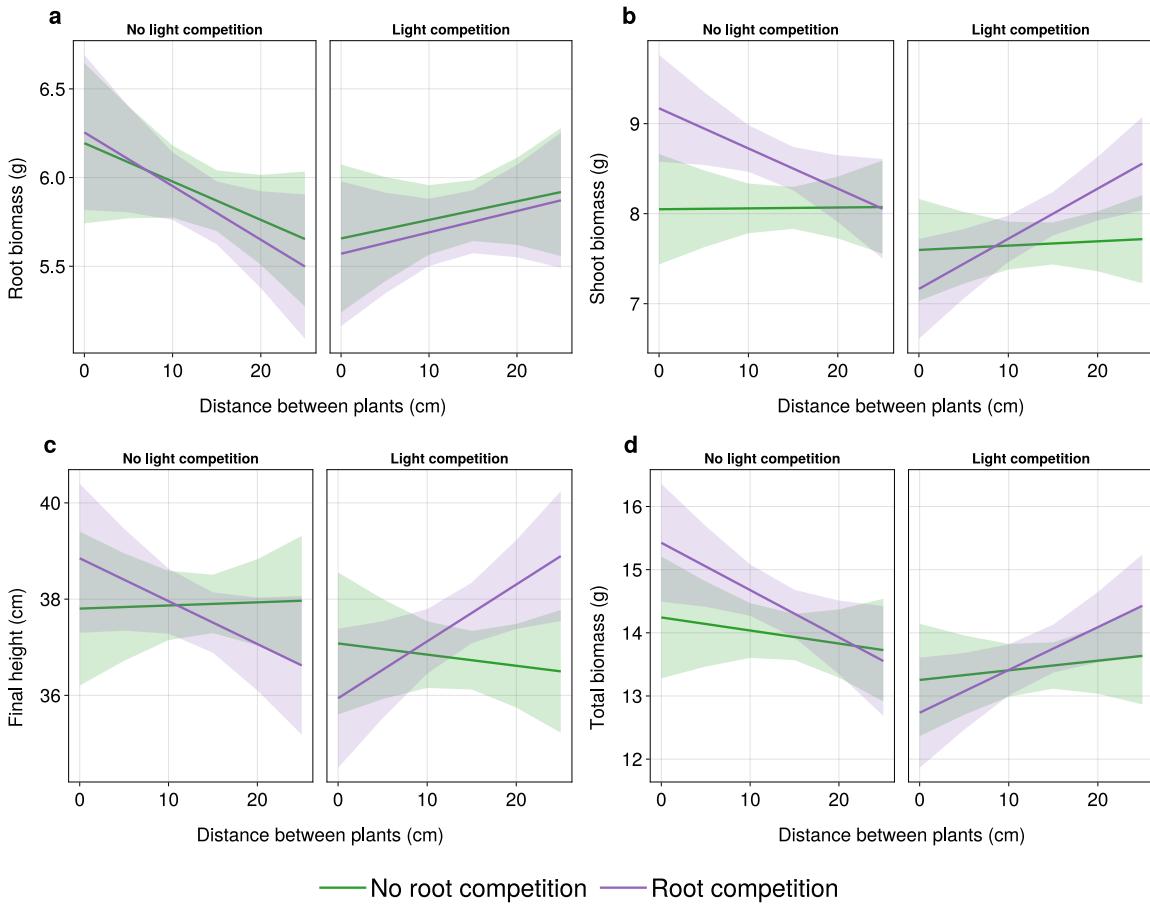


Figure 3.4: Predicted marginal means for (a) root biomass, (b) shoot biomass, (c) final plant height and (d) total biomass.

### 3.4. Discussion

#### Concomitant above- and belowground competition leads to a tragedy of the commons

In our experiment, we found increased relative allocation to roots and decreased total biomass (TOC) when plants were closer to neighbors, but only when plants were simultaneously competing aboveground and belowground. While for the occurrence of a TOC both aboveground and belowground competition were necessary, the aboveground competition was sufficient to lead to a reduction of total biomass. Interestingly, aboveground competition also led to a reduction of final root biomass, not influenced by belowground competition. This was unexpected, as previous results showed the root biomass to increase under a TOC, while we observed only an increased

relative allocation to roots. We assume that at the beginning of our experiment, the plants competing both aboveground and belowground did in fact produce overall more roots, but that this led to a reduction in shoot production and a slower growth rate. Over time, this led to the observed situation where competing plants show an increased relative allocation to roots without producing overall more root biomass.

Our results are in line with Gersani et al. (2001), who manipulated belowground competitive interactions using physical dividers. It is interesting to note that Gersani et al. (2001) identified a potential TOC in setups where competing plants were grown relatively close to each other, without aboveground dividers and in a greenhouse over the winter period requiring supporting artificial light – conditions that typically induce a shade avoidance response. This growth condition was similar to our close-proximity and no-aboveground-separator treatment. Our observations suggest that beyond neighbor distance, aboveground competition may also be important for the expression of a TOC. This could indicate that discrepancies among studies arose due to unanticipated yet important differences in experimental setups (i.e. the presence or absence of light competition due to light quality and quantity as well as neighbor-distance), instead of representing statistical sampling errors (Mobley et al. 2022).

### **Shade avoidance responses under aboveground competition cause a tragedy of the commons**

Responses to increased aboveground competition were consistent with the typical shade avoidance (SA) syndrome described previously (Keuskamp et al. 2010, Green-Tracewicz et al. 2011, Wille et al. 2017, Ballaré and Pierik 2017, Pantazopoulou et al. 2021, Lyu et al. 2023, Golan et al. 2023), i.e. more slender plants and lower allocation to roots. At the same time, plants were smaller when close together and competing aboveground. Plant height typically also increases under shade avoidance (Pierik and De Wit 2014). Thus, this was likely a consequence of the negative effects of shade avoidance on overall plant performance expressed in terms of biomass. Indeed, SA exerted a strong negative effect on overall plant performance. We argue that this was not caused by a limitation of light per se (Figure SM 3.5) but rather by the sensing of neighbors through light quality changes. Even though SA behavior can increase the fitness of individuals at high densities, our results and those of others suggest that the light-foraging strategy invoked by the presence of neighbors is detrimental to the performance of plants overall (Schmitt et al. 1995, Wille et al. 2017, Pantazopoulou et al. 2021).

### Aboveground and belowground signals are processed hierarchically

Competition between two plants ranges from fully symmetric to strongly asymmetric: depending on the extent to which larger plants acquire amounts of resources that are proportionately or disproportionately large relative to their size (Weiner 1990). Light competition is usually asymmetric, because the taller plant can pre-empt light, as light is a largely directional resource. Nutrient competition is more symmetric with nutrient acquisition largely proportional to the size of the root system (Schwinnning and Weiner 1998). From our study, we think that plants adjust their responses to different competition cues hierarchically. In this scenario aboveground signals would have precedence over belowground signals. Indeed, the interaction between aboveground competition and neighbor distance was significant for multiple measurements. However, the interaction between belowground competition and neighbor distance was never significant in absence of an interaction with aboveground competition. An integration of above- and belowground responses to competition on root growth could occur through various processes. For example, plants may “sense” the strength of a neighboring competitor aboveground and adjust their belowground foraging strategy (Gottlieb and Gruntman 2022). Alternatively, as predicted by game theoretical models for belowground interactions (Cabal et al. 2020), aboveground interactions with neighbors may also increase the costs of root production and maintenance. Certainly, extending the current game theoretical models to incorporate interactions in both above- and belowground domains would be revealing.

### Neighbor response should be considered as breeding target in future

Our approach presents a novel and simple way to control for the confounding factors which have long dominated the debate about plant pair competition experiments. We strongly suggest to explicitly account for neighbor distance in future works investigating the effect of individual foraging strategies on plant group performance. While it was anticipated long ago, most famously by Donald (1968), that reducing competitive traits of crop plants may harbor much potential to increase yields (see also Jennings and Herrera 1968, Duvick et al. 2004), this has recently become a renewed focus of the emerging research field of “crop evolutionary ecology” (Denison 2012, Anten and Vermeulen 2016, Weiner et al. 2017, Weiner 2019). Interestingly, the four different genotypes included in the experiment were of different origin and improvement status (e.g. landrace Natsu Kurakake vs. elite breeding line Weber). However, despite strong absolute differences amongst genotypes in all allocation traits measured, there was no evidence that the responses to competition varied amongst them. This indicates that neighbor response has not been targeted by breeding so-far, and should be considered for future crop improvement.

### 3.5. Supplementary materials



Figure 3.5: Experiment under hail protection net. The plants were clearly not light limited.

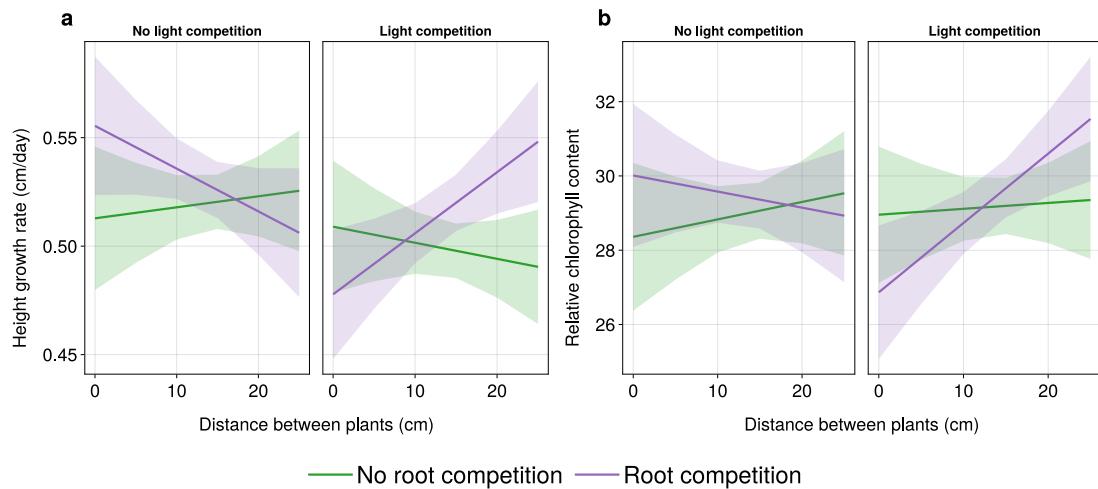


Figure 3.6: Predicted marginal means for (a) height growth rate and (b) relative chlorophyll content.

## **4. Cooperative soybean genotypes mitigate drought stress through water sharing**

Emanuel B. Kopp<sup>1, 2</sup>, Lukas Vonmetz<sup>1</sup>, Simon Jaeggi<sup>1</sup>, Pascal A. Niklaus<sup>2, 3</sup>, Samuel E. Wuest<sup>1</sup>

1: Agroscope, Group Breeding Research, Müller-Thurgau-Strasse 29, 8820 Wädenswil, Switzerland

2: Department of Evolutionary Biology and Environmental Studies and Zurich-Basel Plant Science Center, University of Zurich, Winterthurerstrasse 190, 8057 Zurich

3: University of Zurich Research Priority Program on Global Change and Biodiversity, University of Zurich, 8057 Zurich

## **Abstract**

Over the last 80 years it has repeatedly been postulated that social plant traits were the key to substantially increasing field crop yields, by rendering plants more “communal” or “cooperative”. However, if such theories simply represent anthropomorphic interpretations, or also have a practical utility in breeding, remains unclear. Here we use a highly diverse soybean population and a phenomenological approach to identify cooperative or selfish genotypes by measuring the outcomes of different competitive “games”. In a validation experiment that serendipitously became exposed to a seasonal drought, we found that cooperative genotypes exhibited lower water stress than selfish one, and so did their neighbors. We subsequently confirmed that cooperative genotypes demonstrate a higher water-use efficiency, enabling water-sharing and reducing stress symptoms of neighbors under water scarcity. Genetic association studies in our soybean experiment revealed that cooperation was partially associated with genetic variation in genes encoding Gretchen Hagen 3-type auxin conjugating enzymes, which have been recognized to play a role in drought susceptibility and root growth. Overall, our study shows that investigating the outcomes of plant competitive games can identify genotypes exhibiting cooperative traits, and pinpoint genetic determinants of selfish vs. cooperative behavior. This approach is compatible and to a large degree complementary to modern breeding approaches

## 4.1. Introduction

The traits of modern crop plants have been shaped by millions of years of evolution prior to their domestication. It has therefore been argued that many physiological and biochemical processes have been “optimized” to a high level, leaving little room for improvement for plant breeders trying to enhance the performance of individual plants (Denison 2012, Weiner 2019). At the same time, cash crops like maize, rice, and wheat saw a high yield increase over the last century (Fischer and Edmeades 2010), only partially explained by the large-scale introduction of mineral fertilizers. Indeed, a major component of this increase can be ascribed to the improved performance of plant groups growing in dense stands, rather than to a productivity increase of the individual plants (Duvick et al. 2004). This principle had already been demonstrated at the International Rice Research Institute (Jennings and De Jesus 1968, Jennings and Herrera 1968) during the breeding of the “miracle rice” variety IR8, and later formalized in the concept of a “communal ideotype” by Donald (1968). Donald suggested that natural selection, driven by competition, often leads to the excessive development of certain resource-foraging structures (Donald 1968, Zhang et al. 1999). This represents a special case of a “tragedy of the commons”: all plants will grow better if they compete less with each other, but if only some plants hold back, others might take advantage of them (Hardin 1968). Communal plants may yield less individually but, when grown together in dense fields, produce higher overall crop yields because they compete less with one another (Donald 1968, Zhang et al. 1999). Indeed, many of the traits selected by breeders, often inadvertently through selection for higher population yields, support the idea that “social” traits, improving group performance especially at high densities, have contributed most to the genetic progress over the decades. Examples include reduced plant height, vertical leaf angles or steep roots in monocots (Jennings and Herrera 1968, Hedden 2003, Lynch 2013, Ferrero-Serrano et al. 2019, Gaur et al. 2020, Tian et al. 2024) or improved water-use efficiency in soybean and wheat (Feng et al. 2022, Xi et al. 2023). These changes are paralleled by a reduced competition with neighbors. Accordingly, it has been proposed that a better understanding of the trade-off between individual plant performance and performance of plants as a group might be an important key to increase stand-level yield (Weiner 2019). A recent study showed how this group-versus-individual performance can be quantified for a given set of genotypes by exposing each genotype to competition with a same-genotype neighbor or with different-genotype neighbors. Moreover, it demonstrated the aptitude of this measure for further genetic association studies on the model plant *Arabidopsis thaliana* (Wuest et al. 2022). However, it is unclear whether these methods are applicable to crop species, which have more complex genetics and a long history

of domestication and breeding, and likely a lower genetic diversity compared to wild ancestors (Khoury et al. 2022).

The yield increases over the last century of the most important protein-rich crop plant worldwide, soybean (*Glycine max* - Semba et al. 2021), have been low compared to other important crops (Liu et al. 2020). Therefore, soybean may still have considerable untapped breeding potential. The slower breeding progress in soybean as compared to grain crops is at least partially a consequence of the different plant-architecture between typical grain crops such as rice or wheat and the pod crop soybean. In grain crops, the reduction of plant height led to a reduction of competition in dense plant stands, preventing an excessive development of resource-foraging structures (i.e. reducing the tragedy of the commons) and leading to higher community yields. On the other hand, soybean yield depends on internode number, which in turn depends on plant height (Liu et al. 2020). Thus, the reduction of plant height can indeed reduce competition, but it simultaneously reduces yield. A better understanding of soybean competition and the genetics underlying it might be the first step towards more communal soybean genotypes.

Drought is one of the central stressors threatening plant production worldwide and deserves special consideration in breeding. Plant strategies to cope with water scarcity can vary widely, ranging from escape to avoidance and tolerance (Levitt 1980, Ludlow 1989, Skirycz 2011, Palta and Turner 2019). Drought escape occurs when plants accelerate their development to reproduce before the drought conditions become critical (Shavrukov et al. 2017). Tolerance strategies include adaptations that allow growth and persistence under water stress, e.g. osmotic adjustments by solute accumulation that can reduce the turgor loss point (Bartlett et al. 2012, Delzon 2015). Avoidance strategies involve mechanisms improving water-use efficiency (WUE), e.g. through the reduction of transpiration rate by decreasing stomatal conductance (Anyia 2004, Kooyers 2015). Previous studies and theory indicate that there might be trade-offs between WUE and growth-related traits (Heschel and Riginos 2005, Franks 2011), and high WUE may represent a disadvantage in mixed-strategy neighborhoods, where saving water is evolutionarily unstable (Cohen 1970, Midgley and Moll 1993). That breeding could substantially improve water-use efficiency of crops (Condon 2004, Yang et al. 2020, Xi et al. 2023) is in line with the notion that breeding for higher population yields under water-limited conditions selects for “cooperative” plant genotypes that manage the available water more effectively, e.g. by saving water early during the growth period and leaving more water for the later, drought-sensitive stages (Xi et al. 2023).

Here, we established a method to quantify social strategies in soybean, based on outcomes of “competitive games”. We show a strong correlation between reduced genotypic competition, interpreted as more cooperative strategies, and increased drought avoidance, both at individual

and at group level. Moreover, we identified a small chromosomal region correlated with genotypic competitive ability. In this region we also found a gene belonging to the gene family of Gretchen Hagen 3 auxin-conjugating enzymes, and previously described to play a central role both in drought susceptibility and growth regulation.

## 4.2. Materials and methods

### Data collection

**Experimental population.** A total of 90 soybean varieties were assembled from the USDA germplasm collection and from varieties bred and released by Agroscope (see Table 4.1 for a complete list of the used varieties). These varieties covered a broad range of geographic origins and steps in breeding progress, but all reached maturity when grown outdoors in Switzerland (Figure SM 4.6 and Figure SM 4.7).

**Group-vs-individual (GI) trade-off experiment.** To test for variation in social strategies among these varieties, a first pot experiment was conducted in the summer 2022 in Wädenswil, Switzerland (47.222°N, 8.669°E, 509 m a.s.l.). The experiment was designed as partial diallel: in two blocks the 90 varieties were grown either in same-variety pairs (monocultures, 2 replicates per block) or in combination with one of eight tester varieties, differing in geographic origin and breeding progress (mixtures). This allowed to expose all 90 genotypes to these 8 different competition strategies. Each block was completed by 41 unique pairwise varieties combinations to increase the number of “social environments” a variety is exposed to. Therefore, the experimental design consisted of 90 x 2 monocultures, 90 x 8 mixtures and 41 additional pairs for a total of 941 pots or 1882 plants per block (Figure 4.1a). After inoculating the seeds with commercial rhizobial inoculum (HiStick, BASF, containing *Bradyrhizobium japonicum*), the plants were germinated in propagator trays in the greenhouse and transplanted into 5 l pots (see supplementary materials) after 14 days. Each pot contained two plants spaced 15 cm apart. All plants were inoculated with 10 ml of a solution of 10 g/l of mycorrhiza (*Rhizophagus irregularis*, Lalrise Max WP, Lallemand Inc, Canada). The pots were moved outside at the beginning of May and kept under a hail protection net until harvest (Figure 4.1b). Plants were irrigated daily. At time of harvest, the aboveground biomass of each plant was collected, dried at 60°C, and weighed. Afterwards, all seeds were removed and the seed yield determined. To avoid losses due to physiological and phenological differences between the varieties, every variety was harvested at its physiological maturity instead of harvesting the entire experiment at once. Narrow-sense

heritability ( $h^2$ ) for both aboveground biomass and seed yield was calculated in R using the variances obtained from fitting a model accounting for genotypic relatedness with *asreml* (The VSNi Team 2023).

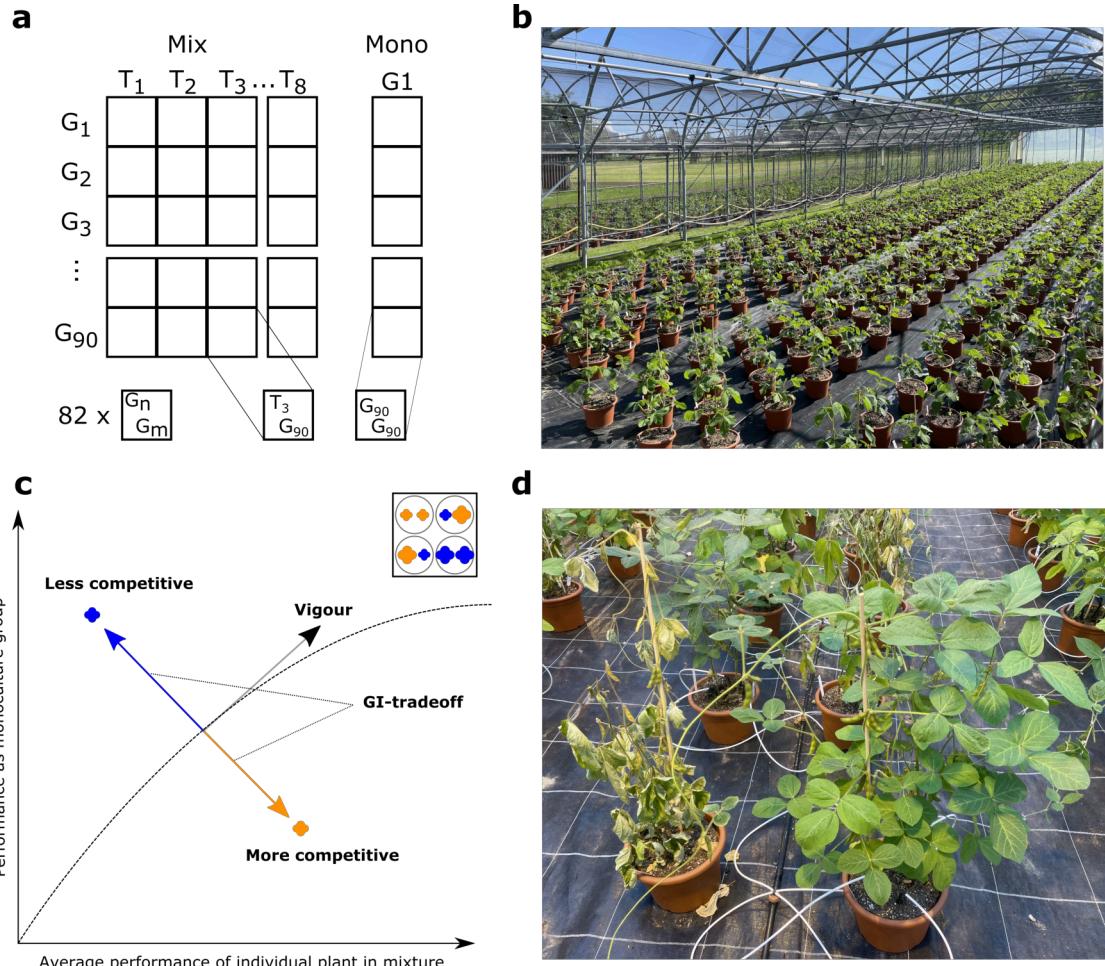


Figure 4.1: (a) Design of GI-experiment. The 90 genotypes were grown in pairs with 8 tester genotypes. Additionally, all genotypes were grown in monoculture. (b) Group-vs-individual (GI) trade-off pot experiment under hail protection net. (c) Conceptualized representation of how GI trade-off is calculated. (d) Drought experiment in greenhouse with a drought treatment plant (left) and a control plant (right).

**Accidental drought experiment.** To validate the results of the first experiment, we performed a half-diallel pot experiment with 12 varieties, chosen based upon preliminary results of the previous experiment (Figure 4.2b) and grown in all possible pairwise combinations, using the same protocols as in the previous experiment. Additionally, for each variety 6 pots with single plants were added to the experiment to establish biomass production in the absence of competition. This resulted in 66 mixtures and 12 monocultures (each with 5 replicates) and 12 single plants (each

with 6 replicates), for a total of 390 pots. Due to unusually high temperatures during the summer 2023 and a failure of the irrigation system, the experiment was terminated prematurely when plants started to show severe drought symptoms, ranging from wilted leaves to dead plants. We therefore scored these drought symptoms (ordinal scale with 3 levels; 1 being almost no symptoms and 3 plants being completely dry), harvested the plant shoots, and dried and weighed the shoots. As plants had not reached maturity, seed yield could not be measured.

**Drought experiment.** In early summer of 2024, an additional pot experiment was conducted in a greenhouse in Wädenswil, Switzerland. Again, all seeds were inoculated with rhizobia and germinated in propagator trays before being planted in 5 l pots filled with the same substrate as before. The same 12 varieties as in the accidental drought experiment were used. Of these 12, two varieties at the opposite end of GI were selected as tester varieties (Weber - PI 548524 and Natsu Kurakake - PI 417187). The plants were again planted in pairs: each of the 12 varieties was planted 10 times with each of the two tester varieties, and for the 10 non-tester varieties 10 monoculture pots were added. This gave a total of  $[(12 \times 2) + 10] \times 10 = 340$  pots. All pots were well-watered until the beginning of flowering (around 6 weeks from seeding), then the 10 replicates were split in two treatment groups. In the control group, the irrigation was kept at about 420 ml per pot and day, while in the drought treatment it was reduced to 210 ml per day. Once the plants in the drought treatment started to show signs of drought (Figure 4.1d) the irrigation in the dry-down treatment was completely stopped. Each pot was weighed three times per week throughout the experiment to track the water consumption of the plants. As soon as the first drought symptoms appeared, their intensity was scored on an ordinal scale from 0-4 (0: no symptoms, 1: low symptoms, 2: medium symptoms, 3: severe symptoms, 4: dead). The scoring was repeated three times, until most plants were dead. At the end of the experiment, all plants were cut at soil level, dried at 60°C for a minimum of 24 h and weighed.

**Sequencing.** DNA was extracted from seedlings of all genotypes using the CTAB extraction method (Doyle and Doyle 1987), and low-pass whole genome sequencing was performed at the BGI and the DNBseq platform (0.5x coverage). Genotypes were imputed against a reference haplotype collection through the commercial service provided by Gencove (Li et al. 2021a). The genetic data was filtered to keep only biallelic SNPs with minor allele frequency higher than 0.05, resulting in a total of ~3.6 mio SNPs.

## Data analysis

**Spatial correction.** To account for spatial variation within our two hail-protected outdoor tunnels in the GI-trade-off experiment, we first removed the block effect. We fitted a linear model using block as predictor, then summed the intercept and the model's residuum for each point as block corrected estimate for the given point. We then checked positional effects along rows and columns and corrected for them using linear models in the same way.

**Group-vs-individual trade-off.** For each variety used in the experiments we calculated a group-vs-individual trade-off by comparing its performance in a monoculture group against its mean performance as an individual growing with a neighbor of a different variety (Wuest et al. 2022). Following game theory, more cooperative genotypes will perform better in monocultures compared to their average performance in mixtures, while more competitive genotypes will perform better as individual in mixtures compared to their monoculture performance. We therefore fit a model of the form  $Y = a + bX + cX^2$ , where  $Y$  is the mean monoculture performance of a genotype and  $X$  is the mean performance of an individual of this genotype in a mixture. This makes more cooperative genotypes lie above the line defined by the model, while more competitive genotypes lie below the line (Figure 4.1c). By taking the orthogonal distance between each genotype's position and the best fit line, we obtain a continuous estimate for the genotype's position along a competitiveness-continuum, which is referred to as group-versus-individual trade-off (GI). We calculated GI on square root transformed data to partially counterbalance the effect of GI having a higher variance with increasing mean genotypic biomass.

**Analysis of accidental drought experiment.** To test whether a genotype's GI, as well as whether the GI of the neighbor plant affects drought tolerance, we run a cumulative link mixed model (CLMM). CLMMs work by modeling the probabilities of an ordered outcome as a function of fixed and random effects, using a cumulative probability framework to capture the ordinal structure of the data. Besides the GI of both plants, we added the biomass of both plants as predictors to ensure that the drought symptoms were not due to size differences only. Genotype identities of the focal and the neighbor plant were added as random effects. In each pot, the two plants were randomly assigned to be either the focal or the neighbor plant before the analysis was run. To avoid the models results to be an artifact of the plants selected as focal or neighbor, this assignment and the analysis were repeated 100 times and the results averaged. All models were done in R using the `clmm` function from the `ordinal` package (Haubo and Christensen 2018).

**Analysis of drought experiment.** We quantified the water consumption over time for every dry-down pot by taking the difference of its weight and the weight of its control pot at each weight measurement. As the control pots were always standing next to the corresponding dry-down pots, this allowed us to account for potential spatial effect within the greenhouse that could lead to differences in evaporation between the pots. To obtain water-use efficiency (WUE) for each genotype, we used the monoculture pots only. For each genotype separately, we performed a linear model with the weight difference between the dry-down and the control pots as response and the days since start of dry-down as predictor. We extracted the slopes, obtaining a measure of water consumption per day for each genotype, and divided each genotype's daily water consumption by its final mean biomass. This gave us an estimate of WUE in grams of biomass produced per gram of water consumed per day. We then analyzed the influence of GI on WUE. All linear models were done in R using the `lm` function of the `stats` package.

To investigate the effect of GI on the genotypic probability to display drought symptoms, we run CLMMs in R with the function `clmm` from the `ordinal` package. For this, we considered only non-monoculture pots. We first investigated the effect of a genotype's GI on its own probability to display drought symptoms. We fitted a CLMM with the drought score of each focal plant at 18 days after start of treatment as response, GI of the scored genotype as fixed term and the focal genotype identity as random term. We added the genotype identity of the neighbor plant (two levels) as fixed term to correct for neighbor effects. Next, we investigated the effect of a genotype's GI on its neighbor plant. For this, we took the score of the plants belonging to the two tester varieties Weber and Natsu Kurakake 18 days after the start of the treatment as response, the genotype identity of the scored plant as fixed effect (two levels), the GI of the neighbor plant as fixed effect and the genotype identity of the neighbor plant as random effect (12 levels). For all cumulative link models, we checked for the proportional odds assumption to hold. Moreover, we run all the models a second time adding the biomass of the plants as predictors to ensure that the effects were not driven by size differences only. As the results did not change when adding biomass as predictors, all reported results refer to the models without biomass.

**Genome-wide association study.** After calculating GI, we performed a genome wide association study (GWAS) using GI as response (GWAS was performed in R using `ASRgwas`, Galli et al. 2022). After performing a standard GWAS, we performed a local score analysis ( $\xi = 1.8$ ) to identify genomic regions with a higher-than-average concentration of markers with medium to high signal (Bonhomme et al. 2019). This method is well suited for detecting relatively weak marker–phenotype associations in GWAS with high marker density. We performed a GWAS

with subsequent local score calculation both for GI based on above-ground biomass data and on seed yield data and filtered for the genomic regions that were found significant in both analyses. Afterwards, we intersected the resulting regions with the soybean reference genome (*Glycine\_max\_v2.1.58*) and filtered for genes.

### 4.3. Results

Average aboveground biomass and seed yield (per plant) of the GI-experiment was 66.8 g (S.E. 28.8) and 20.3 g (S.E. 9.6 g) respectively. Both aboveground biomass and seed yield was highly variable, mainly due to the genotypic diversity in the experimental population. Nevertheless, narrow-sense heritability for aboveground biomass and seed yield was 0.53 (S.E. 0.04) and 0.58 (S.E. 0.04) respectively.

**Group vs. individual trade-off.** The varieties used in our experiments showed variation in their GI-trade-off (Figure 4.2a). This suggests that there is genetic variation in the elements driving competition and cooperation in soybean. Narrow-sense heritability for GI equalled 0.19

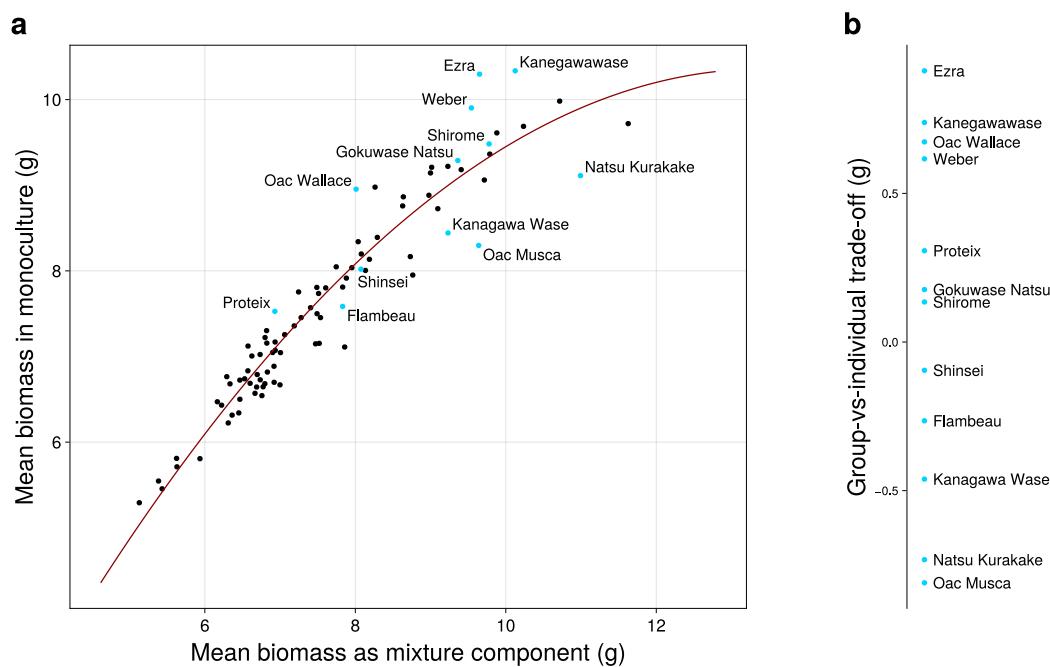


Figure 4.2: (a) Group-vs-individual (GI) trade-off of experimental population as calculated from biomass data from GI-experiment. In blue: the 12 genotypes selected for subsequent experiments. (b) The selected genotypes for subsequent experiments, spanning over the entire range of the GI trade-off.

(S.E. 0.14) and 0.21 (S.E. 0.16), when calculated on aboveground biomass and seed yield respectively. This suggests that it can potentially be targeted in breeding efforts.

**Accidental drought experiment.** More competitive genotypes tended to be more drought susceptible than less competitive genotypes (odds ratio for GI of focal plant: 0.6, 95% CI [0.31, 1.16]). Plants tended to show lower drought susceptibility with decreasing competitiveness of their neighbor plant (odds ratio for GI of neighbor plant: 0.61, 95% CI [0.35, 1.09]). The results were not statistically significant ( $p$ -values for GI of focal and GI of neighbor were 0.15 and 0.12 respectively). This was not surprising, given the experiment was not designed to investigate the effect of GI on drought susceptibility. However, the direction of the effect suggested that less competitive genotypes and plants growing in community with less competitive genotypes to be less prone to drought stress (Figure SM 4.8).

**Drought Experiment.** Less competitive genotypes had a higher water-use efficiency than more competitive genotypes ( $F_{1,10} = 6.88$ , ANOVA  $p = 0.025$ ; Figure 4.3a). Moreover, less competitive genotypes had a lower probability of displaying severe drought symptoms compared to more competitive genotypes ( $\chi^2_1 = 5.192$ ,  $p = 0.023$ ; Figure SM 4.9). Finally, plants growing next to a competitive neighbor had a higher probability of exhibiting severe drought symptoms (orange line on the left of Fig. 3b), while plants growing next to a more cooperative neighbor had a higher probability of displaying only low drought symptoms ( $\chi^2_1 = 4.272$ ,  $p = 0.039$ ; green line on the right of Figure 4.3b).

**Genome-wide association study.** GWAS of GI using a Bonferroni-corrected significance level of  $\alpha = 0.05$  did not result in the identification of any significant peaks (Figure 4.4a). This was not surprising due to the high number of analyzed SNPs and the relatively low number of genotypes in our population. After scanning the genome for regions with a high accumulation of markers with medium to high signal we found small significant genomic regions for both GI based on aboveground biomass and GI based on seed yield on multiple chromosomes. We focused on region 9'218'000 - 9'270'000 of chromosome 12 that revealed the highest signal of the regions identified in both the GWAS using GI based on biomass and in the one using GI based on seed yield (Figure 4.4b). The SNP at position Chr12:9258726 explained 8.26% of the total genetic variance for GI based on biomass, and genotypes carrying a thymine (T) at this position had a significantly lower GI compared to genotypes carrying a cytosine (C) at the same position ( $F_{1,80} = 7.033$ , ANOVA  $p < 0.01$ ; Figure 4.5). After screening the soybean reference genome, *Glyma.12G103500* (position 9250665 - 9254646 on chromosome 12) was identified as the closest gene to the most

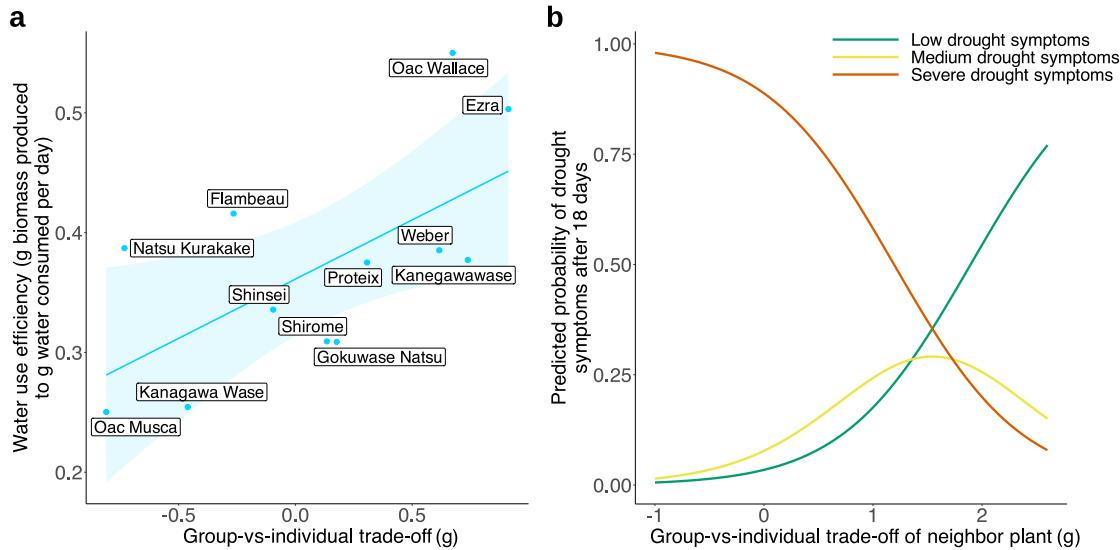


Figure 4.3: (a) Water-use efficiency correlates with group-versus-individual (GI) trade-off. (b) Drought tolerance increases with decreasing GI trade-off of neighbor.

significant SNP in our analysis. This gene is an ortholog of dwarf in light 1 gene (DFL1), coding for an auxin-responsive Gretchen Hagen 3 (GH3) family protein in *Arabidopsis thaliana*. *A. thaliana* genotypes over-expressing DFL1 are characterized by a shorter hypocotyl when grown in light compared to their wild-type (Nakazawa et al. 2001). Moreover, auxin-responsive GH3 proteins are known to link stress adaptation with growth regulation in *A. thaliana* (Park et al. 2007), as well as to affect drought tolerance in rice and cotton (Du et al. 2012, Kirungu et al. 2019). In *Arabidopsis*, the inactivation of the orthologue of *Glyma.12G103500* and seven of its closest homologs has been described to confer drought and salinity tolerance to plants (Casanova-Sáez et al. 2022).

#### 4.4. Discussion

We used a pot experiment encompassing 90 highly diverse soybean genotypes, each exposed to a range of “social neighborhoods”, to classify genotypes along a group-vs-individual trade-off into more cooperative and more competitive types. In agreement with the results from a similar experiment in the model plant *Arabidopsis thaliana* (Wuest et al. 2022), our results showed that genotypes vary in competitiveness, and that our population contains some relatively competitive and some relatively cooperative genotypes. In a validation experiment that was accidentally affected by a summer drought, we serendipitously discovered that the more competitive genotypes

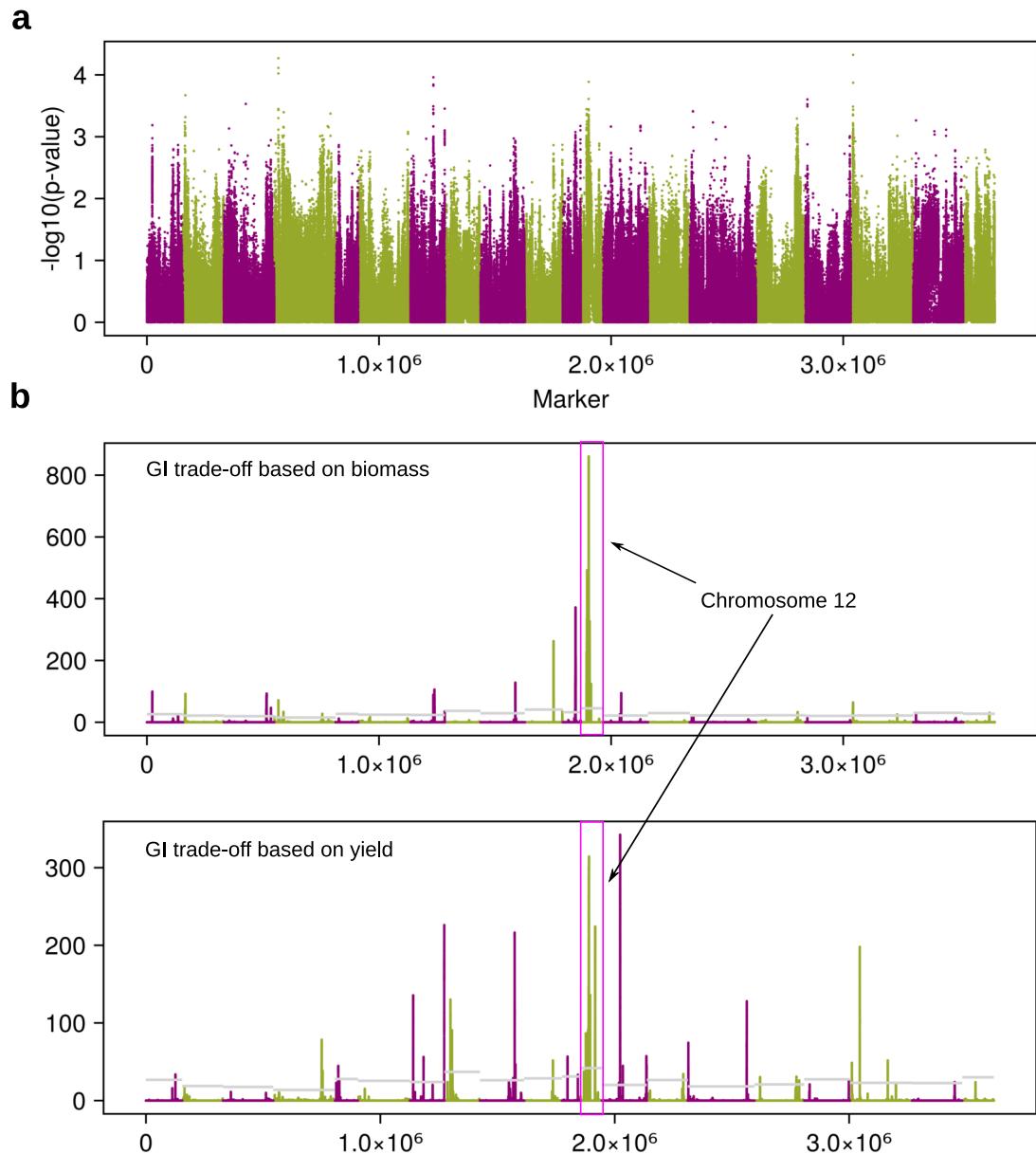


Figure 4.4: (a) Manhattan plot of GWAS using group-versus-individual (GI) trade-off (based on biomass) as response (no significant SNP). (b) Local score plot with chromosome-based significance thresholds. The pink rectangle highlights chromosome 12, which was found to have the highest significance among the regions significant both in GI based on biomass and GI based on yield.

suffered more from drought than the more cooperative genotypes. Therefore, we investigated the correlation between competitiveness and drought susceptibility more closely and found that more cooperative genotypes had a higher water-use efficiency and lower drought susceptibility, and this effect of the focal plant manifested also in the neighbor plant. We also identified a genomic

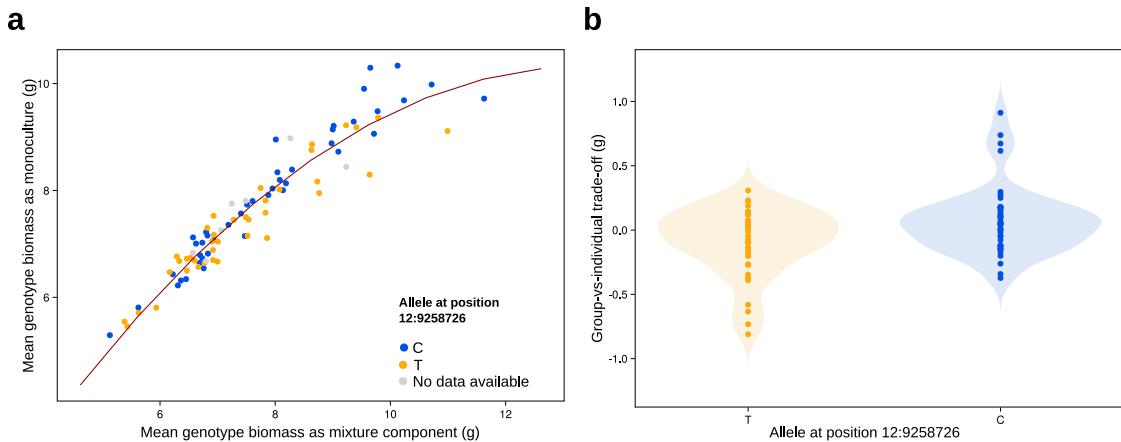


Figure 4.5: (a) Group-vs-individual trade-off colored by variant at position 9258726 on chromosome 12. (b) Effect of allele variant on group-versus-individual trade-off (based on biomass).

region that was associated with variation in competitiveness. The region contained a gene whose homologues in other plant species have been described to be involved in growth regulation, or to render plants more susceptible to drought (Park et al. 2007, Du et al. 2012, Kirungu et al. 2019, Casanova-Sáez et al. 2022).

### Competitiveness matters

Over the last century, there has been an important increase in yield production of the major food crops worldwide, in particular rice, wheat, and maize (Calderini and Slafer 1998, Kucharik and Ramankutty 2005, Xu et al. 2020). This yield increase has been termed the Green Revolution and is attributed to several causes, one of which is the plants' ability to efficiently grow under an increased nitrogen fertilizer regime without lodging. The introgression of dwarfing genes from exotic into elite germplasm in rice and wheat has contributed greatly to this by reducing plant height (Hedden 2003, Makino 2021). Alongside with these morphological improvements, an important component of the Green Revolution which has long time been underestimated was the possibility to massively increase planting density per area (Liu et al. 2020). In fact, common-garden experiments in maize showed that the historic breeding progress is most apparent under high planting densities, while the increase in yield under low planting densities was relatively small over the last century (Duvick et al. 2004).

Natural selection acts predominantly at the level of the individual, but the agricultural focus lies on the yield of a plant group, making little reference to the contribution of the individual plant to total yield. Since natural selection has been optimizing individual plants and their

traits for thousands of years, it seems implausible that there is much potential for trade-off free improvements of traits increasing individual fitness by breeding (Denison 2012, Weiner 2019). At the same time, as natural selection doesn't normally act on groups of plants, it is plausible to assume that there is high potential for breeding in the optimization of traits which increase group performance, as long as these traits are not also increasing individual performance (as they otherwise would already have been selected for). This scenario is likely to occur in response to biotic interactions, especially competition (Weiner 1990).

Higher planting densities will result in stronger competition for limiting resources between plants, which can result in allocation of resources into competitive structures rather than into harvestable products. Indeed, the soybean yield increase over the last century was smaller than the one for grain crops, and yield responses of modern soybean varieties to increased planting densities are saturating at relatively low levels already (Kokubun 1988, Carciochi et al. 2019). Thus, elite soybean varieties still seem to spend a high amount of their resources in response to competitive environments (Suhre et al. 2014). Efforts to elevate yields through increased planting densities should therefore focus on the reduction of competitive interference between neighboring plants (Kokubun 1988, Denison 2012, Weiner 2019, Wuest et al. 2022). To do so, we need a measure that allows researchers and breeders to quantify genotypic competitiveness. Here we show that the GI trade-off, first described by Wuest et al. (2022) for the model plant *A. thaliana*, is apt to quantify competitiveness in a major crop plant. Moreover, being agnostic with respect to the underlying mechanisms, the GI trade-off allows a quantification of competitiveness without having to make assumptions about the mechanisms of competition. This and related methods would be especially well-suited to identify relevant traits for breeding in gene banks that harbor large and highly diverse germplasm collections, a considerable part of which is absent from elite soybean breeding populations.

### Drought susceptibility matters

Drought is a major threat for global crop production, and, at least locally, severe drought periods are likely to increase in frequency and intensity due to anthropogenic climate change (Loukas et al. 2008, Cook et al. 2018). It is therefore of utmost importance to better understand how crop plants differ in their mechanisms to handle drought stress and to develop drought tolerant varieties of major food crops. Drought tolerance is often associated with conservative growth strategies, whereby slower growth and reduced stomatal conductance promote tolerance (Ouédraogo et al. 2013, Rodrigues et al. 2019, Gupta et al. 2020). These traits trade off with competitive strategies like faster growth and higher allocation to leaf area that leave individuals

more vulnerable to drought. Given this tolerance-competition trade-off, the improvement of crop production under drought could be a highly desirable side-effect of breeding strategies aimed to decrease within-field plant competition (Yang et al. 2020, Feng et al. 2022).

Here, we show that in soybean water-use efficiency and drought tolerance are correlated with the GI trade-off, i.e. more cooperative genotypes exhibited higher water use efficiencies and suffered less from water restrictions, and so did their neighbors. This may represent a fitness trade-off between resource acquisition and stress resistance, similar to that described for other organisms (see e.g. De Tombeur et al. 2023, describing a growth-defense trade-off in rice). However, as GI was calculated in a well irrigated experiment and the models correlating GI and drought tolerance were corrected for plant biomass, it seems unlikely that different resource acquisition strategies are directly driving stress susceptibility responses or vice-versa. Rather, we assume that both resource acquisition strategies and stress tolerance responses might be mediated by a common, not yet identified mechanism.

### **Candidate gene linking social strategies to drought susceptibility**

Upon our GWAS we identified a candidate gene coding for a GH3 family protein, known to be involved in both growth regulation and stress adaptation in multiple plant species (Nakazawa et al. 2001, Park et al. 2007, Du et al. 2012, Kirungu et al. 2019). A recent study showed that the inactivation of the entire group II GH3 gene family confers an increased degree of drought tolerance in *A. thaliana* (Casanova-Sáez et al. 2022). This is an at first sight surprising finding, since it may appear counterintuitive that genes rendering plants more susceptible to an important abiotic stress would be maintained throughout evolutionary history. Therefore, we can assume that drought susceptibility represents a trade-off with another important trait that the genes are relevant for. The mutants also showed higher lateral root density (without penalty for the primary root system), longer hypocotyls and increased concentrations of abscisic acid (ABA) in the shoots under stress free conditions. Elevated ABA concentrations promote stomatal closure (Chen et al. 2021b, Bharath et al. 2021, Postiglione and Muday 2023), shifting the water-carbon trade-off associated with stomatal regulation towards a more conservative growth strategy. Moreover, stomatal strategies may be altered not only by drought but also by competition (Zenes et al. 2020). These observations make it plausible that the identified gene *Glyma.12G103500* is a key determinant in competitive interactions among plants. Additional work is needed to better understand the relevance and function of the identified gene in shaping social strategies and drought susceptibility in soybean genotypes.

### **Conclusion**

Using a diverse soybean population, we identified “team players” and “selfish” genotypes through competitive games. Remarkably, our measure for genotypic competitiveness correlated strongly with genotypic water-use-efficiency. Cooperative plants not only had a higher water-use-efficiency and showed lower drought symptoms themselves, their neighbors also exhibited lower drought stress symptoms. This indicates that the cooperative genotypes were sharing water with their neighbors. A genomic analysis led to the identification of a gene involved in the auxin regulation pathway which might play a crucial role both in shaping a genotype’s social strategy and its drought susceptibility. Further research is needed to validate and better understand the role of the identified gene. The innovative approach presented here offers an exciting, complementary path to modern breeding, promising more resilient and sustainable agriculture.

## 4.5. Supplementary materials

Table 4.1: List of genotypes used throughout the thesis with USDA PI number (when available), genotype name, genotype origin (collection site for landraces and breeding site for elite varieties) and maturity group.

PI Number	Genotype Name	Origin	Maturity Group
PI 548524	Weber	USA	I
PI 317336	Shinsei	Japan	O
PI 423954	Shirome	Japan	O
PI 548382	Manitoba Brown	Canada	OO
PI 603424 A	Ben di huang dou/ZDD007871	China	O
PI 467343	Yan-nong No. 2	China	I
PI 548593	Maple Arrow	Canada	OO
NA	Protéix	Switzerland	OO
NA	Galice	Switzerland	OOO-OO
PI 417187	Natsu kurakake	Japan	O
NA	Ezra	Canada	O
PI 507354	Tokei 421	Japan	I
NA	Paprika	Switzerland	OO
NA	No 90093	Switzerland	OO
NA	No 90113	Switzerland	OO
NA	Gallec	Switzerland	OOO
PI 290120	Amurskaja 42	Russia	OO
PI 504499	Ta li tsao shen wu tou	Taiwan	OOO
PI 548296	Agate	Japan	OO
PI 549054	Ao 75	Japan	OO
PI 548417	Soysota	Italy	I

PI Number	Genotype Name	Origin	Maturity Group
PI 194626	634-7-64-1	Sweden	O
PI 194643	748-5	Sweden	O
PI 238920	Brillamayerova Giesenska	Czechoslovakia	O
FC 30685	Cha Kura Kake	Japan	O
PI 548544	Crest	Canada	OO-O
PI 189963	Geant Vert	France	OOO
NA	Himekogane	Japan	O
PI 317332	Isuzu	Japan	O
PI 442025	ISZ 14	Hungary	O
PI 467323 A	Jiu nong 13	China	O
PI 458531	Ke shuang	China	O
PI 506900	Kitamusume	Japan	O
PI 548545	Merit	USA	O
PI 180519	N.134	Germany	O
PI 189952	N15 Nordeutsch Schwarts Matt	France	O
PI 189940	N3 Rouest 13 A1 2	France	O
PI 180517	No. 184	Germany	OO-O
PI 548535	Norman	USA	O
PI 290156	Wielnska Brunatna	Hungary	O
PI 290158	Ke schuan	China	OO-O
PI 347540 B	(F59-244) B	Romania	O
PI 347540 C	(F59-244) C	Romania	O
PI 347551	S-16	Romania	O
PI 347556	F60-156	Romania	O
PI 361110	Secca	Italy	OO-O

PI Number	Genotype Name	Origin	Maturity Group
PI 361111	Shimshirazu nr. 1	Japan	O
PI 361114	Trzic Rana	Romania	O
PI 361123 A	Warszawska	Poland	O
PI 372403 B	(Caloria) B	Austria	O
PI 399074	PI 399074	Korea	O
PI 548607	Portage	Canada	OO-O
PI 442037	Prawda	Russia	O
NA	Keshan 1	China	O
NA	Heike 49	China	O
NA	Suinong 24	China	O
NA	Fengshou 26	China	O
NA	Kenfeng 13	China	O
NA	Mentor (EGA 702)	France	OO
PI 538406	Sapporo Midori	Japan	OO
PI 358321 B	PI 358321 B	China	OO
PI 513382	Glenwood	USA	O
PI 548539	Comet	Canada	O
PI 416890	Gokuwase natsu daizu	Japan	O
PI 548643	Maple Glen	Canada	OO
Coll 10795	Naya	Canada	O
PI 548398	Pagoda	Canada	OO-O
PI 548582	McCall	USA	OO-O
PI 548638	OAC Libra	Canada	O
PI 548644	OAC Musca	Canada	O
PI 153262	Roumanie	Belgium	O

PI Number	Genotype Name	Origin	Maturity Group
PI 548379	Mandarin (Ottawa)	China	O
PI 229330	Kanegawawase	Japan	O
PI 506832	Kanagawa Wase	Japan	O
PI 507201	Saishuu Daizu	Japan	O
PI358320	PI 358320	China	OO-O
PI 548378	Mandarin	China	O-I
Coll 10877	OAC Wallace	Canada	O-I
PI 232999	No. 155/49	Germany	OO
PI 240079	6-3-1	Japan	OO
PI 360958	Kamishumbetzu	Japan	OO
PI 360965	Toshi 7910	Japan	OO
PI 423869	Tenpoku wase	Japan	OO
PI 507232	Shimo Shirazu	Japan	OO
PI 548325	Flambeau	Canada	OO
Coll 10712	AC Albatros	Canada	OO
FC 30687	Kosodiguri Extra Early	Japan	OOO
P I506895	Kitahomare	Japan	O-I
PI 548504	Altona	Canada	OOO
PI 567787	OAC Vision	Canada	OOO

**Substrate used in all experiments described in this chapter:**

RICOTER Erdaufbereitung AG, Frauenfeld, Switzerland

25% Landerde top max. Grösse 15 mm / 60% Weisstorf 0-30 mm / 15% Perlit 2/6 mm

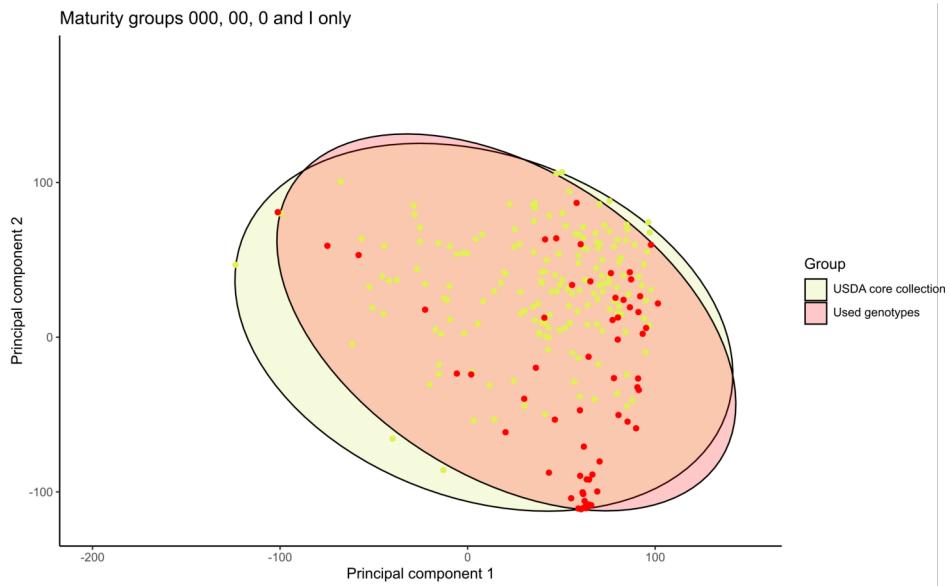


Figure 4.6: Principal component analysis based on 40k SNP data (available from USDA) showing the genetic diversity of the early-maturing genotypes in the USDA core collection and the genotypes used in this thesis (for which SNP data was available through USDA). The high overlap shows that the population used in this thesis covers a considerable part of the overall genetic variation of the early-maturing genotypes present in the USDA core collection.

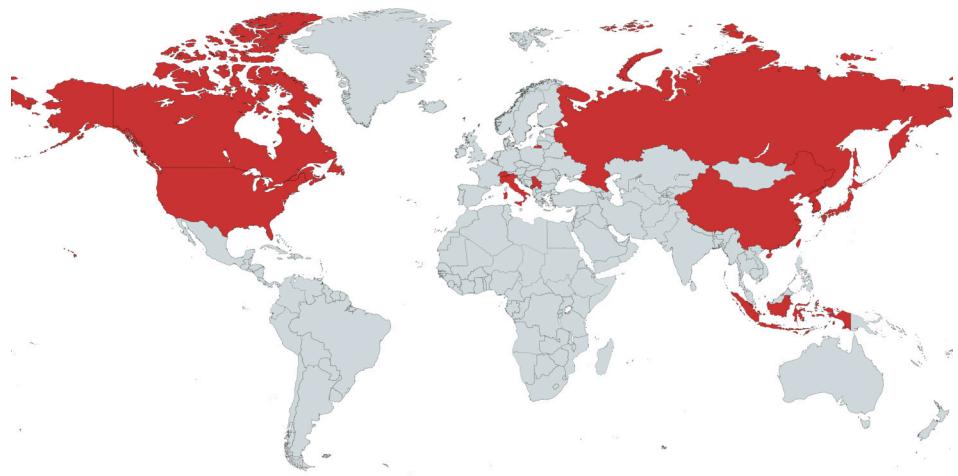


Figure 4.7: Map showing all countries of origin of the 90 soybean genotypes used throughout the thesis.

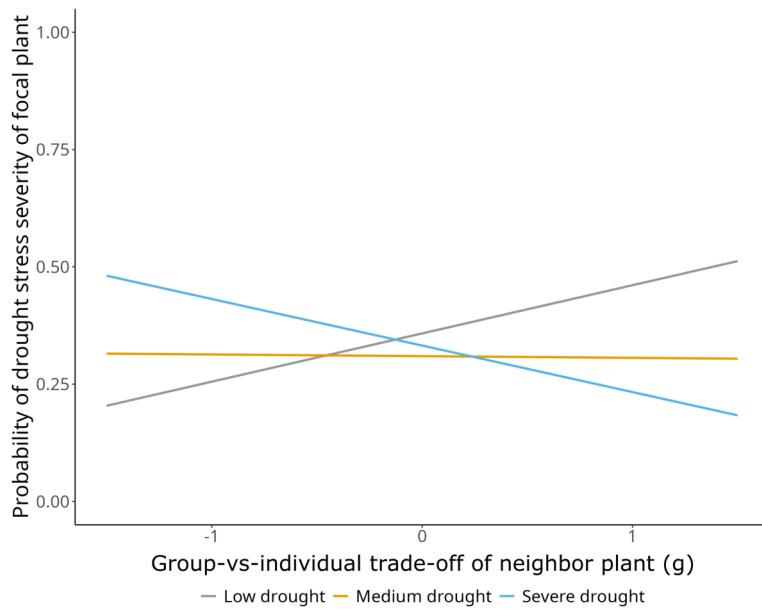


Figure 4.8: Analysis of accidental drought experiment: drought susceptibility decreases with increasing group-versus-individual trade-off of neighbor plant.

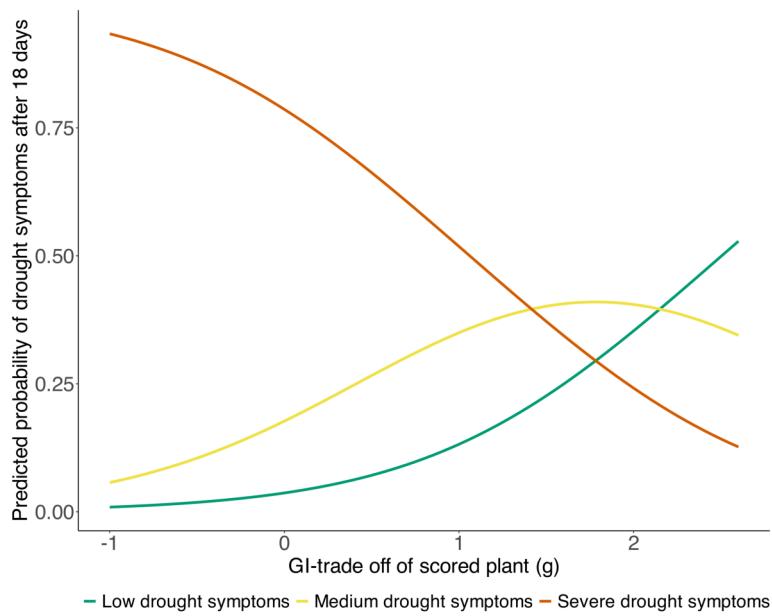


Figure 4.9: Drought susceptibility decreases with increasing group-versus-individual trade-off.

## **5. Ecological principles to guide the development of crop variety mixtures**

Emanuel B. Kopp<sup>1, 2</sup>, Pascal A. Niklaus<sup>2, 3</sup>, Samuel E. Wuest<sup>1</sup>

1: Agroscope, Group Breeding Research, Müller-Thurgau-Strasse 29, 8820 Wädenswil, Switzerland

2: Department of Evolutionary Biology and Environmental Studies and Zurich-Basel Plant Science Center, University of Zurich, Winterthurerstrasse 190, 8057 Zurich

3: University of Zurich Research Priority Program on Global Change and Biodiversity, University of Zurich, 8057 Zurich

Published as:

Kopp E B, Niklaus P A, Wuest S E (2023) Ecological Principles to Guide the Development of Crop Variety Mixtures. *Journal of Plant Ecology* 16:rtad17. <https://doi.org/10.1093/jpe/rtad017>

## **Abstract**

Crop variety mixtures can provide many benefits, including pathogen suppression and increased yield and yield stability. However, these benefits do not necessarily occur in all mixtures, and the benefits of diversity may be compromised by disadvantages due to increased crop heterogeneity. In-field development of mixtures by assembling many combinations of crop genotypes without prior expectation about which genotypes need to be combined to produce well-performing mixtures results in prohibitively large designs. Therefore, effective tools are required to narrow down the number of promising variety mixtures, and to then identify in experiments which of these deliver the highest benefits. Here, we first review current knowledge about the mechanisms underlying effects in ecological diversity experiments and in current agricultural applications. We then discuss some of the principal difficulties arising in the application of this knowledge to develop good variety mixtures. We also discuss non-conventional approaches to solve some of these issues. In particular, we highlight the potential and limitations of trait-based methods to determine good variety mixing partners, and argue that non-traditional traits and trait-derived metrics may be needed for the trait-based approach to deliver its full potential. Specifically, we argue that good mixing partners can be identified using modern genetic and genomic approaches. Alternatively, good mixtures may be obtained by combining varieties that respond differently to environmental variation; such varieties could easily be identified in standard variety testing trials. Preliminary analyses show that niche differences underlying the different environmental responses can indicate functional complementarity and promote mixture yield and yield stability.

## 5.1. Introduction

Increasing crop diversity is a known path towards a more sustainable and stable crop production (Zhu et al. 2000, Yang et al. 2018, Renard and Tilman 2019, Cassman and Grassini 2020; Li et al. 2020, 2021b). Crop diversification can be implemented at different levels, each with distinct advantages and constraints: at the very large scale, a larger diversity of crops that are grown within or between farms will contribute to food security by temporally stabilizing crop yields at the regional to global level (Renard and Tilman 2019). Within fields, intercropping of different plant species has traditionally been used to leverage benefits of beneficial interactions among crop species (mainly reduced competition, pathogen suppression, and facilitation; Brooker et al. 2015, 2023). A somewhat similar effect can be achieved by crop rotation, i.e. by diversification in time instead of space. Finally, diversification can also be achieved within a field by co-cultivating different varieties of the same crop – which is the topic on which we focus here. Variety mixtures promise benefits similar to those seen in crop species mixtures, yet appear more compatible with modern mechanized agriculture and yield processing because they exhibit only limited heterogeneity in relevant agronomic traits (Wolfe 1985, Finckh et al. 2000, Mundt 2002, Newton et al. 2009, Barot et al. 2017, Wuest et al. 2021). Interest in the use of such variety mixtures in Western agriculture has waxed and waned over the years: the topic has a long research history (Frankel 1939, Gibson 2022), saw an interest peak in the 1980ies, specifically driven by phytopathological research on diversity-mediated disease suppression, and is currently seeing renewed attention (Finckh et al. 2000, Borg et al. 2018, Wuest et al. 2021). In this review, we focus on variety mixtures and the current understanding of the mechanisms that underlie the benefits of such mixtures. We highlight potential bottlenecks that arise in the development of mixtures that excel on the market, in particular the problems encountered in the development and implementation of predictive methods. Finally, we re-visit ecological theory and present, based on ecological principles, a range of potential solutions to particular challenges.

## 5.2. From monocultures to mixtures: known benefits and future challenges

Today, breeding programs mostly focus on developing pure lines (or hybrids) that are typically deployed in monocultures consisting of a single genotype. Breeders improve varieties by directional selection within populations that are typically created by crossing elite breeding material, which allows the accumulation of an ever-increasing number of beneficial alleles in the lines

released to the market. This breeding strategy works well and has greatly contributed to increased yield and improved quality characteristics of modern varieties (Duvick et al. 2004, Voss-Fels et al. 2019). The breeding process itself may be associated with diversity bottlenecks, whereby selection against undesirable alleles results in the loss of rare but potentially beneficial alleles. This can result in a reduced representation of global genetic diversity in elite line breeding populations (Litrico and Violle 2015, Louwaars 2018, Bourke et al. 2021), and the overall genetic diversity within a crop may drop over time (Reif et al. 2005, Bonnin et al. 2014 but see Chai et al. 2022). This problem is addressed by periodically re-introducing genetic diversity into breeding populations (Reif et al. 2005, Kovach and McCouch 2008, Jannink 2010, Gorjanc et al. 2016), which over time can even increase the levels of diversity among popular varieties and improve the buffering of a crop against new challenges such as emerging diseases (Chai et al. 2022).

The use of pure lines grown in monoculture rests on the assumption that for a crop there is an optimal value for a trait, and that this value remains the same or at least very similar across multiple sites and years and under different management practices. However, this is not necessarily the case. It is now established that increasing within-field phenotypic variation (i.e. increasing the range of trait values) can improve across-year and across-site crop stability through a portfolio effect, and also deliver other benefits ranging from higher yield to reduced requirements for management or input (Litrico and Violle 2015). High within-field genetic variation also provides, on average, protection against disease epidemics (Box 5.3). However, positive mixture effects are not guaranteed (Beillouin et al. 2021). Indeed, some experiments have also found that genetic diversity can increase the abundance of diseases, or diversity benefits can vary strongly between experiments and even within studies (Smithson and Lenné 1996, Cowger and Mundt 2002, Mundt 2002, Gibson and Nguyen 2021, Montazeaud et al. 2022). In accordance, meta-analyses, while supporting the idea that mixture effects are on average positive, also show that effects are highly variable (Smithson and Lenné 1996, Kiær et al. 2009, Huang et al. 2012, Koricheva and Hayes 2018, Reiss and Drinkwater 2018, Borg et al. 2018, Kristoffersen et al. 2020b, Gibson and Nguyen 2021). Specifically, they show that (1) overyielding benefits average around 2-4% (slightly higher effects are found under low-input conditions); (2) diseases are, on average, suppressed in mixtures, but effect estimates are difficult to extrapolate from experiments to real-world applications, because many epidemiological phenomena only appear at scales larger than the ones used in experiments (Box 5.3); and (3) mixture yield stability is often slightly improved compared with monocultures. Increased yield and stability are, however, often more pronounced in mixtures under more variable conditions or disease pressure.

To summarize, the transition from traditional cropping systems that utilized genetically diverse populations (landraces) to monocultures of pure lines decreased within-field heterogeneity (Harlan 1975, Dawson and Goldringer 2011). While management and production were simplified in this process, the potential advantages that diversity provides were lost. This may become particularly critical when facing future environmental challenges such as climate extremes and the emergence of novel pathogens (Bonnin et al. 2014). Indeed, breeders are concerned about the increasing difficulty to identify new genotypes that perform well across the whole range of their target environments (Dario Fossati, personal communication).

### **5.3. Box: Pathogen suppression – the ultimate ecosystem service?**

The use of genetic diversity to mitigate large-scale disease epidemics and to suppress local pathogen spread has been of interest for more than a century (Gibson 2022), and mixture-mediated reductions in pesticides can likely make a major contribution towards sustainable agriculture (Zhu et al. 2000, Mundt 2002). Much of the phytopathological work so far has focused on the question whether multilines, mixtures or heterogeneous populations suppress pathogens and pests better than do monocultures, and under what conditions (e.g. ratio of resistant to susceptible plants, planting densities, level of disease pressure or type of pathogen). Numerous studies now summarized in several meta-analyses and reviews have established that, on average, mixtures are effective at suppressing diseases, and increase yield under high pathogen pressure (McDonald et al. 1988, Smithson and Lenné 1996, Zhu et al. 2000, Finckh et al. 2000, Cowger and Mundt 2002, Mundt 2002; Kiær et al. 2009, 2012; Finckh et al. 2017, Reiss and Drinkwater 2018, Yang et al. 2019, Kristoffersen et al. 2020b). However, many questions still remain to be answered before a richer understanding or effective predictive methods for disease suppression can be reached. So far, a number of potential mechanisms for disease suppression in mixtures have been proposed (Finckh et al. 2000, Mundt 2002, Borg et al. 2018), for example epidemiological effects through a reduced density of susceptible hosts, or even non-hosts that act as barriers for the spread of an inoculum. Further, an infection of a host with a pathogen strain that is incompatible may trigger increased host immunity, and therefore represent an early-warning signal and increased resistance against compatible pathogen strains. Competitive

interactions may also increase basal pathogen resistance, through mechanisms not fully understood (Barton and Bowers 2006, Pélišsier et al. 2021). Finally, resistant hosts may compensate for the losses suffered by susceptible neighbor plants, leading to higher fractions of resistant host tissue within a field (Finckh 1992). It is worth noting that disease suppressive effects not only arise in mixtures that were deliberately designed to do so (e.g. based on combining varieties with different resistance genes), but also in cases they were not (Kellerhals et al. 2003, Kristoffersen et al. 2020b), suggesting that plant-pathogen interactions are more complex than early models have implied (Flor 1971, Wu et al. 2018) A still understudied potential of mixtures is therefore their ability to simultaneously suppress multiple pathogens (Finckh et al. 2000, Wuest et al. 2021), e.g. when each component of a mixture shows resistance to a different pathogen. Such mixture properties could be specifically used in breeding programs, and future research may extend the focus on the general nature of host-pathogen network structures (i.e. interactions of varieties and pathogens or pathogen strains) that allow for particularly effective disease suppression in mixtures. Provided many studies that have shown a disease-suppressive effect of mixtures compared to the corresponding monocultures, the research focus should now shift towards the underlying mechanisms, using methods that compare different mixtures, ideally in large numbers, to determine or confirm predictions when diversity effects differ in size. Diallel or factorial designs, typically used in hybrid breeding, are ideally suited for this purpose (see main text and Figure 5.3). However, some epidemiological control mechanisms, such as reduced host density or barrier effects are typically dependent on the scale at which mixtures are grown and evaluated, and on the mode of pathogen dispersal (Finckh et al. 2000, Mundt 2002): diversity effects on long-distance wind-dispersed wheat mildew or rusts are hard to score in small experimental evaluation plots due to large effects of external inoculum pressures. Such scale-dependencies are per se an important research topic that will likely benefit from support by epidemiological models (Garrett and Mundt 1999, Mikaberidze et al. 2015). Finally, repeated breakdowns of resistances upon the evolution of new virulent pathogen strains, also referred to as “boom-and-bust cycles” in plant pathology (Wolfe 1980, McDonald and Linde 2002, Mundt 2014, Brown 2015), can have devastating consequences for crop production and diminish the pool of resistance genes available for the breeder. Combining a variety of resistances at different spatial or temporal scales has been proposed as a solution to this problem (Finckh et al. 2000, McDonald and Linde 2002, Mundt 2014, Brown 2015; Rimbaud et al. 2018, 2021). Mixtures may contribute to such resistance

management strategies in the future, but investigating processes at such large temporal and spatial scales using experiments represents an enormous challenge (Garrett and Mundt 1999, Finckh et al. 2000, Kristoffersen et al. 2020a, Rimbaud et al. 2021).

#### **5.4. Challenges in breeding for variety mixtures**

Overall, existing research demonstrates that mixtures can provide tangible benefits, especially when good variety compositions can be identified. Which compositions are best depends on goals and applications, but in general the properties aimed for will be high yield and yield stability, a good suppression of pathogens, and a high quality of the product. However, developing mixtures adds another level of complexity to today's crop development programs. Breeders are already facing growing lists of breeding aims to address new problems such as increasing climatic uncertainty or the unprecedented spread of new pathogens (Fones et al. 2020, Xiong et al. 2021). Breeders are therefore wary of additional constraints such as the requirement to minimize genetic variation for traits for which uniformity is desirable (e.g. agronomic traits such as phenology including maturity date) but to maximize genetic variation for the traits that underpin the diversity benefits (Litrico and Violle 2015, Wuest et al. 2021). At the same time, concerns about conflicts of phenotypic diversity with market demands for uniform products are often unfounded (Mundt 2002, Newton et al. 2009, D. Fossati, personal communication). Mixture developers, however, focus more strongly on negative selection criteria (for example agronomic constraints that demand the exclusion of specific variety combinations like large maturity date differences, very different quality traits, etc.), further increasing the disincentive to develop variety mixtures at all. A general mechanistic understanding of diversity effects would likely change this: for example, a recent meta-analysis of wheat variety mixture experiments showed relatively higher overyielding values reported in studies from the 1980ies than in more recent ones, likely due to a strong focus on disease suppression mechanisms during that time period and leading the development of more disease resistant – and more productive – mixtures (Borg et al. 2018).

Another challenge lies in the combinatorial properties of mixtures: designs to combine pure lines into different variety mixtures quickly escalate because in the case of 2-component mixtures the number of possible combinations increases with the square of the number of components, and even more rapidly for more complex mixtures (Barot et al. 2017, Wuest et al. 2021). This makes it impractical to test even a fraction of all possible compositions. More efficient designs can alleviate this problem (see below), but finding variety combinations that result in a mixture benefit remains difficult and is often perceived as requiring larger efforts for field-testing than can be justified

based on the expected potential returns. For these reasons, the full potential of mixtures likely remains untapped, although mixtures promise to reduce pathogen epidemics and improve the crop's resource-use efficiency, goals that rank high on today's agendas. One key innovation to promote mixtures would therefore be to identify effective methods to design high-performing mixtures whose benefits outweigh the disadvantages. In the following, we therefore discuss the following questions: (1) what insights have been gained in ecological studies of species diversity effects and in intercropping, and (2) how could such insights suggest positive selection criteria for variety mixture development. We also identify specific applications of ecological principles that have not received much attention but could be of great use in practical mixture designs, especially when combined with current agronomic and breeding methods.

## **5.5. Ecological theory related to biodiversity benefits**

Positive species diversity effects on grassland productivity are well-documented: a large body of observational and experimental studies have shown that more diverse plant communities often are more productive and their productivity temporally more stable than less diverse plant communities (Proulx et al. 2010, Cardinale et al. 2011, Morin et al. 2014, Isbell et al. 2015). In general, the diversity-productivity relationship is “positive decelerating”, i.e. adding additional species results in decreasing marginal increases in productivity as community diversity increases. There is broad consensus that biodiversity effects can be attributed to some form of functional complementarity among species. This functional complementarity has been described in terms of niche differences among species, reasoning that distinct species compete less because they “partition niches”, and, together, constitute a broader “community niche” (Salles et al. 2009, Turnbull et al. 2016). However, the exact mechanisms that underlie such complementarity have proven very difficult to identify. An early idea was that abiotic resources such as soil nutrients are partitioned, but to date there is little direct empirical evidence for this (Silvertown et al. 1999, McKane et al. 2002, von Felten et al. 2012), despite decades of research, except for simple and relatively obvious cases (e.g. symbiotic nitrogen fixation in legumes, see Spehn et al. 2002). Another idea is that plant community members “partition” interactions with enemies such as herbivores and pathogens, i.e. enemies are specialized to different hosts. There is evidence that this supports biodiversity effects (Maron et al. 2011, Schnitzer et al. 2011, Mikaberidze et al. 2015, Huang et al. 2022), but also here many aspects remain unclear. Species may further facilitate the growth of other species, for example by environmental niche construction (Wright et al. 2017). Overall, subsuming all these (and likely also further) mechanisms in the concept of ecological niches results

in a niche concept that is highly abstract (Chase and Leibold 2003). It is useful in mathematical models because many phenomena can be parameterized using the same form (e.g. coefficients characterizing species interactions). However, the concrete niches and niche dimensions remain difficult to quantify in practice.

Because niches are so difficult to quantify directly, an appealing alternative is to infer niches indirectly from plant traits (Box 5.6). The underlying reasoning is that niche differences must be related to phenotypic differences that reflect the capabilities of a species to “deal” with the challenges posed by its abiotic and biotic environment. While this certainly is true, a number of difficulties arise. First, the traits considered may not be aligned with the fundamental niche dimension of interest (Kunstler et al. 2012, 2016; Funk and Wolf 2016, D'Andrea and Ostling 2016, Blonder 2018). For example, a shift along a specific niche dimension that underpins complementarity may result in a whole suite of trait changes, and these trait changes may be nonlinearly related to the original shift along the niche axis (Hoogenboom and Connolly 2009, Porter and Kearney 2009). Second, traits typically form syndromes of correlated changes because they together mediate an environmental adaptation (Wright et al. 2004, Díaz et al. 2016), or because trade-offs exist in terms of the underlying physiology, morphology, or genetics. Third, not all trait differences may be functionally important, at least not in the studied context. And finally, there may be traits that are difficult to observe because they are related to specific physiological functions that are not reflected in morphological differences. One example may be different root uptake kinetics for nutrients. Probably for all these reasons, trait-based approaches have not lived up to the expectations with which they were originally adopted in experimental biodiversity research. Today, there is consensus that single traits only have limited predictive power (Plas et al. 2020), and that niche complementarity likely is higher-dimensional in terms of the associated trait space. Also, it may be that the specific traits mediating complementarity differ between different species compositions of mixtures, for example between different species pairs considered (Kraft et al. 2015).

### **5.6. Box: Trait-based methods in ecology**

Trait-based methods allow to establish general rules applicable to a broad range of ecological patterns and processes, for example those underlying the assembly and functioning of plant communities (Violle et al. 2007, Grime and Pierce 2012, Shipley et al. 2016). A prominent example is the leaf economic spectrum (LES) that describes a continuum of plant growth

strategies that is readily described by six functional leaf traits related to fundamental resource allocation trade-offs (Wright et al. 2004). This relationship holds across a wide range of habitats, despite the vast diversity of species and their often idiosyncratic properties. Many readily observable plant traits also reflect adaptations to specific environments and are therefore interpreted as proxy for a species' niche position along specific dimensions of the Hutchinsonian niche (Figure 5.1). These traits successfully predict species abundances and habitat filtering processes when moving along gradients in e.g. precipitation, altitude, or latitude (Cadotte et al. 2015, Chalmandrier et al. 2017, Schellenberger Costa et al. 2017). Unsurprisingly, trait-based methods have become a central part of functional ecology (Viole et al. 2007, Shipley et al. 2016). However, the application of trait-based concepts to predicting the performance of mixtures of species (or varieties) has turned out to be surprisingly challenging. In general, traits seem to predict only a relatively small fraction of the observed variation in ecosystem functioning (Roscher et al. 2012, Ebeling et al. 2014, Plas et al. 2020). A reason may be that the traits used in these studies do not reflect the complementarity of co-existing species adapted to the same set of broad environmental conditions, and that the particular traits mediating complementarity are not well understood to date (Figure 5.1).

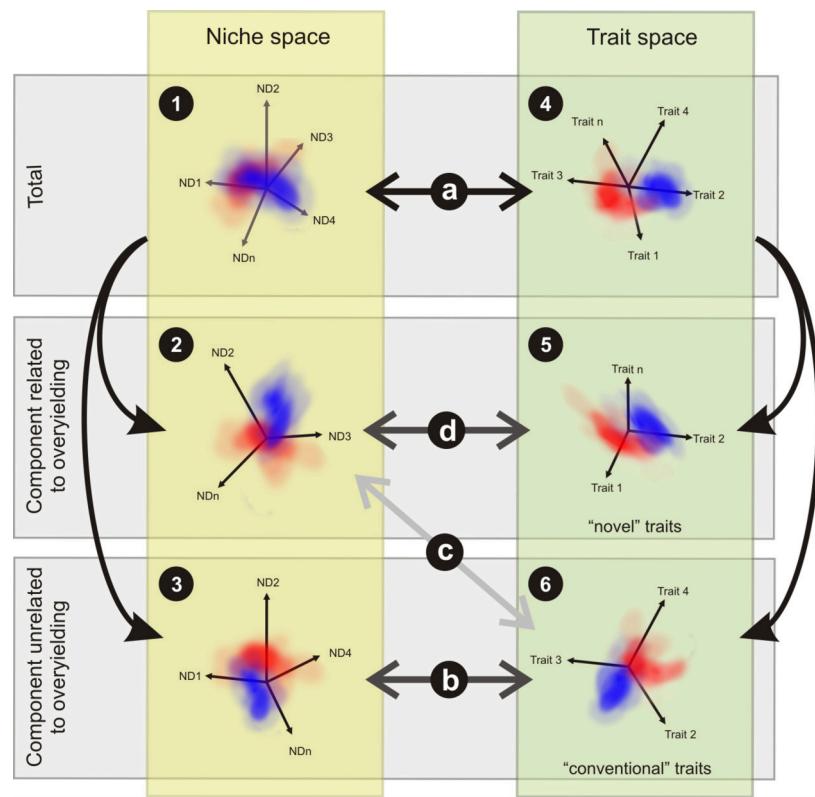


Figure 5.1: Niches and traits. Niches are often defined in a high-dimensional space spanned by axes ( $ND_1, ND_2, \dots, ND_n$ ) that correspond to a species' environmental requirements (e.g. climate, edaphic conditions, soil nutrients; Hutchinson 1978) and interactions with other organisms (e.g. competitors, consumers, symbionts, pathogens; Chase and Leibold 2003). The space ("hypervolume") that contains the conditions under which the species persists then defines its niche (1, red and blue volumes indicating the niches of two hypothetical species. Note that the sketch is conceptual because niche overlap cannot be adequately shown in two dimensions). Because the niche of a species is related to its interactions with its biotic and abiotic environment, which in turn depend on the species' functional characteristics, it is evident that niches have a correspondence (a) in trait space (4). In functional ecology, niche dimensions therefore often are approximated by sets of observable traits. When applying this idea to the analysis of biodiversity effects, difficulties arise: diversity effects evidently root in some form of niche partitioning; however, it is equally evident that not all niche differences promote diversity effects. Hence, only a component of the total interspecific niche differences underlies diversity effects (2); the remaining niche differences (3) are unrelated to diversity effects but matter in other contexts. Here, we posit that the suite of traits commonly used in ecological research (e.g. SLA, LDMC, leaf N contents) (6) strongly correlate (b) with niche dimensions related to fundamental growth strategies (e.g. leaf economic

spectrum) and environmental adaptations, but are not or only weakly related to niche dimensions underlying diversity effects (c). “Novel” trait metrics (5) that capture less-obvious characteristics of a species, and that are less related to environmental adaptations may better predict species differences that promote diversity effects (d). An example of such traits are pathogen susceptibilities and defense mechanisms, characteristics that are known to matter for enemy-related diversity effects (Box 5.3) but are not included in the “conventional trait set”.

## 5.7. Ecological principles applied to agriculture

How do these ecological concepts perform in agricultural applications? Interestingly, for species mixtures (intercropping), there are many cases in which the mechanisms responsible for the community-level benefits are relatively well understood (Brooker et al. 2015, Homulle et al. 2022, Schöb et al. 2023). An obvious case is the use of legumes, which can meet their nitrogen demand from atmospheric N<sub>2</sub>, thereby avoiding competition for soil nitrogen with other species. In the long term, N<sub>2</sub> fixation results in a general enrichment of the ecosystem with nitrogen and hence higher productivity also in non-fixers (Annicchiarico et al. 2019, Cowden et al. 2020, Engbersen et al. 2021). Many further examples exist: in agroforestry systems, trees provide shade to crops and increase water availability to the less deep-rooting crops by hydraulic lift – a case of facilitation by environmental niche construction (Brooker et al. 2015, Alagele et al. 2021, Homulle et al. 2022). Peas often are co-cultivated with cereal crops (Hauggaard-Nielsen et al. 2001). Here, the cereal prevents pea lodging and suppresses weeds by improving ground cover – mechanisms that could be termed niche construction and indirect biotic facilitation. In maize-faba bean intercrops, beans were found to mobilize phosphorus through local modifications of soil properties, thereby improving growth conditions also for maize (Li et al. 2007). Because in these (and similar cases) the main mechanism of interaction is relatively well-understood, such systems are amenable to improvement by “engineering approaches”.

At the intra-specific level, in general, genotypes of a species are more similar than different species. Nevertheless, positive effects of genotypic diversity have not only been documented in agricultural variety mixtures but also in ecological studies (Schmid 1994, Crawford and Whitney 2010, Cook-Patton et al. 2011, Bukowski and Petermann 2014). The challenge in identifying the mechanisms underlying these intraspecific diversity effects is that here the mixture components do not combine very dissimilar functions (e.g. a N<sub>2</sub>-fixer with a non-fixers, or a plant providing mechanical support to another). Instead, complementarity is rooted in traits that often are less visible to the human observer, because the varieties’ morphologies are relatively uniform.

A prominent example of diversity effects in variety mixtures are differences in the susceptibility of varieties to particular pathogens or pathogen strains. By “partitioning” the host-pathogen interactions, varieties will benefit from lower host densities available to specialized pathogens. Because the dynamics of epidemics are strongly density-dependent, a general suppression of diseases and associated yield losses often is observed (Reiss and Drinkwater 2018; Kristoffersen et al. 2020b, 2022). However, mechanisms beyond simple density-dependence also contribute much to disease suppression (Finckh et al. 2000, Mundt 2002 and Box 5.3).

## 5.8. Untapped potential for complementarity

In both ecological experiments and in agriculture, overyielding is frequently observed – yet, it seems rather the exception (e.g. Finn et al. 2013) than the rule that the specific mechanisms responsible for these effects are understood in detail (Barry et al. 2019, Plas et al. 2020), in particular in variety mixtures in which the observable differences between mixture components are more subtle (Montazeaud et al. 2018, 2020). What are the reasons for this difficulty to identify traits that cause overyielding? First, ecological research related to traits has largely focused on niche dimensions that are related to adaptations to environmental conditions or characteristic of a relatively narrow set of ecological strategies. Hence, these traits mainly describe the suitability of a species to a particular environment, for example, adaptation to aridity, or whether it is characterized by rapid, acquisitive or slow, conservative growth (Wright et al. 2004, Díaz et al. 2016). While these are important traits, we argue here that other traits that receive less attention likely are more important mediators of complementarity within a given environment (c.f. Box 5.6). Second, complementarity likely is related to a (possibly non-linear) combination of a larger trait suite, especially when the traits considered are not the ones closely linked to the functional differences that ultimately cause complementarity (Chacón-Labella et al. 2023). Third, within a given environment, the relevant trait differences may be relatively subtle and plastic, so that complementarity only becomes apparent in visible trait differences once the respective components interact. Based on all these considerations, we therefore posit that a large fraction of existing complementarity currently remains hidden because it remains inaccessible using the traditional trait-based methods. It thus seems that alternative approaches are required to uncover this potentially large “complementarity reservoir”.

## 5.9. Novel ideas to tackle an old problem

At the root of the problem discussed so far is the lack of clear hypotheses regarding the concrete biological mechanisms that underlie complementarity – therefore, the selection of traits is rather opportunity-driven than hypothesis-driven (i.e. the focus is on traits that are already well described and easy to measure). Given that plant interactions are complex and not particularly well understood, this is unlikely to change in the near future. In the end, this means that agnostic methods are required that broaden the trait domain investigated to date.

One idea is that genotype-by-environment interactions could be used to this end (Figure 5.2). In classical ecological work, niche overlap is frequently determined by counting the occurrence of individuals of different species in different microhabitats, or their use of different food sources (MacArthur 1958, Colwell and Futuyma 1971; Pianka 1973, 1974). Species that co-occur frequently across microhabitats, or that share the same food sources, are considered to occupy similar niches. In analogy, proxies of niche overlap between crop varieties could be derived from pure line yield changes across multi-year, multi-site variety trials, which are commonly performed during breeding and variety testing. The underlying rationale of using this data to identify complementary varieties is that year  $\times$  site combinations differ in many aspects (e.g. edaphic and climatic conditions, disease pressure), some of which remain unknown, and that the “specialization” of varieties is expressed when they are confronted with such different environments. Similarly, genotype-by-environment interactions in disease susceptibility, which can be considerable in field trials (Beukert et al. 2020), could be a sign that genotypes are specialized in their “pathogen niches” and resistant against different strains (which occur at different frequencies at different sites or years), or that some unknown environmental factors differentially modulate resistance traits in different genotypes. In a recent proof-of-concept analysis we have shown that overyielding of variety mixtures in Danish wheat trials can indeed be predicted from variety-specific pure stand yield-variation across years and sites (Wuest et al. 2021). This suggests that the specialization of varieties, and the resulting “division of labor” that supports overyielding, indeed is reflected in differences in environmental reaction norms, i.e. as differences in variety-specific trait changes across environments (Figure 5.2B). In the case of variety multi-environment trials, such reaction norms likely represent multiple environmental dimensions, which is maybe better aligned with multivariate theories of plant species coexistence (Kraft et al. 2015). This approach allows for the estimation of variety niches based on data readily available from standard variety development trials, by calculating differences in environmental reaction norms between genotypes instead of focusing on single trait differences between varieties within environments. Another advantage of

this approach is that differences in the function of varieties are assessed in the environmental range for which the cultivar is suitable. This somehow contrasts with approaches that focus on determining the environmental range of a species – here the focus is rather on the margins of the niches.

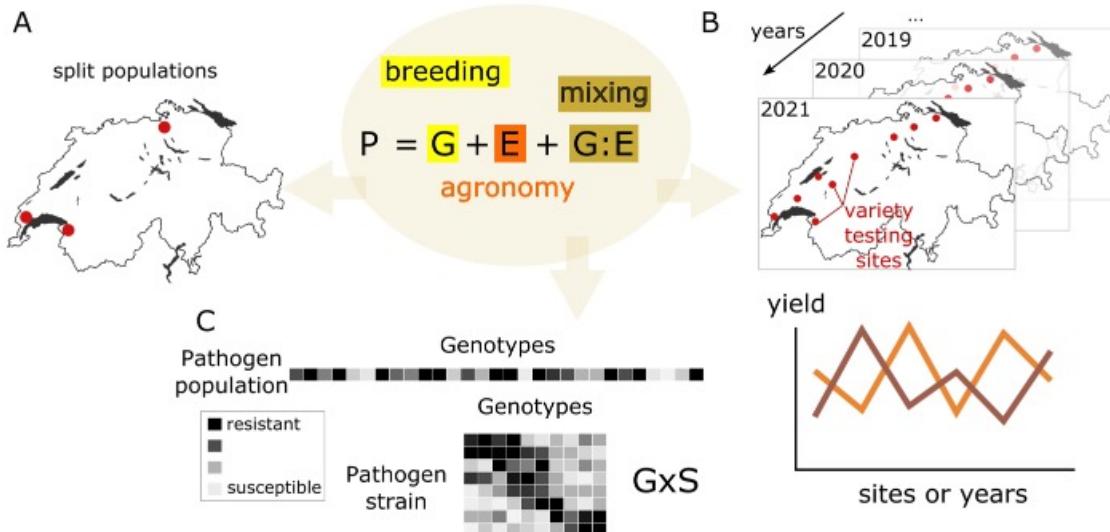


Figure 5.2: Ecological specialization of varieties through diverse strategies. Breeder's partitioning of phenotypic variation ( $P$ ) into various components highlights different levers for crop improvement.  $G$  represents a genotypic variance component, the additive part of which influences the response to selection that can be achieved through breeding.  $E$  represents an environmental variance component (e.g. variation between sites, year, or even management), which is also influenced by agronomic practices such as fertilization or irrigation. The  $G \times E$  component is typically considered a nuisance for breeders but could be exploited to some degree in mixture development. Three examples are given here, as follows: (A) Splitting a breeding population across sites with different pedoclimatic conditions or pathogen pressures etc, representing relevant niche axes along which genotypes could be specialized for optimal complementarity. (B) Asynchrony of genotypes across variety testing sites or years in multi-environment trials typically conducted in breeding and variety testing may stabilize mixture yields and even lead to overyielding or other mixture benefits, as described in Wuest et al. 2021. (C) Genetic resistances are typically evaluated upon infecting different genotypes with a pathogen population or a single pathogen strain. However, modeling work shows that genotypic specialization for different pathogen strains (i.e. how specific genotypes interact with specific pathogen strains – genotype-by-strain ( $G \times S$ ) interactions), or alternatively, genotype-by-environment interactions in disease susceptibility, should result in maximal mixture benefits and could also slow the evolution of new virulent strains.

This idea can not only be applied to mixture productivity (by increasing overyielding) but also to tackle the stability of mixture yields. A tenet in community ecology is that the greater stability exhibited by more diverse communities is driven by higher asynchrony among component species populations. Again, we argue that such asynchrony can be measured in analogy in pure stand plots assessed across different sites and years. A preliminary test we performed using the same Danish trial dataset indicated that this approach indeed predicted a significant amount of variation in mixture stability (Wuest et al. 2021). The idea to leverage such differences runs counter prevailing breeding practice. The phenotypic variation observed in breeding trials can be partitioned into genotypic variance, environmental variance, and variance due to genotype-by-environment (GxE) interactions. Breeders generally concentrate on the genotypic variance component, in particular its additive part, on which selection can operate and which drives breeding progress. GxE interactions within the defined production environment targeted by the breeders are seen as a nuisance because further niche breeding is not economical. However, niche differences and asynchrony as evidenced in growth differences of pure stands across different environments are drivers of such GxE interactions. In other words, mixture development could build on a component of phenotypic variance that is inaccessible and generally undesired in pure-line breeding. This offers the potential for important synergies between breeding and mixture development.

A fundamental alternative to focusing on traits as predictors of complementarity is to use genetic methods to identify the genetic underpinnings of community overyielding (Wuest and Niklaus 2018, Sato et al. 2021, Montazeaud et al. 2022, Wuest et al. 2023). In some pilot studies with mixtures of model plants, we have been able to link overyielding to genetic diversity across genotypes at particular loci. Such insights may be directly applicable in breeding. In the longer term, understanding the genetic basis of diversity effects may also lead to the identification of associated functional traits. In our studies, some diversity effects were mediated indirectly through soil factors, (e.g. Wuest and Niklaus 2018, Wuest et al. 2023), and associated with root physiology. These functional differences were not apparent in the widely-adopted “conventional” traits such as leaf angle, plant height, or specific leaf area, and if they had been apparent, then these differences likely would have been a consequence of modified growth rather than its primary cause. Once, in one way or another, causal mechanisms or specific trait combinations leading to complementarity have been identified, it will be relatively straightforward to select complementary varieties that optimize yield, yield stability, suppress diseases, or promote other ecosystem functions (Barot et al. 2017). With the leading niche dimensions identified along which varieties can be specialized for optimal complementarity (e.g. specific nutrients, soil properties,

pathogen communities), new selection strategies could be implemented in breeding programs. For example, breeding populations could be split among field sites that span the relevant niche axis, and divergent adaptation to local sites and conditions should then select for complementarity among these split populations (Figure 5.2A). Such niche breeding would also account for genotype-by-environment interactions that are normally deemed unfavorably by breeders, and which are often caused by trade-offs between alternative plant strategies. Directional selection in split populations will follow already established breeding methods but lead to local adaptations that can lead to within-mixture complementarity. Furthermore, populations can be genetically differentiated in analogy to genetic differentiation that is used in hybrid breeding schemes, where heterotic groups are bred separately before hybrid crosses. Modern genomic technologies could make such differentiation highly efficient (Technow et al. 2014, Zhao et al. 2015, Li et al. 2022).

Nevertheless, the combinatorial challenges associated with mixture testing remain: optimizing mixtures, regardless of the procedure, requires the screening of a large candidate pool of mixtures. In ecological diversity experiments, overyielding is assessed relative to the performance of monocultures of the components. For variety mixture development, this is inefficient because monocultures are not the communities of interest, yet monocultures need to be highly replicated to provide statistically stable benchmark values to assess overyielding. Also, in mixture development the goal has become to identify the best mixture in a set of mixtures, and no longer to quantify performance of these mixtures relative to monocultures not intended for cultivation. Both problems can be circumvented using diallel designs (Figure 5.3) in which the average performance of a mixture is assessed relative to the average performance of the components across all tested mixtures, and potentially even in a range of environments (Griffing 1956, Norrington-Davies 1967, Schmid et al. 2002, Wuest and Niklaus 2018, Forst et al. 2019). Such designs represent very effective starting points to determine additive contributions of varieties to mixtures, and to evaluate how varietal differences improve their interactions and thus the performance of specific mixture compositions. Such designs, and variations thereof (Forst et al. 2019), therefore shift the focus from comparisons between monocultures and mixtures towards the identification of characteristics that make some variety mixtures more effective than others. Whether these characteristics are best described by differences in genes, in biochemical properties, in physiological or morphological traits, in reaction norms, or in some other metrics is currently unclear. However, identifying such predictors is highly relevant for both mixture development and for ecologists studying biodiversity effects and species coexistence.

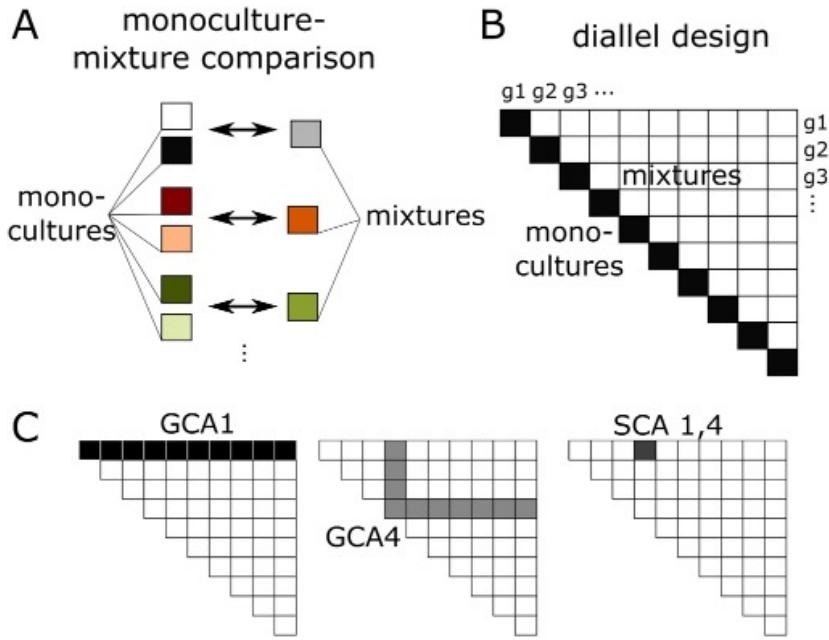


Figure 5.3: Shift from a focus on monoculture-mixture comparisons (A) towards systematic comparisons between different mixtures will require alternative experimental designs. Diallel designs (B) systematically pair different combinations of genotypes or varieties ( $g_1, g_2, g_3, \dots$ ) and focus on comparisons amongst mixtures. (C) Additive performance contributions (biomass, yield, ....) of genotypes across all mixtures (i.e. how much a genotype on average contributes to the mixture performance, termed general combining abilities - GCA; examples shown are GCA1 and 4, which are estimated from the average performance of the shaded compositions) can be modeled from diallel designs without the need to grow monocultures, and allow the formulation of expected mixture performances based on additive contributions only. Expected deviations of genotype combinations from such additive expectations are termed specific combining abilities (SCA) and can be utilized to understand how trait, gene, or other differences between genotypes contribute to complementarity and the added value of specific mixtures.

## 5.10. Conclusions

Increased diversification, ideally at different spatial or temporal scales, represents a known strategic element of a sustainable agricultural intensification. In isolation, variety mixtures are unlikely a “silver bullet” to overcome all major challenges in mechanized agriculture. At the same time, their widespread use could make important contributions to solving some important problems or current trade-offs, and could contribute at multiple levels to improved crop functioning and reduced input. However, transitioning from pure-line breeding to the widespread use of

variety mixtures is associated with different challenges. The ways in which such challenges can be overcome will determine if this transition is deemed economical by the wider breeding and seed-production community, and if high-performing mixtures - with multiple benefits - can be developed that are then embraced by farmers and the processing industry. As we outlined here, ecological principles can guide efforts to address many of these challenges. We've highlighted that ecological theory and the knowledge gained from the study of biodiversity-ecosystem function relationships can be combined with the resources and data obtained by default from breeding or variety testing activities, generating synergies that could tip the perceived economic imbalance between pure cultures and mixtures. Whereas trait-based approaches to predict mixture performances are conceptually attractive and have shown their merits in specific situations, they should not distract from a larger pool of alternative strategies, also rooted in ecological theory, to be explored in the future, namely the utilization of reaction norms to approximate variety niches, mechanism-agnostic genetic approaches and diallel designs which shift the focus from mixture-monoculture comparisons to mixture-mixture comparisons. Time will tell if the more empirical, "engineering" approaches or the more theory-driven methods discussed here will provide better guidance for mixture development, but both will likely enable a better use and understanding of the positive effects of diversity in agro-ecosystems.

## **6. Novel approaches to identify functionally complementary genotypes promoting overyielding in soybean variety mixtures**

Emanuel B. Kopp<sup>1, 2</sup>, Noelle Koenig<sup>2</sup>, Lukas Vonmetz<sup>1</sup>, Samuel E. Wuest<sup>1</sup>, Pascal A. Niklaus<sup>2, 3</sup>

1: Agroscope, Group Breeding Research, Müller-Thurgau-Strasse 29, 8820 Wädenswil, Switzerland

2: Department of Evolutionary Biology and Environmental Studies and Zurich-Basel Plant Science Center, University of Zurich, Winterthurerstrasse 190, 8057 Zurich

3: University of Zurich Research Priority Program on Global Change and Biodiversity, University of Zurich, 8057 Zurich

## Abstract

Positive effects of plant diversity on the functioning of communities have been described to occur frequently. However, the underlying drivers remain poorly understood, and, for applications in agriculture, positive effects of specific mixtures remain essentially unpredictable from the properties of single components. There is consensus that some form of niche complementarity between mixture components drives diversity effects, and that niche dimensions have a correspondence in measurable functional traits. However, current methods typically used to approximate niche complementarity from trait measurements lack the complexity needed to accurately depict this relationship. Multivariate statistical analyses and machine learning methods may become crucial tools to introduce this complexity in analyses of biodiversity-functioning relationships. At the same time, two major challenges are inherent in these methods, namely (1) they require relatively large datasets to work and (2) they require careful model evaluation through cross-validation. Here, we compare a range of statistical learning approaches to predict the diversity effects of two-way soybean genotype mixtures based on the characteristics of the component genotypes when grown in isolation. We combine trait reaction norm differences, expected to better approximate niche complementarity than traditional mean trait differences, with different flexible models allowing for non-linearities between niche complementarity and diversity effects. Diversity effects were measured in more than 600 pairwise soybean mixtures in a systematic pairing design, and single-genotype trait values for all mixture components in completely independent experiments. We then evaluated the models using two cross-validation schemes: (i) one standard scheme that splits mixtures into test and evaluation sets but allows mixtures with shared components to occur in either set, and (ii) a strict scheme that splits mixtures into sets with completely independent components. The results show that (1) strict model validations suggested that overyielding was only poorly predictable even from reaction norms and complex models, but (2) the standard cross-validation grossly overestimated the predictive abilities of several models. The latter result indicates strong data leakage: the models were able to infer information about the mixture components identities from the data instead of training on the trait measurements, strongly influencing the cross-validation results. We conclude that in most biodiversity-ecosystem functioning studies, where models are rarely tested on communities with a completely different composition, this represents a major challenge for explanatory as well as predictive statistical analyses.

## 6.1. Introduction

It is well established that more species-rich plant communities are on average more productive than monocultures, a phenomenon termed overyielding (Hector et al. 1999, Hooper et al. 2005, Cardinale et al. 2007, Tilman et al. 2012, Huang et al. 2018). In agriculture, variety mixtures, i.e. the mixed cultivation of different varieties of the same crop, promise higher yields and could easily be integrated into modern production schemes (Wuest et al. 2021, Huang et al. 2024). However, mixtures differ greatly with respect to overyielding (Reiss and Drinkwater 2018), and it is currently very challenging to predict which mixtures overyield and to which extent (Borg et al. 2018). This limits the development and large-scale introduction of variety mixtures in agricultural practice, and with it the benefits that variety mixtures can bring to production besides overyielding (including more stable yields and higher pathogen suppression, see Tooker and Frank 2012, Wuest et al. 2021).

Ecological theory suggests that the overyielding of mixtures of species or varieties results from niche complementarity among species or varieties (Godoy et al. 2020, Amyntas et al. 2023). However, the niche concept is elusive and the direct quantification of niches difficult. Niches are traditionally described as a set of environmental conditions allowing a species' population to maintain a positive growth rate (Chase and Leibold 2003) and have geometrically been conceptualized as hypervolumes where the dimensions are environmental conditions (Hutchinson 1957). Recent advances in theoretical ecology led to the formulation of more complex niche geometries, including concave shapes and holes within these hypervolumes (Blonder 2018).

Ecologists have often used functional traits as proxies of niche dimensions. Under this premise, distinct functional traits may indicate functional niche complementarity (Wagg et al. 2017, Mahaut et al. 2023). However, the relationship between niche-space and trait-space is not straightforward (e.g. Violle and Jiang 2009), and niches may look even more geometrically complicated in trait-space than in niche-space. To do justice to this fact, it therefore is necessary to consider more complex relationships between niches and trait spaces, including non-linearities or discontinuities. Moreover, both for the approximation of niches and niche complementarity, new measures must be found to be used alongside the established traits. For example, in a recent review article we suggested that to better predict niches, traits might be complemented by trait-responses across environments, which we hereafter refer to as *reaction norms* (Kopp et al. 2023). Following this reasoning, functional complementarity might be modelled not just based on differences among genotypic traits but also by reaction norm differences among these genotypes.

Selecting among analytical models in ecological analyses is typically based on information criteria, which balance the fit of a model with its complexity (e.g. the number of parameters a parametric model uses, see Yates et al. 2023). For more complex models, and in particular in practical applications of machine learning where predictive ability rather than model simplicity is the primary concern, cross-validation techniques are typically used to evaluate model performance. During cross-validation, datasets are split into “training data” that are used to parametrize a model, and “test data” that allow for an independent evaluation. Typically, this approach provides the means to evaluate the predictive ability of a model. One crucial component of cross-validation is to avoid data leakage. Data leakage is the “illicit sharing of information between the training data and the test data” (Bennett et al. 2024). This may, for example, occur when the training and validation data overlap, or when training and validation data somehow are structurally correlated through information which at time of inference, when the model is applied to completely new data, is no longer available. In multi-environment studies, for example, samples often have IDs which they receive based on the environment in which they were grown. ID numbers should have no predictive importance. However, samples from the same environments often have similar ID numbers, and machine learning algorithms will learn that the similar ID numbers originating from particular environments have a higher probability to show certain characteristic trait values.

Here, we use soybean variety mixtures as model to explore a range of approaches that may overcome current limitations in predicting the overyielding of binary mixtures based on traits. Specifically, we investigated whether (1) reaction norm differences among genotypes and (2) the consideration of non-linearities between predictors and mixture overyielding are superior to just using linear models and “simple” traits.

## 6.2. Materials and methods

### Experiments

Ninety soybean (*Glycine max*) varieties were selected based on their geographic origin and domestication history (ranging from landraces to elite cultivars – see supplementary materials 4.5 for details). For all experiments, seeds were inoculated with rhizobia (HiStick, BASF, containing *Bradyrhizobium japonicum*) and germinated in propagator trays in the greenhouse.

Variety mixture data was obtained from a pot experiment (Figure 6.1A, hereafter called mixture experiment) conducted during the summer 2022 in Wädenswil, Switzerland (47.222°N, 8.669°E, 509 m a.s.l.). After 14 days in the propagator trays, the seedlings were transplanted

pairwise in 5 l pots, inoculated with 10 ml of a solution of 10 g/l of mycorrhiza (*Rhizophagus irregularis*, Lalrise Max WP, Lallemand Inc, Canada) and grown in two blocks under hail protection nets until maturity, the plants were irrigated daily. From the 90 used varieties, 8 varieties differing in geographic origin and domestication history were selected as so-called “testers”. All 90 varieties were paired with these 8 testers for a total of 712 distinct two-way mixtures and 8 tester monocultures. Additionally, all varieties were grown twice in monocultures. Each block was completed by 41 distinct two-way mixtures from the 82 non-tester varieties to increase the total number of mixtures. Hence, the final experiment consisted of 2 blocks with 712 mixtures containing tester varieties, 41 mixtures not containing tester varieties, 24 tester monocultures and 164 non-tester monocultures each, for a total of 1882 pots or 3764 plants (Figure 6.1B). Each variety was individually harvested after reaching maturity to avoid losses due to significant differences in maturity date. After harvest, all plants were dried for at least 24h at 60°C and shoot biomass was weighed.

Genotypic functional traits were obtained from two distinct experiments conducted in 2022 and in 2023 (hereafter called trait experiments). In the first one, the 90 varieties were grown as single plants in 5 l pots in four blocks in the greenhouse. We tried to generate blocks that were as environmentally different as possible, so that the measured traits could vary along many environmental variables. For each block, the pots were filled with a different substrate, thus exposing the varieties to a different soil environment. We used lawn soil, sand-mix, soybean breeding substrate with added mycorrhiza and potting soil. The four substrates strongly differed in organic matter content, pH and texture (see Table SM 6.1 for details). There were no within block replicates, so the experiment consisted of a total of 360 plants. Plant height, relative growth rate, leaf area, number of leaves, stomatal conductance, flowering date, chlorophyll content, number of branches, leaf weight, leaf dry matter content, specific leaf area, aboveground biomass and seed yield were recorded per plant.

In the second trait experiment, the same 90 varieties were grown outdoors at two different locations (Wädenswil, 47.222°N, 8.669°E, 509 m a.s.l. and Zürich-Irchel, 47.396°N, 8.551°E, 506 m a.s.l; distance ca. 21 km) and in two different soil-types each (for a total of four different environments). At Irchel, the plants were grown in big containers (approx. 1 x 1 m) filled once with organic and nutrient rich soil and once with nutrient poor sandy soil. In each container, 6 varieties were grown at 20 cm from each other. Per variety, four plants were planted at 4 cm from each other to obtain more stable genotypic trait values. In Wädenswil, the plants were grown on two fields, both of which have received standard fertilizer treatments every year, but one of which has additionally been treated with compost over the last 11 years. Again, four plants per genotype

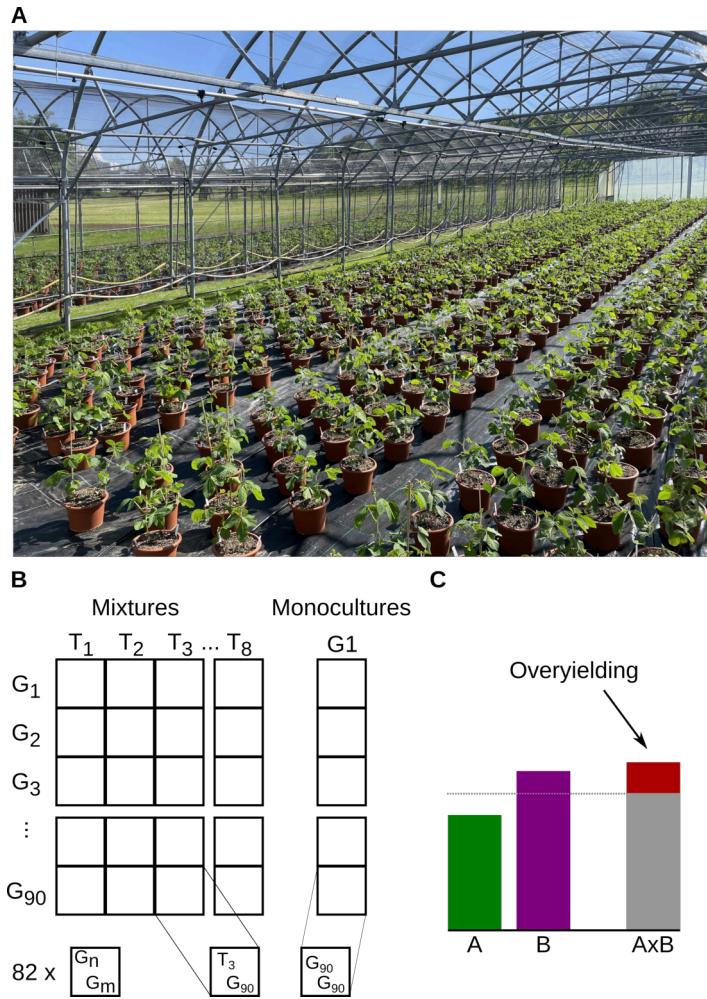


Figure 6.1: (A) Photo of the mixture experiment. (B) Experimental design of overyielding experiment. In two identical blocks we grew two-way mixtures consisting of one of 90 genotypes and one of 8 tester genotypes (the tester genotypes were also part of the overall 90 genotypes). Moreover, we grew all 90 genotypes as monocultures in both blocks. Additionally, we grew 82 mixtures consisting of unique combinations of the 82 non-tester genotypes (41 additional pots per block). (C) Overyielding (red) is calculated as the deviation from the mean between the two mixture components.

were grown close to each other, at 50 cm from the next group of four plants. At both locations, the varieties were replicated between one and three times (depending on seed availability) in order to use all the available space, but each trait measurement was averaged per genotype and environment. The experiment was randomized in a way to ensure that replicated genotypes never had the same neighbors. In this experiment, the height of the plants was measured twice and averaged, and the aboveground biomass and seed yield was weighed after harvesting and drying the plants.

## Data processing

Data from the outdoor pot experiment was corrected for block effects and for spatial effects within blocks using linear models. Then, overyielding per genotype mixture was calculated as difference between the observed mixture performance and the average monoculture performance of the two mixture components (Figure 6.1 C).

Data for both trait experiments were processed separately but in the same way. For each measured trait and genotype, the mean value across all four environments was calculated. Then, for each two-way genotype combination, mean trait differences were calculated. Additionally, per genotype, trait reaction norms were calculated as the difference of the trait values between all pairwise combinations of environments. Then, similar to how trait differences among genotypes are used as potential predictors of overyielding, we calculated reaction norm differences for all genotype pairs (Figure 6.2).

A problem inherent in the analysis of trait data is that traits are not independent because they are determined by fundamental ecological strategies and trade-offs. To reduce the linear dependency of the many trait values, we reduced the dimensionality of the trait data set by principal component analysis (PCA), which we ran for both mean trait differences and reaction norm differences after scaling and centering the data using Z-score normalization.

## Cross-validation and modelling

Two distinct cross-validation (CV) strategies were compared (Figure 6.3). For the first, which we refer to as “standard”, the test set contained all 82 mixtures without tester varieties and a random selection of 2 mixtures per tester varieties (added to reach a ratio of 85/15 between training and test set). This allowed us to avoid any model to potentially learn patterns based on the identity of one of the eight tester varieties only, as such a model would poorly generalize on data not containing any tester varieties. Indeed, if one genotype was present in many more mixtures than other genotypes both in the training and the test set, this might lead to the model assessing the presence/absence of the said genotype instead of giving reliable estimates of genotype interaction effects (Roberts et al. 2017). Moreover, the random addition of two mixtures per tester genotype to the test dataset allowed us to run the models multiple times with slightly different training and test datasets, which is useful to assess the stability of the predictive model. This CV approach is the closest to the current standard in BEF studies where multiple different mixtures assembled from the same components are typically compared without explicitly considering their referencing to the same component monocultures (where, however, usually no explicit CV is done). The presence of the same varieties (in different mixtures) in training and test data

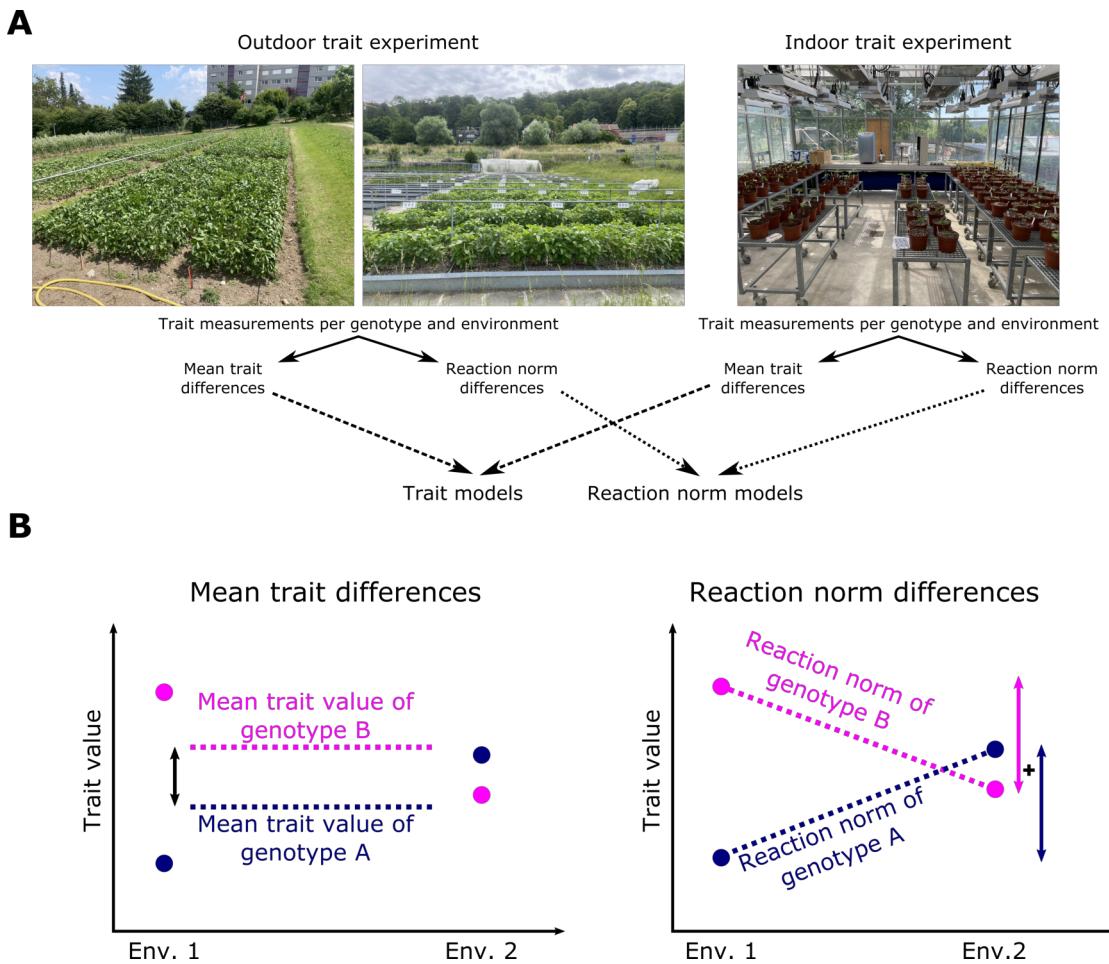


Figure 6.2: (A) Pictures of the two independent trait experiments and how the measured trait measurements were processed. (B) Calculation of mean trait differences and reaction norm differences.

could potentially lead to data leakage through correlated error structures. This is because all mixtures are considered as statistically independent, but they refer to the same monocultures (both in terms of predictors and in terms of response). Therefore, although the measurement error is random by definition, it could lead to structural correlations in the mixtures (both on predictor and response side) as it propagates in the same way for all mixtures sharing one component.

For the second CV strategy, here referred to as “strict”, we first randomly selected one of the eight tester varieties and 19 of the 82 non-tester varieties. The test set then consisted of all the mixtures assembled from these 20 varieties, while the training set consisted of all the mixtures composed from the remaining 70 varieties. A genotype could therefore be present only either in the training or in the test data, preventing the models to use any genotype-specific information.

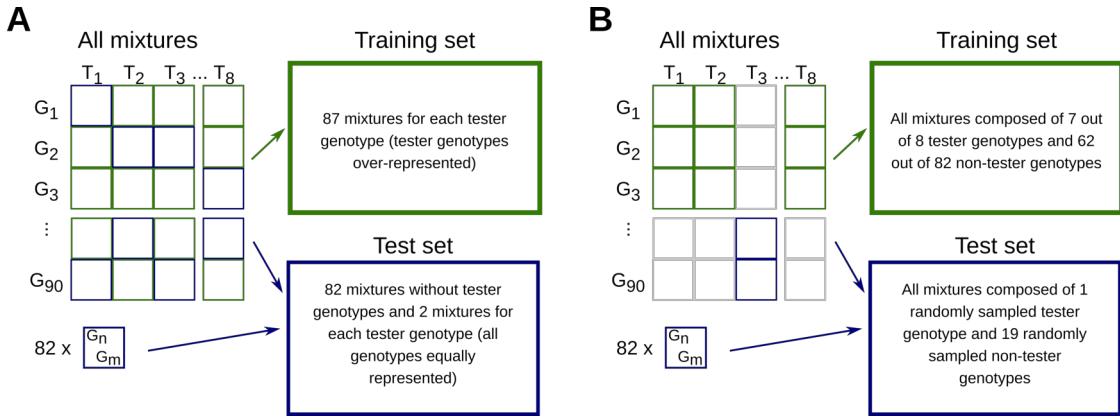


Figure 6.3: Comparison of the two cross-validation (CV) strategies. (A) “Standard” strategy. The test set is balanced (no genotypes are over-represented), but the same genotypes can be present both in training and in test data (but never in the same mixtures). (B) “Strict” strategy. Training and test data are composed from mixtures without any overlapping genotypes.

We trained three different model types: linear regressors, random forests regressors and  $k$ -nearest-neighbor (kNN) regressors (Fixt and Hodges 1989, Ho 1995). In the linear models, all mean trait differences, or all reaction norm differences respectively were included as predictors. Random forests build multiple decision trees based on a bootstrapped subset of the provided data and a subset of the predictors (Ho 1995). At each step in each decision tree the variable explaining most of the remaining variance is selected and the data split based on this variable. This process is repeated until the maximum depth of the trees is reached. Eventually, the individual trees are averaged into a final ensemble model. This trained random forest regressor can then be applied to the test data to predict the dependent variable. A kNN regressor simply places the training data in a Euclidean space defined by the provided predictors (Fixt and Hodges 1989). It then predicts the values of the dependent variable of the test data by placing the test data in the same predictor space and averaging the values of the  $k$  nearest training data points. For each model type, we used mean trait differences or reaction norm differences as predictors, and overyielding as response. Moreover, for each model we once used the untransformed predictors and once used principal components as predictors (keeping all PC needed to preserve at least 99% of the data variance, up to a maximum of 15 PCs). Additionally, we run the kNN models with all possible three-way combinations of PCs to identify the best set of PCs. All kNN models were run with  $k = 3$ , meaning that the OY value of each point in the test dataset was predicted by averaging

the OY values of the three closest point of the training dataset in the feature space. Euclidean distance was used to compute neighbor proximity.

Each model was run 100 times with different splits of the data into training and validation subsets. For each run, the models were fit on the training dataset. Then, the expected OY of the test dataset was predicted from the model and the correlation ( $R^2$ ) with the observed OY of the test data was calculated to quantify its predictive ability.

### Statistical analysis

All analysis was performed in Julia (v1.9.1) or in R (v4.3.1). Linear models were implemented through GLM.jl (Bates et al. 2023), random forest models were implemented through DecisionTrees.jl (Sadeghi et al. 2022) within the MLJ.jl framework (Blaom et al. 2024). kNN models were implemented through NearestNeighbors.jl (Carlsson et al. 2024) within the MLJ.jl framework. PCA was done with MultivariateStats.jl.

## 6.3. Results

### Overyielding

The performance of 30 mixtures could not be assessed due to dead or not well-looking plants. Over the 764 mixtures analyzed, mean biomass overyielding was 4.4 g (3.5 %). However, overyielding was highly variable, ranging from -51.4 to 94.2 g (-35% to 72%, Figure 6.4). While negative overyielding is rare in natural communities (Cardinale et al. 2007), our findings are in line with previous studies done in crops (see Reiss and Drinkwater 2018 for a meta-analysis). Moreover, our OY estimates carry the compound biomass measurement error of the mixture biomass and the biomass of the two reference monocultures, and some of the observed variation in OY thus is due to random measurement error.

### Variability of prediction

The variability of model prediction accuracy was higher when using strict CV than when using standard CV (Figure 6.5). This is one hand because in standard CV, all 82 non-tester mixtures are always included in the test dataset, leading to higher similarity between the test sets of different model runs. In strict CV, on the other hand, the test datasets were much more variable among model runs. Moreover, in strict CV the test dataset is smaller than in standard CV, further increasing the variability of the prediction. Contrary to our hypothesis, models using reaction norm differences as predictors did not systematically outperform models using mean trait differences as predictors. However, limiting ourselves to standard CV might have led us

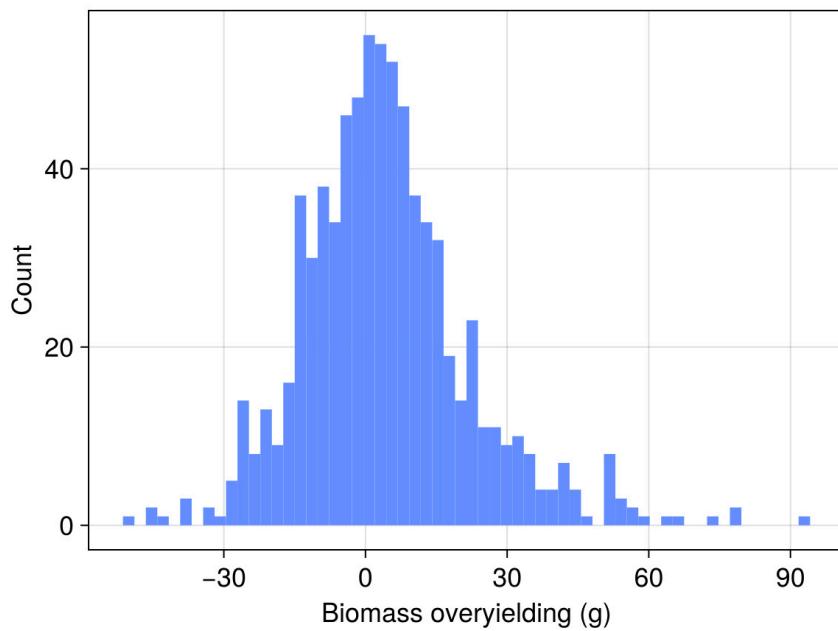


Figure 6.4: Biomass overyielding distribution of 764 measured mixtures. Mean OY was 4.4 g, but values were highly variable.

to make this conclusion: when using reaction norm differences as predictors, the random forest with standard CV outperformed all other models. When using strict CV, no model performed significantly better or worse than the others.

### Data leakage

The two CV strategies differ only in the fact that in standard CV, mixtures in the training and the test set share components, while in the strict CV, the mixtures in the two sets are assembled from different varieties. This implies that the difference in the predictive ability of the models appearing between the two CV strategies is due to data leakage related to genotype identities. Specifically, this difference suggests that in standard CV, the models can infer information about mixture components identities from the trait data during training. In the test phase, the model infers the same information from the test data and predicts OY based on mean genotypic OY values and not on functional entities (Figure 6.6). This behavior can be tested by randomly reshuffling genotype IDs in the trait data prior to computing mean trait differences and reaction norm differences. If the models worked with functional values, the predictive ability should now tend towards zero (as randomly computed trait differences should be uninformative with respect to overyielding). On the other hand, if there is data leakage based on genotype identity, the

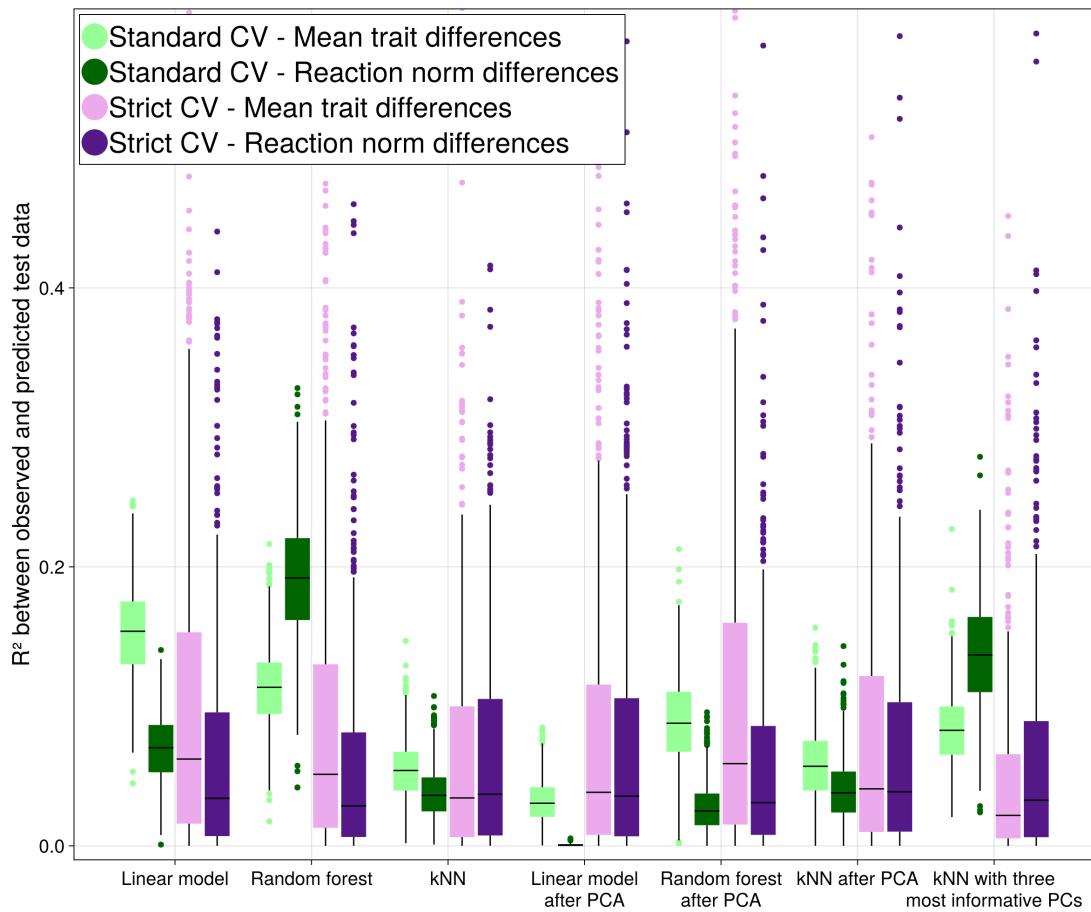


Figure 6.5: Model prediction accuracy ( $R^2$  between prediction and observation of test set) of the different models depending on the predictors used (light vs. dark) and the cross-validation strategy applied (green vs. purple).

models should show some predictive ability even when randomly shuffling genotype IDs. Shuffling the genotypes IDs 20 times and running all models with standard CV confirmed our suggestion: models still showed consistent predictive ability, especially when using reaction norm differences as predictors (Figure 6.7). The effect was particularly pronounced for the more flexible models, like random forest and dimensionality reduced kNN.

## 6.4. Discussion

### Different mixtures referring to the same monocultures lead to data leakage

Here, we aimed to leverage ecological niche theory to develop novel models to predict genotype mixture overyielding from traits determined in monoculture stands or on single plants,

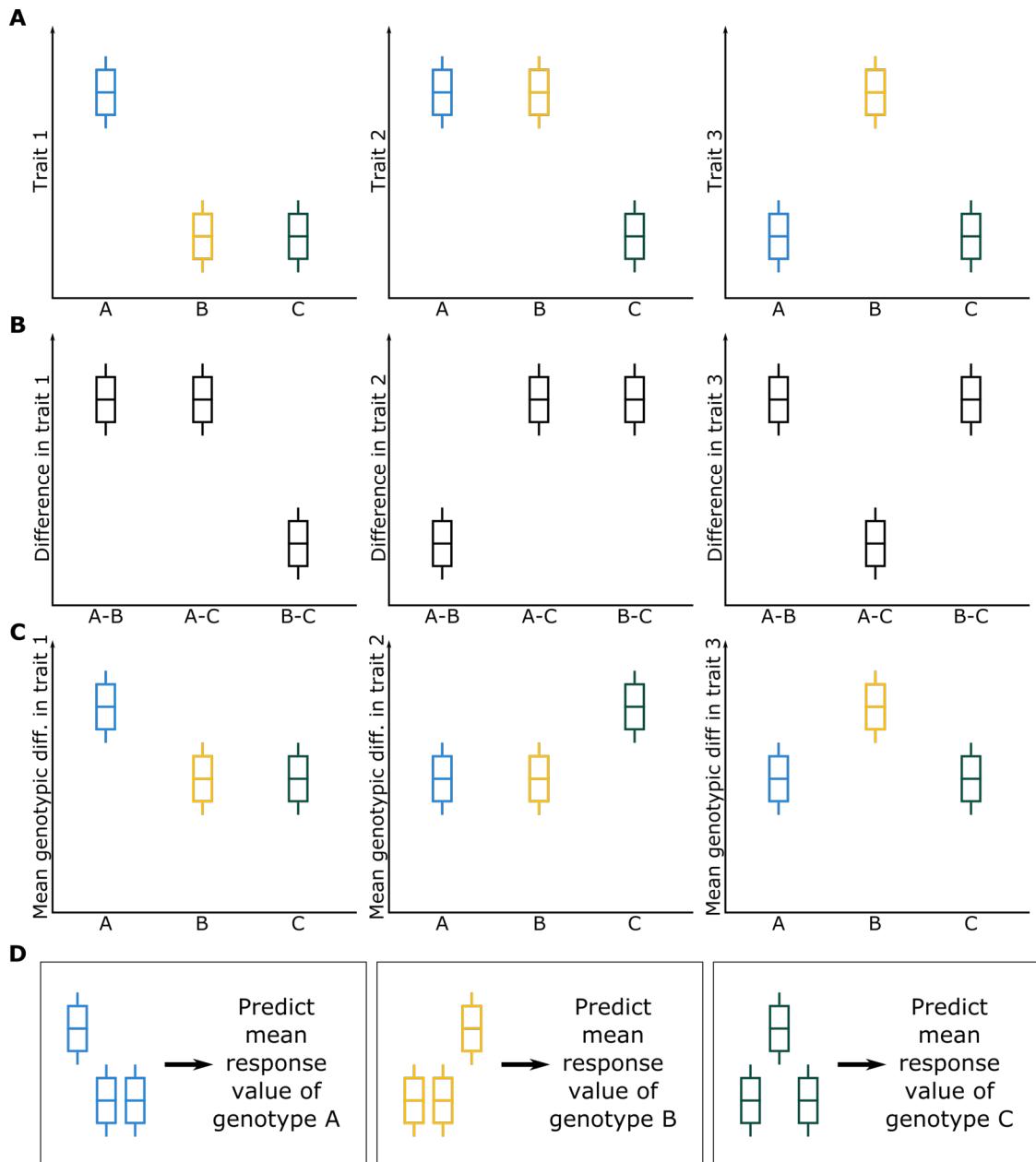


Figure 6.6: Example of how genotype identities could inform prediction. (A) Three traits are measured for the three genotypes A, B and C. (B) Pairwise trait differences for each genotype pair and each trait are computed. (C) Mean trait differences values aggregated by component genotype differ between genotypes. (D) This difference can be used to predict the response based on the mean response value of all mixtures containing a certain genotype.

and tested these models using soybean as model system. We first replaced traditionally-used mean trait differences by reaction norms differences as model predictors. Second, we replaced the traditional linear models by models capable of non-linearly mapping trait differences to

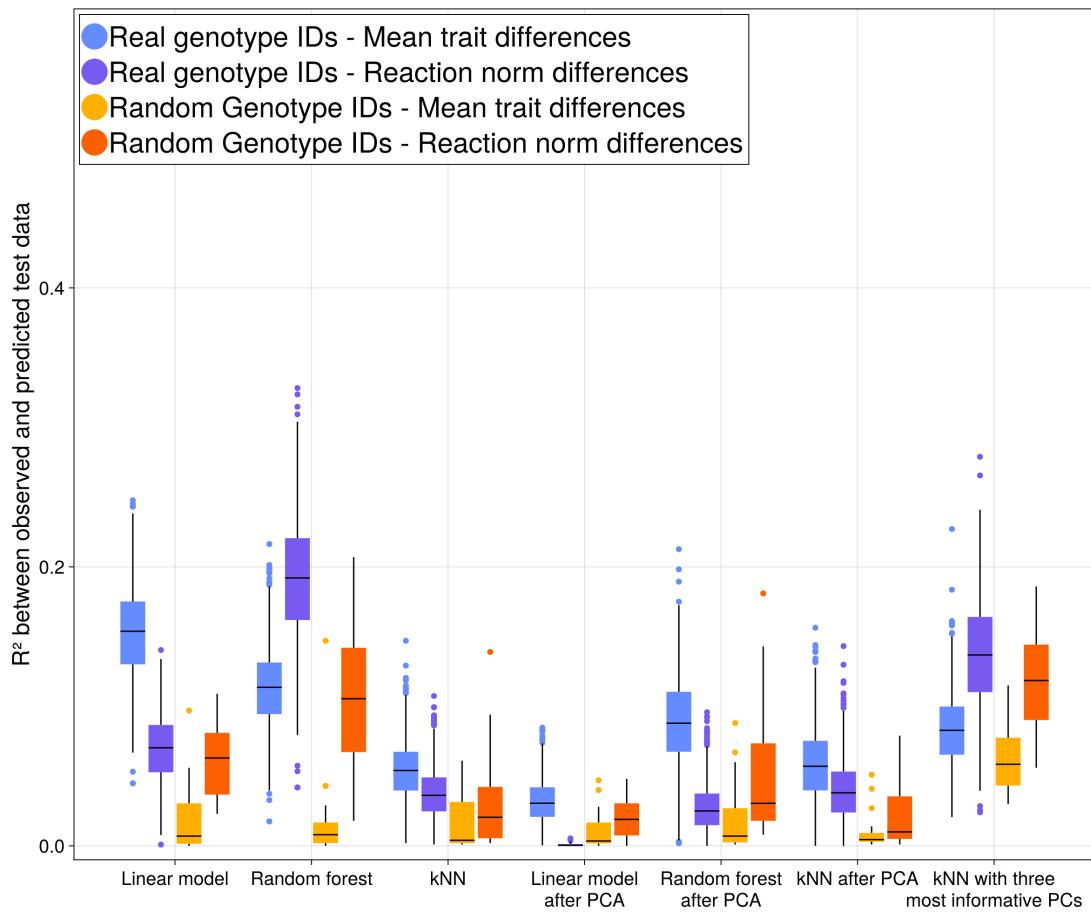


Figure 6.7: Comparison of predictive ability of models where the trait data are correctly assigned to the genotypes before computing mean trait differences and reaction norm differences (blue and purple) against predictive ability of models where the genotypes IDs are randomly shuffled before before computing mean trait differences and reaction norm differences (orange and red).

overyielding. Finally, we reduced the dimensionality of the predictor data with a PCA to account for collinearity among predictors, and selected only three of the remaining principal components as predictors for our final model.

We tested our model using two distinct cross-validation strategies. The first considers mixtures that share one component as independent, even though both their trait differences (mean trait differences or reaction norm differences) and their overyielding measurements are inherently correlated through their common referencing to the same monoculture biomass or trait measurements (a special case of pseudoreplication). Ecological studies often aim at using some measured variables to explain the variation within a set of observations, rather than at predicting how observations would look like based on values of the measured variables not directly present in the

dataset (see Shmueli 2010 for an in depth discussion about the differences between explanatory and predictive modelling). Indeed, many studies investigating biodiversity-ecosystem functioning (BEF) relationships use the entire available data to fit their statistical models, abstaining from rigorous cross-validation. This makes it difficult to directly compare our results to previous ones. However, the assumption of independence of overyielding measurements is ubiquitous, and many studies focus on mixtures which share at least one component. One previous study in durum wheat, for example, used 19 trait means and the corresponding trait differences to explain overyielding in 220 pairwise mixtures composed from 179 different genotypes (Montazeaud et al. 2020). After model selection, it reported to have explained 17% of the variation in biomass overyielding (representing the model fit).

The other CV strategy consisted in testing predictions only on mixtures composed of genotypes different from those used in model training. This ensured that all the information available to the model was at mixture level (in the mean trait differences or reaction norm differences) but not at genotypes level (through correlated error structures between mixtures sharing components). Even in relatively simple models (linear models using all mean trait differences or all reaction norm differences as predictors, respectively), this can lead to quite different interpretations of how well overyielding in new mixtures is predictable (Figure 6.8). Whereas the coefficient of determination of the linear model fits or even of the standard CV runs are in a range that agrees well with model fits reported previously (e.g. in Montazeaud et al. 2020), the ones calculated from strict CV are substantially lower.

The difference in performance between the two CV methods can be explained by a special case of data leakage, appearing when there are “dependencies between training and test data that do not exist between training and inference-time data” (Bennett et al. 2024). Indeed, in the standard CV method, the training and test data share information about mixture performance of the same genotypes (in different mixtures). In the strict CV, on the other hand, the training and test data do not share any information at genotype level: all information is about genotypic differences of the mixture components. Therefore, the standard CV is adequate only to predict the performance of mixtures composed of the same genotypes as those used in the training set. This limits its broader applicability as well as its suitability to inform fundamental ecological questions such as the relationship between functional trait differences and mixture productivity. The strong discrepancy of the results obtained with the two CV strategies highlights the importance of considering the non-independence of measurements when dealing with overyielding. This is an ubiquitous issue for the current BEF studies, and fundamentally questions what we know about the drivers of overyielding both in agricultural and in ecological systems.

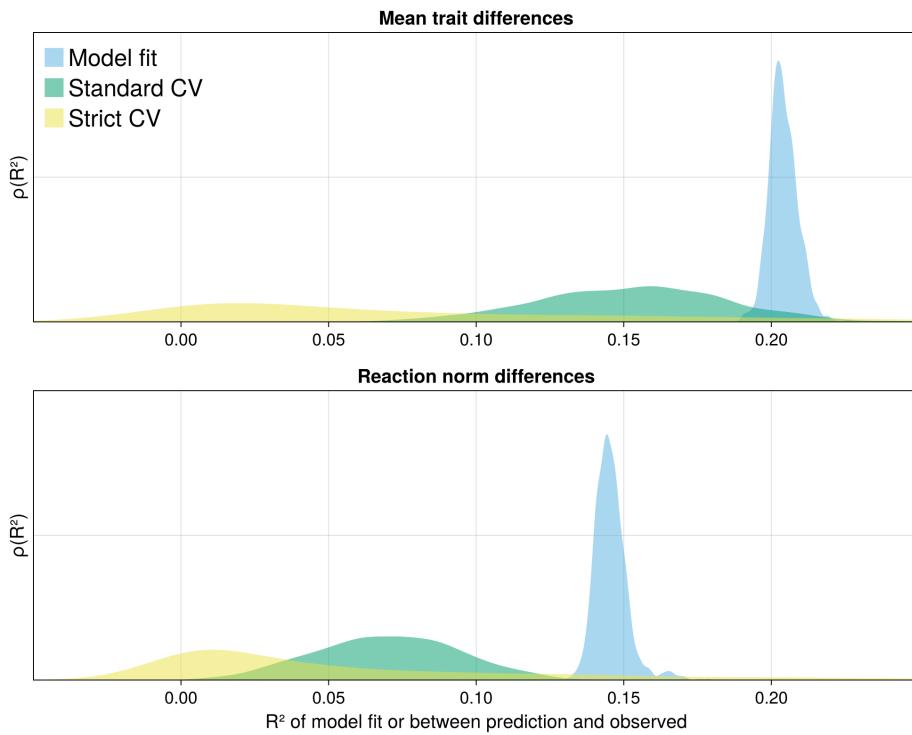


Figure 6.8: Results obtained by linear models using all mean trait differences or all reaction norm differences as predictors. Density distribution of coefficient of determination ( $R^2$ ) of model fit (on training data, in blue), standard CV (on test data with mixtures sharing some components with the mixtures from training data, in green) and strict CV (on test data with mixtures which have no shared components with the mixtures in training data, in yellow).

### Measurement errors propagate into composite measurements like overyielding

In addition to the data leakage problem, there is a problem with the statistical propagation of measurement errors. Indeed, as overyielding is calculated using mixture performance and mixture-components monoculture performance, both the measurement errors of the mixture performance and of the monoculture performances will propagate into the overyielding estimate. We explored our variation in overyielding to estimate how much of it was real and how much was a consequence of measurement error propagation in the calculation of overyielding (see Appendix 6.6). We found that, when considering all overyielding measurements as independent, around 46% of variation in overyielding was potentially explainable, with the rest being “random noise”. However, when considering that all 764 mixtures are referring to the same 90 monoculture measurements, this value dropped to ~42%. While this difference might appear marginal, it shows how measurement errors can propagate into a derived metric such as overyielding. Again,

the propagation of measurement error into the overyielding value is ubiquitous and usually not accounted for. Indeed, it is present by design as soon as multiple mixtures reference to the same monocultures. Experimentally, it could be avoided only by growing separate monoculture references for each mixture. Additionally, it is plausible to assume that the error propagation will have more severe effects in experiments with less power.

### **Conclusion and outlook**

Relying on the strict cross-validation results, we could not confirm the postulated hypothesis about reaction norm differences being better predictors of diversity effects than mean trait differences. However, this work shows how considering non-independent (community level) trait and overyielding values as independent ones can lead to erroneous results and incorrect conclusions due to data leakage and error propagation. The impact of this in past BEF studies is hard to quantify and depends on the specific study designs. Nevertheless, it raises questions if the commonly adopted methods to analyze BEF experiments, especially experiments using many explanatory variables (e.g. traits and trait differences), lead to the over-interpretation of findings due to the confounding of community level with community-components level effects.

## 6.5. Supplementary materials

Table 6.1: Content details of the four soil types used in the greenhouse experiment.

Soil type	pH	KCl (g/l)	N (mg/l)	$P_2O_5$ (mg/l)	$K_2O$ (mg/l)	Mg (mg/l)	S (mg/l)	Composition
Lawn soil	7.0	1.6	120	80	600	220	120	Peat free, high percentage of sand and perlite
Sand-mix (ED73 + sand)	5.8	2.0	340	260	330	100	-	Culture substrate: organic soil, peat, raw clay, fertilizer (calcium carbonate, NPK)
Soy-mix substrate	3-4	-	-	-	-	-	-	See supplementary materials 4.5
Potting soil	5.8-6.7	2.0	200-500	200-2000	400-3000	-	-	Peat, green compost, wood fiber, mineral N-fertilizer, organic NPK-fertilizer

## 6.6. Appendix: Error propagation

### Error propagation in overyielding

#### Error propagation formulas:

Mean:  $\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$

Sum:  $\sigma_f = \sqrt{\sigma_x^2 + \sigma_y^2}$

Difference:  $\sigma_f = \sqrt{\sigma_x^2 + \sigma_y^2}$

#### Overyielding, general formula:

$$\begin{aligned} OY &= Y_{TG} - \frac{1}{2}(Y_{TT} + Y_{GG}) \\ &= Y_{TG} - \frac{1}{2}\left(\frac{Y_{T_1} + Y_{T_2} + Y_{T_3}}{3} + \frac{Y_{G_1} + Y_{G_2}}{2}\right) \end{aligned}$$

#### Overyielding, error propagation:

TG: Mixture, one pot per block

T: Tester, monoculture replicated 3 times per block

G: Genotype, monoculture replicated twice per block

Part left of minus sign:

$$\epsilon_{T_{TG}} = \sigma$$

Part right of minus sign:

$$\epsilon_{\frac{Y_{T_1} + Y_{T_2} + Y_{T_3}}{3}} = \frac{\sigma}{\sqrt{3}}$$

$$\epsilon_{\frac{Y_{G_1}+Y_{G_2}}{2}} = \frac{\sigma}{\sqrt{2}}$$

$$0.5 * (\epsilon_{\frac{Y_{T_1}+Y_{T_2}+Y_{T_3}}{3}} + \epsilon_{\frac{Y_{G_1}+Y_{G_2}}{2}}) = 0.5 * \sqrt{(\frac{\sigma}{\sqrt{3}})^2 + (\frac{\sigma}{\sqrt{2}})^2} = 0.5 * \sqrt{\frac{\sigma^2}{3} + \frac{\sigma^2}{2}}$$

Togther using rules for difference:

$$\epsilon_{OY_{block}} = \sqrt{\sigma^2 + (0.5 * \sqrt{\frac{\sigma^2}{3} + \frac{\sigma^2}{2}})^2} = \sqrt{\sigma^2 + 0.25 * (\frac{\sigma^2}{3} + \frac{\sigma^2}{2})} = \sigma * \sqrt{1 + 0.25 * (\frac{1}{3} + \frac{1}{2})} = \sigma * \sqrt{\frac{29}{24}}$$

Final formula for two blocks:

$$\epsilon_{OY} = \frac{\sigma * \sqrt{\frac{29}{24}}}{\sqrt{2}} \approx \sigma * 0.777$$

### After correction for block, row and column and removing outliers:

```
using DataFrames, CSV, GLM, RCall, Statistics
using Idefix: overyielding
dat = CSV.read("Data/IntermediateData/2022_potdata_biomass_ready.csv",
DataFrame, missingstring="NA")
dat = dropmissing(dat, :biomass_sum_cor)
@rput dat
R"library(pascal); dat$GenoComp <- sorted.code(dat$Comp1, dat$Comp2, collapse = '-')"
@rget dat
sigma = lm(@formula(biomass_sum_cor ~ block * GenoComp), dat).model |> dispersion
```

17.017646342712272

```
OY_error = sigma * 0.777
```

13.222711208287436

### Maximal possible $R^2$

Formula:

$$R_{max}^2 = 1 - \left( \frac{\sigma_{noise}^2}{var(x)} \right)$$

```
dat_OY = overyielding(dat, "Comp1", "Comp2", ["biomass_sum_cor"];  
    percent = false, OY_composition_mean = true) # OY  
dat_OY = dat_OY[dat_OY.Comp1 != dat_OY.Comp2, :] # remove monocultures  
varOY = var(dat_OY.biomass_sum_cor_OY)
```

324.0078036151318

```
R_OY_max = 1 - (OY_error^2 / varOY)
```

0.4603830841511103

---

## Check with simulation

Create a dataset that has the same structure as the one used in the experiments.

```
R""""  
# Simulation of error propagation  
  
# experimental design  
testers <- sprintf("T-%02d", 1:8)  
nontesters <- sprintf("G-%02d", 9:90)  
allgenos <- c(testers, nontesters)  
  
monos <- data.frame(  
    genotype = c(  
        rep(testers, each = 6), # 48  
        rep(nontesters, each = 4) # 328  
    )  
)  
  
tester_mixtures <- NULL  
for (tst in testers) {  
    tester_mixtures <- rbind(  
        tester_mixtures,  
        data.frame(geno1 = tst, geno2 = setdiff(allgenos, tst))  
    )  
}
```

```

extra_mixtures <- data.frame(
  geno1 = nontesters,
  geno2 = nontesters[c(2:82, 1)]
)

## Auswahl des Simulationssets

mixtures <- rbind(tester_mixtures, tester_mixtures, extra_mixtures)

mixtures$y <- 0
mix_means <- aggregate(y ~ geno1 + geno2, FUN = mean, data = mixtures)

monos$y <- 0
mono_means <- aggregate(y ~ genotype, FUN = mean, data = monos)

gvec <- sort(unique(monos$genotype))
idx1 <- match(mono_means$geno1, gvec)
idx2 <- match(mix_means$geno2, gvec)
"""

```

Populate the dataset 10'000 times with random numbers and calculate OY.

```

R"""
## simulation
indi_OY <- replicate(1e4, {
  t1 <- sample(testers, 1)
  g1 <- sample(nontesters, 1)
  mo1 <- subset(monos, genotype %in% c(t1, g1))
  mi1 <- subset(mixtures, geno1 %in% c(t1, g1) & geno2 %in% c(t1, g1))
  mo1$y <- rnorm(nrow(mo1))
  mo1_means <- aggregate(y ~ genotype, FUN = mean, data = mo1)
  mi1$y <- rnorm(nrow(mi1))
  mi1_means <- aggregate(y ~ geno1 + geno2, FUN = mean, data = mi1)
  mi1_means$oy <- mi1_means$y - 0.5 * (sum(mo1_means$y))
})
sd_OY_ind <- sd(indi_OY)
round(mean(sd_OY_ind), 3)
"""

RObject{RealSxp}
[1] 0.774

```

The conversion factor of 0.777 found algebraically is confirmed.

### Considering that mixtures with shared components are not independent

```
R"""
sd_OY <- replicate(1e4, {
  monos$y <- rnorm(nrow(monos))
  mono_means <- aggregate(y ~ genotype, FUN = mean, data = monos)
  mixtures$y <- rnorm(nrow(mixtures))
  mix_means <- aggregate(y ~ geno1 + geno2, FUN = mean, data = mixtures)
  all <- merge(mix_means, mono_means, by.x = "geno1", by.y = "genotype")
  all <- merge(all, mono_means, by.x = "geno2", by.y = "genotype")
  names(all) <- c("geno2", "geno1", "y", "mono_gen01", "mono_gen02")
  all$oy <- all$y - 0.5 * (all$mono_gen01 + all$mono_gen02)

  sd(all$oy)
})
round(mean(sd_OY), 3)
"""

RObject{RealSxp}
[1] 0.806
```

Calculating maximal possible  $R^2$  considering OY as not independent:

```
R_OY_max = 1 - ((sigma * 0.806)^2 / varOY)
```

```
0.4193511082879746
```

## 7. General discussion

Agriculture in the last century has reached a level of productivity never seen before in human history. However, this was obtained at a high cost for the environment due to increased water eutrophication, air pollution, pesticide contamination of non-target species or other similar side-effects of modern high-intensity agricultural systems. Thus, it appears clear that agriculture is in urgent need for innovation, with the overarching objective to increase sustainability while maintaining high productivity.

Evolutionary agroecology rethinks agricultural systems from an evolutionary and ecological perspective. This allows to overcome the outdated mental model of agriculture as a primarily linear “input → output” system. Rather, the eco-evolutionary perspective offers a model of agriculture based mainly on trade-offs, where both the sustainability and the environmental effects of the system are no longer mere externalities, but intrinsic part of the system itself. Soybean is an ideal model crop to study eco-evolutionary approaches. Indeed, its importance will further increase in the next decades as its symbiosis with nitrogen fixing rhizobia make it attractive for sustainable production systems, while its breeding progress still lags behind the major cereal crops (Peoples et al. 2009, Zimmer et al. 2016).

In this dissertation, eco-evolutionary concepts have been applied to the model crop soybean to investigate their potential. To do so, experimental work, literature review and state-of-the-art statistical and computational approaches were combined. Experiments were conducted at different organizational levels, investigating single plants and communities. This allowed to balance the focus on underlying processes (i.e. allocation strategies in plant organs) and the broader view on the general drivers (i.e. competition and complementarity). Overall, this work gives novel insights into soybean competition, cooperation and complementarity which might have profound implications for agricultural practice (i.e. variety mixture development) and soybean breeding (i.e. breeding for cooperativeness). At the same time, it significantly contributes to the advancement of ecology by resolving long-debated methodological issues (Chapter 3), outlining a specific example of competition-costs associated trade-off (Chapter 4) or improving our understanding of the methodological challenges occurring when investigating the relationship between functional trait differences and community productivity (Chapter 5 and Chapter 6).

## 7.1. From individual plants to genotypes

While Chapter 3 focused on within-genotype effects of individual plant allocation patterns on collective performance, Chapter 4 moved the focus to differences in competitiveness between genotypes. At the within-genotype level, plants exhibited significant variation in allocation patterns depending on competition cues. For instance, plants exposed to both aboveground and belowground competition prioritized root growth over aboveground biomass, suggesting a strategy to outcompete neighbors for resources. At the between-genotype level, significant variation in competitiveness across a representative soybean population was found. Interestingly however, the experiments performed in Chapter 3 revealed little genotypic differences in the measured trait responses, indicating that the drivers underlying the patterns observed in the two experiments were presumably not the same. This highlights the complexity of the interplay between phenotypic plasticity upon competition cues and the genetic architecture of competition in soybean. The transition from individual-level traits to genotype-level performance is still to be examined in further depth.

### 7.1.1. Tragedy of the commons, allocation strategies and phenotypic plasticity

A central theme emerging from Chapter 3 is the “tragedy of the commons,” wherein individual optimization can lead to collective inefficiency (Hardin 1968). Plants can increase allocation to competitive structures such as the root system; but in communities where all individuals pursue similar strategies, shared resources are quickly depleted, resulting in reduced overall productivity (Gersani et al. 2001, Zea-Cabrera et al. 2006, Zhu et al. 2019b, Cabal et al. 2020). It is important to notice how a certain degree of phenotypic plasticity is a prerequisite for this phenomenon, as individuals need to adapt their allocation patterns upon the presence of a competitor or cues of changes in resource availability (Cabal 2022). To better understand how individual-level competition traits and genotype-level competitiveness are linked it would therefore be of interest to further investigate the genotypic variation in phenotypic plasticity and how phenotypic plasticity affects or is affected by plant-plant interactions. Eventually, this would allow to investigate how variation in phenotypic plasticity upon competition cues influences plant group performance.

The impact of plant-plant interactions on phenotypic plasticity in response to abiotic conditions remains largely unknown (Wang and Callaway 2021). However, it is widely recognized that phenotypic plasticity is, at least partially, regulated by genetic factors (Zhai et al. 2014, Kadam et al. 2017, Mangin et al. 2017, Cooper et al. 2019). This suggests that plant plasticity

can evolve as an adaptive trait (Ghalambor et al. 2007, Brooker et al. 2022), especially important for species living in rapidly changing environments. While plasticity therefore represents an important resource for plants to face climate change (Lande 2009, Franks et al. 2014, Valladares et al. 2014, De La Mata and Zas 2023), it also represents an under-exploited breeding target to develop crop varieties more resilient to future environmental conditions, predicted to show higher climatic variability (Sambatti and Caylor 2007, Kadam et al. 2017). To investigate the genetic basis of phenotypic plasticity in crops and its breeding potential, recently published genome data of geo-references crop landraces and crop wild relatives are a precious resource (Maccaferri et al. 2019, Sansaloni et al. 2020, Cheng et al. 2024).

## 7.2. Genotypic interactions: cooperation and competition

The dynamics of cooperation and competition in plants provide a unique perspective within the broader framework of ecological and evolutionary interactions. Unlike many animals, plants are sessile, which imposes distinct constraints on how they interact with their neighbors. Cooperation in plants often manifests as shared resource use (Duvick et al. 2004, Weiner et al. 2017) or mutual benefits such as nutrient facilitation (Li et al. 2014), and lacks the direct behavioral and communicative coordination observed in animals. Instead, it operates through indirect mechanisms, such as root exudates (Li et al. 2016) or resource partitioning (Engbersen et al. 2021). Recent theoretical and methodological advances in the field of plant evolutionary ecology allowed to obtain a better understanding of the genetic structure of cooperation, notably through the dissection of genetic effects into direct and indirect genetic effects (Montazeaud et al. 2023) and the development of a mechanism-agnostic method to quantify cooperativeness in plants (Wuest et al. 2022).

### 7.2.1. Cooperation, competition and trade-offs in genotypic strategies

From both a phenotypic and a genetic perspective, cooperation and competition in plants are shaped by complex trade-offs. Traits that enhance individual competitiveness – such as rapid growth or pre-emptying resource capture – often come at the expense of group performance, especially in dense populations (Weiner et al. 2010). Conversely, genotypes exhibiting more cooperative traits, like resource-sharing or moderated competitive responses, could be outcompeted by selfish neighbors but may be favored in breeding schemes that select, at least partially, among groups of closely related lines (Denison 2012, Weiner 2019). The genetic architecture

underpinning these strategies are not well explored yet, and the work presented in Chapter 4 provides an intriguing starting point for more in-depth studies about the genetics of cooperation in crops and their influence on the phenotypic exhibition of growth strategies and trade-offs.

Cooperative strategies typically involve costs, such as reduced growth potential or slower reproductive rates, which can only be sustained under specific ecological conditions (e.g. drought stress in Chapter 4). Therefore, identifying these trade-offs has the potential to highlight the evolutionary trajectory that lead to cooperative behaviors. The discovery that cooperative strategies can persist in different settings (e.g. in natural populations of *A. thaliana*, as in Wuest et al. 2022, or in a population of a crop plant shaped by breeding as shown here) suggests that social traits might emerge as evolutionary exaptations (Gould and Vrba 1982), arising through the co-option of pre-existing traits unrelated to cooperation. The identification of cooperation-related trade-offs could thus give hints about the nature of such pre-existing traits. Overall, these findings bridge plant evolutionary agroecology with broader evolutionary theories, offering avenues for future exploration in the genetics of cooperation and competition.

### **7.3. From genotypes to communities**

The transition from genotypic-level processes to community-level outcomes highlights the intricate interplay between diversity, complementarity, and collective performance. Understanding these dynamics is crucial for both ecological theory and the practical management of agricultural systems, where genotypic diversity can enhance ecosystem functioning (Finn et al. 2013, Brooker et al. 2023) and stabilize crop yields (Proulx et al. 2010, Morin et al. 2014, Wuest et al. 2021).

#### **7.3.1. Complementarity and variety mixtures**

In agricultural systems, genotypic diversity within a crop species can significantly influence productivity (Barot et al. 2017, Reiss and Drinkwater 2018). Variety mixtures, where multiple genotypes are sown together, exemplify this principle. Mixtures often outperform monocultures due to complementary resource-use strategies among genotypes. For example, genotypes with contrasting root architectures or nutrient uptake efficiencies can improve overall resource acquisition (Bakker et al. 2021, Homulle et al. 2022). Moreover, genotypic diversity can buffer against environmental stressors, such as diseases (Kristoffersen et al. 2020b, 2020a) or fluctuating climatic conditions (Morin et al. 2014, Engbersen et al. 2021). And a diverse mixture of genotypes reduces the risk of community collapse if one genotype is particularly susceptible to a stressor. This

stabilizing effect, known as insurance effect in biodiversity - ecosystem functioning (BEF) studies (Isbell et al. 2018), underscores the value of maintaining diversity to safeguard productivity over time.

However, the success of variety mixtures depends on the balance between complementarity and competition (Montazeaud et al. 2018). Therefore, designing well performing mixtures requires a careful understanding of the trade-offs between individual performance and community yield, informed by insights about genotypic competitiveness and allocation strategies. Chapter 5 and Chapter 6 focused mainly on the identification of novel measures and methods to predict crop variety mixture performances from monoculture stands. Yet, the integration of the genotypic competitiveness of the mixture components as additional mixture performance predictors is still wanting. The tools presented in Chapter 4 and Chapter 6 however could be combined. This might provide novel insights into the relationships between genotypic competitiveness and inter-genotypic complementarity.

## 7.4. Ideotypes and ideotype breeding

The concept of the ideotype – a hypothetical “ideal plant” tailored to specific environmental and management conditions – has a long history. Introduced by Donald in the 1960s, the ideotype aimed to guide breeding toward traits that optimize yield in a predictable way rather than simply enhancing individual plant performance without knowing *why* the new plant yields better than its ancestors (Donald 1968). Interestingly, already in his original publication Donald postulated the successful crop ideotype to be a “weak competitor”. In practice, however, directed breeding for reduced competition has remained rare. Indeed, it has long been unclear if considering certain crop varieties to be “cooperative” is substantially an anthropomorphic interpretation without practical implications or if it entails real-world value in breeding. At the same time, the development of “cooperative” traits like dwarfing in rice seems to have been more a by-product of other breeding goals like the reduction of lodging (Ferrero-Serrano et al. 2019, Gaur et al. 2020, Jiao et al. 2024). The work presented in Chapter 4 resolves this issue by showing that cooperativeness can be measured in a population of soybean varieties without making assumptions about the underlying traits or mechanisms. The genotypic variation in cooperativeness makes it a valid target for plant breeding, and the finding that the identified variation in cooperativeness correlates with genetic variation at specific genomic positions might even allow for breeding using modern techniques like marker assisted selection.

#### **7.4.1. Ideotypes in the context of soybean**

In soybean, ideotypes have traditionally been defined by traits such as erect upper leaves, reduced branching, short internode lengths, high internode number, steep angled root systems, early canopy cover and late leaf senescence (Kokubun 1988, Roth et al. 2022, Lyu et al. 2023). Even though rarely stated explicitly, many of these are traits that reduce excessive competition among neighbors. Ideotypes that exhibit traits promoting resource sharing or complementarity rather than competitive dominance can improve overall yield stability and efficiency. Indeed, erect upper leaves, reduced branching and short internodes (i.e. shorter plants) reduce light competition, and steep angled root systems reduce competition for belowground resources. Hence, these traits could enhance collective performance, aligning with the principles of community-level optimization discussed earlier. A clearer focus on the “communal” characteristics of the crop ideotype and trade-offs between individual and group level performance in ideotype breeding might allow to identify other more subtle target traits (Weiner 2019).

The potential of ideotypes to decrease within-field competition is especially relevant in high-density cropping systems – and in soybean, where excessive intra-specific competition often limits yield potential under intensive cultivation and the breeding progress during the Green Revolution has been comparably small (Suhre et al. 2014, Liu et al. 2020), it is particularly crucial. Testing ideotypes in this context requires moving beyond single-genotype evaluations toward community-level trials. Advances in high-throughput phenotyping now enable such evaluations (Roth et al. 2022, Jiang and Zhu 2024), offering new opportunities to refine and validate ideotype concepts under realistic agricultural conditions.

#### **7.4.2. Ideotypes and sustainable agroecosystems**

In the broader framework of sustainable agroecosystems, ideotypes must address more than yield maximization. Traits that enhance ecosystem services – such as nitrogen fixation, maintenance of soil fertility, or resilience to climatic variability – should be essential components of modern ideotype design. In soybean, communal ideotypes could help transition cropping systems toward greater sustainability without sacrificing productivity: a Green Revolution for soybean (Liu et al. 2020). Moreover, the ideotype approach aligns well with diversification strategies such as variety mixtures or intercropping (Borg et al. 2015, Louarn et al. 2020). By breeding plants optimized for cooperative performance within diverse communities, ideotype breeding could reduce reliance on external inputs, such as fertilizers and pesticides, while maintaining high yields. This synergy

between ideotypes and agroecological principles highlights their relevance in meeting the dual challenges of food security and environmental sustainability.

## 7.5. Limitations

While shedding light on many interesting and highly relevant questions, it is nevertheless important to keep in mind the limitations of the performed investigations. In Chapter 3, only four genotypes were used. It is unknown if the observation that the genotypes do not differ between the treatments would still hold if more genotypes were used. Moreover, the study didn't comprise any temporal dynamics, while competition is not a static process but happens over time. Therefore, we expect the temporal dynamics of competition to play an important role in the relative resource allocation towards the plant's organs and thus the occurrence of a tragedy of the commons.

In Chapter 4, even though a large population of soybean genotypes was used, all these genotypes were early maturing. This was necessary, as later maturing genotypes would not have reached maturity in the relatively short vegetation period available in Switzerland. Nevertheless, expanding the study to later maturing genotypes would be important to make the results more relevant for regions with longer vegetation phases, like the tropics or the sub-tropics. Additionally, in the same study a candidate locus potentially influencing soybean cooperativeness and water use efficiency was identified. However, the role of the respective gene was not validated, and the mechanisms with which the protein coded by the gene might influence the observed behaviors was not assessed. This could be done either by genetically modifying soybean plants and performing near-isogenic lines (NILs) experiments to quantitatively test the effect of the candidate gene. Alternatively, a first step towards this could consist in performing a crossed competition and drought stress experiment using NILs with a model plant which is simpler to genetically modify than soybean, such as *A. thaliana*.

In Chapter 6 finally, the differences between the environments in which the monocultures were grown were not specified. This means that the reaction norms that were calculated rely on many environmental factors and it is not possible to trace the effect back to a few major driving factors. Moreover, the way in which the measurements were processed and the algorithms that were chosen to predict variety mixture performance impede to explain the underlying mechanisms or the plant traits driving these mechanisms. However, the main limitation of Chapter 6 is also its fundamental message: the analysis of the relationship between community-level effects (i.e. biodiversity effects) and genotype-level traits of the communities' components is challenging, and apparently promising results can generate as artifacts through multiple statistical processes (i.e.

data leakage and error propagation). It requires careful model evaluation and cross-validation to ensure the validity of the results and the avoidance of artifacts – and results of past and future BEF studies have to be scrutinized in detail prior to base novel hypotheses on them.

## 7.6. Concluding remarks and future prospects

The work presented here aims at bridging ecology and agricultural sciences to investigate fundamental ecological questions within a model ecosystem provided by crops - and exploring novel approaches to solve pressing agricultural problems with guidance from ecology and evolutionary biology. To do so, an interdisciplinary, multi-faceted approach entailing large-scale experiments, original experimental design, advanced statistics, genetic analyses and machine learning was chosen. The implications of this work both for agriculture and ecology are extensive, encompassing advancements in fundamental ecology as well as applications in agricultural practice and breeding strategies.

Chapter 3 shows that (1) tragedies of the commons can arise upon competition between soybean plants, (2) competition cues might be processed hierarchically in soybean, with above-ground competition cues being weighed higher than belowground competition cues and (3) explicitly accounting for inter-plant distance in plant-plant interaction pot experiments allows to overcome the long debated confounding of competition treatments with available pot volume. Chapter 4 shows that (1) there is variation in competitiveness within a representative soybean population, (2) this competitiveness trades off with a higher susceptibility to drought stress and (3) the variation in competitiveness is associated with variation in specific genomic regions of competing genotypes.

Chapter 5 provides a theoretical framework for the application of ecological principles to guide the development of variety mixtures, and Chapter 6 shows some of the well-hidden methodological pitfalls that can arise when doing so.

A more nuanced understanding of plant competition, cooperation and complementarity, ranging from the genetic factors involved in plant competition, to the shifts in allocation strategies upon competition cues observed in individual plants, to competitive strategies seen at genotype level and up to diversity-productivity relationships studied in multi-genotypic communities, can advance ecology and ecological theory. At the same time, the identification of variability in soybean cooperativeness, linked to drought susceptibility and associated with a small genetic region, is an important advance directly applicable to breeding and with high agronomic potential. Additionally, the methodological advances in the prediction of variety mixture performance

from multi-environmental monoculture stands is an important step towards more successful crop variety mixture development. Especially breeding for more cooperative soybean genotypes is strongly facilitated by this work.

At the same time, this work raises many intriguing new research questions. For example, the role of plant phenotypic plasticity in shaping both inter- and intra-genotypic plant-plant interactions is still to investigated. Moreover, both the relationship between drought susceptibility and competitiveness and the role of the identified candidate gene is still largely unknown. Additionally, this work could provide the impetus to develop more refined game theoretical plant-plant competition models, which could give insight in the hypothesized hierarchical competition signal processing in soybean. Similarly, this work can provide inspiration for the refinement of our understanding of biodiversity-ecosystem functioning relationships. The high degree of interdisciplinarity of this work finally clearly demonstrates the great potential of evolutionary agroecology.

## **8. Acknowledgments**

Thank you to Samuel Wuest for trusting me with this ambitious project. Your enthusiasm for science, your support and your expertise where both admired and appreciated during the last four years. Thank you to Pascal Niklaus for the many inputs and discussions. One of the things I value highest in my PhD is how many thing both of you taught me about agroecology, genetics, breeding, experimental design, statistics and many other topics. You both allowed me to become an independent and confident researcher.

Thank you to my committee members, Niels Anten and Jordi Bascompte for your advice and collaboration.

Thank you to the Breeding Research Group at Agroscope for many interesting discussions. Special thanks go to Lukas Vonmetz and Simon Jäggi, who over the years became both fundamental collaborators in this project - helping conceptualizing and carrying out the experiments - and friends. Thanks to Jürgen Krauss, without your advice and help a pot experiment with almost 4000 plants would not have been possible.

Thank you to the members of the IEU at the University of Zurich for the many coffee breaks. A special thank you to Simon Landauer for sharing the office with me for more than three years, and becoming a dear friend while jointly experiencing the emotional rollercoaster of a PhD. Thank you to Noelle Koenig for spending a year working with me, and providing valuable data to the project.

Thank you to all my friends who spent time with me during these four years, being it on a long bike ride, while climbing a mountain, playing music or having a nice dinner. A special thank you goes to Flavian Tschurr: the many early morning coffees, accompanied by fruitful discussions about both our relative research projects and our lives in general, were highly appreciated and helpful.

Thank you to my parents and my siblings who always supported me and always were around when I needed something. I appreciated the time spent at home among the beautiful vineyards and the fun and interesting discussions we always have with a good meal and a good glass of wine.

Finally, the greatest thank you goes to Sarah - for everything.

## **8.1. Funding**

This work was supported by the Swiss National Science Foundation, project 310030\_192537. The author acknowledges support by the University of Zurich Research Priority Program Global Change and Biodiversity. We thank Dario Fossati (Agroscope) for fruitful discussions on pure-line breeding and mixture development for Chapter 5.

## 9. Bibliography

- Alagele S M, Jose S, Anderson S H, Udawatta R P (2021) Hydraulic Lift: Processes, Methods, and Practical Implications for Society. *Agroforestry Systems* 95:641–657. <https://doi.org/10.1007/s10457-021-00614-w>
- Alexander J M (2023) Evolutionary Game Theory, 1st edn.. Cambridge University Press
- Amyntas A, Berti E, Gauzens B, et al (2023) Niche Complementarity among Plants and Animals Can Alter the Biodiversity–Ecosystem Functioning Relationship. *Functional Ecology* 37:2652–2665. <https://doi.org/10.1111/1365-2435.14419>
- Annicchiarico P, Collins R P, De Ron A M, et al (2019) Chapter Three - Do We Need Specific Breeding for Legume-Based Mixtures?. *Advances in Agronomy* 157:141–215
- Anten N P, Vermeulen P J (2016) Tragedies and Crops: Understanding Natural Selection To Improve Cropping Systems. *Trends in Ecology & Evolution* 31:429–439. <https://doi.org/10.1016/j.tree.2016.02.010>
- Anyia A (2004) Water-Use Efficiency, Leaf Area and Leaf Gas Exchange of Cowpeas under Mid-Season Drought. *European Journal of Agronomy* 20:327–339. [https://doi.org/10.1016/S1161-0301\(03\)00038-8](https://doi.org/10.1016/S1161-0301(03)00038-8)
- Aschehoug E T, Brooker R, Atwater D Z, et al (2016) The Mechanisms and Consequences of Interspecific Competition Among Plants. *Annual Review of Ecology, Evolution, and Systematics* 47:263–281. <https://doi.org/10.1146/annurev-ecolsys-121415-032123>
- Bakker L M, Barry K E, Mommer L, Van Ruijven J (2021) Focusing on Individual Plants to Understand Community Scale Biodiversity Effects: The Case of Root Distribution in Grasslands. *Oikos* 130:1954–1966. <https://doi.org/10.1111/oik.08113>
- Ballaré C L, Pierik R (2017) The Shade-Avoidance Syndrome: Multiple Signals and Ecological Consequences: Neighbour Detection and Shade Avoidance. *Plant, Cell & Environment* 40:2530–2543. <https://doi.org/10.1111/pce.12914>
- Barot S, Allard V, Cantarel A, et al (2017) Designing Mixtures of Varieties for Multifunctional Agriculture with the Help of Ecology. A Review. *Agronomy for Sustainable Development* 37:13. <https://doi.org/10.1007/s13593-017-0418-x>
- Barry K E, Mommer L, Ruijven J van, et al (2019) The Future of Complementarity: Disentangling Causes from Consequences. *Trends in Ecology & Evolution* 34:167–180. <https://doi.org/10.1016/j.tree.2018.10.013>
- Bartlett M K, Scoffoni C, Sack L (2012) The Determinants of Leaf Turgor Loss Point and Prediction of Drought Tolerance of Species and Biomes: A Global Meta-analysis. *Ecology Letters* 15:393–405. <https://doi.org/10.1111/j.1461-0248.2012.01751.x>
- Barton K E, Bowers M D (2006) Neighbor Species Differentially Alter Resistance Phenotypes in *Plantago*. *Oecologia* 150:442–452. <https://doi.org/10.1007/s00442-006-0531-z>
- Bates D, Noack A, Kornblith S, et al (2023) JuliaStats/GLM.Jl: V1.9.0
- Beillouin D, Ben-Ari T, Malézieux E, et al (2021) Positive but Variable Effects of Crop Diversification on Biodiversity and Ecosystem Services. *Global Change Biology* 27:4697–4710. <https://doi.org/10.1111/gcb.15747>

- Bennett J, Blumenthal D B, Grimm D G, et al (2024) Guiding Questions to Avoid Data Leakage in Biological Machine Learning Applications. *Nature Methods* 21:1444–1453. <https://doi.org/10.1038/s41592-024-02362-y>
- Beukert U, Thorwarth P, Zhao Y, et al (2020) Comparing the Potential of Marker-Assisted Selection and Genomic Prediction for Improving Rust Resistance in Hybrid Wheat. *Frontiers in Plant Science* 11:. <https://doi.org/10.3389/fpls.2020.594113>
- Bharath P, Gahir S, Raghavendra A S (2021) Abscisic Acid-Induced Stomatal Closure: An Important Component of Plant Defense Against Abiotic and Biotic Stress. *Frontiers in Plant Science* 12:615114. <https://doi.org/10.3389/fpls.2021.615114>
- Blaom A, Lienart T, Simillides Y, et al (2024) Alan-Turing-Institute/MLJ.Jl: V0.20.3
- Blonder B (2018) Hypervolume Concepts in Niche- and Trait-Based Ecology. *Ecography* 41:1441–1455. <https://doi.org/10.1111/ecog.03187>
- Bonhomme M, Fariello M I, Navier H, et al (2019) A Local Score Approach Improves GWAS Resolution and Detects Minor QTL: Application to *Medicago truncatula* Quantitative Disease Resistance to Multiple *Aphanomyces euteiches* Isolates. *Heredity* 123:517–531. <https://doi.org/10.1038/s41437-019-0235-x>
- Bonnin I, Bonneuil C, Goffaux R, et al (2014) Explaining the Decrease in the Genetic Diversity of Wheat in France over the 20th Century. *Agriculture, Ecosystems & Environment* 195:183–192. <https://doi.org/10.1016/j.agee.2014.06.003>
- Borg J, Kiær L, Lecarpentier C, et al (2018) Unfolding the Potential of Wheat Cultivar Mixtures: A Meta-Analysis Perspective and Identification of Knowledge Gaps. *Field Crops Research* 221:298–313. <https://doi.org/10.1016/j.fcr.2017.09.006>
- Borg J, Enjalbert J, Gauffreteau A (2015) Concevoir Des Associations Variétales de Blé Par l'Idéotype Participatif. In: Colloque Présentant Les Méthodes et Résultats Du Projet Climage (Métaprogramme ACCAF). INRA, p 223–p
- Bourke P M, Evers J B, Bijma P, et al (2021) Breeding Beyond Monoculture: Putting the “Intercrop” Into Crops. *Frontiers in Plant Science* 12:. <https://doi.org/10.3389/fpls.2021.734167>
- Brooker R W, Hawes C, Iannetta P P M, et al (2023) Plant Diversity and Ecological Intensification in Crop Production Systems. *Journal of Plant Ecology* 16:rtad15. <https://doi.org/10.1093/jpe/rtad015>
- Brooker R W, Bennett A E, Cong W-F, et al (2015) Improving Intercropping: A Synthesis of Research in Agronomy, Plant Physiology and Ecology. *New Phytologist* 206:107–117. <https://doi.org/10.1111/nph.13132>
- Brooker R, Brown L K, George T S, et al (2022) Active and Adaptive Plasticity in a Changing Climate. *Trends in Plant Science* 27:717–728. <https://doi.org/10.1016/j.tplants.2022.02.004>
- Brown J K M (2015) Durable Resistance of Crops to Disease: A Darwinian Perspective. *Annual Review of Phytopathology* 53:513–539. <https://doi.org/10.1146/annurev-phyto-102313-045914>
- Bukowski A R, Petermann J S (2014) Intraspecific Plant–Soil Feedback and Intraspecific Overyielding in *Arabidopsis thaliana*. *Ecology and Evolution* 4:2533–2545. <https://doi.org/10.1002/ece3.1077>

- Cabal C (2022) Root Tragedy of the Commons: Revisiting the Mechanisms of a Misunderstood Theory. *Frontiers in Plant Science* 13:960942. <https://doi.org/10.3389/fpls.2022.960942>
- Cabal C, Martínez-García R, De Castro Aguilar A, et al (2020) The Exploitative Segregation of Plant Roots. *Science* 370:1197–1199. <https://doi.org/10.1126/science.aba9877>
- Cadotte M W, Arnillas C A, Livingstone S W, Yasui S-L E (2015) Predicting Communities from Functional Traits. *Trends in Ecology & Evolution* 30:510–511. <https://doi.org/10.1016/j.tree.2015.07.001>
- Calderini D F, Slafer G A (1998) Changes in Yield and Yield Stability in Wheat during the 20th Century. *Field Crops Research* 57:335–347. [https://doi.org/10.1016/S0378-4290\(98\)00080-X](https://doi.org/10.1016/S0378-4290(98)00080-X)
- Carciochi W D, Schwalbert R, Andrade F H, et al (2019) Soybean Seed Yield Response to Plant Density by Yield Environment in North America. *Agronomy Journal* 111:1923–1932. <https://doi.org/10.2134/agronj2018.10.0635>
- Cardinale B J, Matulich K L, Hooper D U, et al (2011) The Functional Role of Producer Diversity in Ecosystems. *American Journal of Botany* 98:572–592. <https://doi.org/10.3732/ajb.1000364>
- Cardinale B J, Wright J P, Cadotte M W, et al (2007) Impacts of Plant Diversity on Biomass Production Increase through Time Because of Species Complementarity. *Proceedings of the National Academy of Sciences* 104:18123–18128. <https://doi.org/10.1073/pnas.0709069104>
- Carlsson K, Karrasch D, Bauer N, et al (2024) KristofferC/NearestNeighbors.Jl: V0.4.20
- Casanova-Sáez R, Mateo-Bonmatí E, Simura J, et al (2022) Inactivation of the Entire *Arabidopsis* Group II GH3s Confers Tolerance to Salinity and Water Deficit. *New Phytologist* 235:263–275. <https://doi.org/10.1111/nph.18114>
- Cassman K G, Grassini P (2013) Can There Be a Green Revolution in Sub-Saharan Africa without Large Expansion of Irrigated Crop Production?. *Global Food Security* 2:203–209. <https://doi.org/10.1016/j.gfs.2013.08.004>
- Cassman K G, Grassini P (2020) A Global Perspective on Sustainable Intensification Research. *Nature Sustainability* 3:262–268. <https://doi.org/10.1038/s41893-020-0507-8>
- Chacón-Labella J, Hinojo-Hinojo C, Bohner T, et al (2023) How to Improve Scaling from Traits to Ecosystem Processes. *Trends in Ecology & Evolution* 38:228–237. <https://doi.org/10.1016/j.tree.2022.10.007>
- Chai Y, Pardey P G, Silverstein K A T (2022) Scientific Selection: A Century of Increasing Crop Varietal Diversity in US Wheat. *Proceedings of the National Academy of Sciences* 119:e2210773119. <https://doi.org/10.1073/pnas.2210773119>
- Chalmandrier L, Münkemüller T, Colace M-P, et al (2017) Spatial Scale and Intraspecific Trait Variability Mediate Assembly Rules in Alpine Grasslands. *Journal of Ecology* 105:277–287. <https://doi.org/10.1111/1365-2745.12658>
- Chase J M, Leibold M A (2003) Ecological Niches: Linking Classical and Contemporary Approaches. University of Chicago Press
- Chen B J W, During H J, Vermeulen P J, et al (2015) Corrections for Rooting Volume and Plant Size Reveal Negative Effects of Neighbour Presence on Root Allocation in Pea. *Functional Ecology* 29:1383–1391. <https://doi.org/10.1111/1365-2435.12450>

- Chen B J W, Huang L, During H J, et al (2021a) No Neighbour-Induced Increase in Root Growth of Soybean and Sunflower in Mesh-Divider Experiments after Controlling for Nutrient Concentration and Soil Volume. *AoB PLANTS* 13:plab20. <https://doi.org/10.1093/aobpla/plab020>
- Chen Q, Bai L, Wang W, et al (2021b) COP1 Promotes ABA-induced Stomatal Closure by Modulating the Abundance of ABI/HAB and AHG3 Phosphatases. *New Phytologist* 229:2035–2049. <https://doi.org/10.1111/nph.17001>
- Chen Z, Liu Y, Yin Y, et al (2019) Development of Dwarfish and Yield-Effective GM Maize through Passivation of Bioactive Gibberellin. *Transgenic Research* 28:589–599. <https://doi.org/10.1007/s11248-019-00172-z>
- Cheng S, Feng C, Wingen L U, et al (2024) Harnessing Landrace Diversity Empowers Wheat Breeding. *Nature* 632:823–831. <https://doi.org/10.1038/s41586-024-07682-9>
- Cober E R, Morrison M J, Ma B, Butler G (2005) Genetic Improvement Rates of Short-Season Soybean Increase with Plant Population. *Crop Science* 45:1029–1034. <https://doi.org/10.2135/cropsci2004.0232>
- Cohen D (1970) The Expected Efficiency of Water Utilization in Plants under Different Competition and Selection Regimes.. *Israel Journal of Botany* 19:50–54
- Colwell R K, Futuyma D J (1971) On the Measurement of Niche Breadth and Overlap. *Ecology* 52:567–576. <https://doi.org/10.2307/1934144>
- Condon A G (2004) Breeding for High Water-Use Efficiency. *Journal of Experimental Botany* 55:2447–2460. <https://doi.org/10.1093/jxb/erh277>
- Cook B I, Mankin J S, Anchukaitis K J (2018) Climate Change and Drought: From Past to Future. *Current Climate Change Reports* 4:164–179. <https://doi.org/10.1007/s40641-018-0093-2>
- Cook-Patton S C, McArt S H, Parachnowitsch A L, et al (2011) A Direct Comparison of the Consequences of Plant Genotypic and Species Diversity on Communities and Ecosystem Function. *Ecology* 92:915–923. <https://doi.org/10.1890/10-0999.1>
- Cooper H F, Grady K C, Cowan J A, et al (2019) Genotypic Variation in Phenological Plasticity: Reciprocal Common Gardens Reveal Adaptive Responses to Warmer Springs but Not to Fall Frost. *Global Change Biology* 25:187–200. <https://doi.org/10.1111/gcb.14494>
- Cowden R J, Shah A N, Lehmann L M, et al (2020) Nitrogen Fertilizer Effects on Pea-Barley Intercrop Productivity Compared to Sole Crops in Denmark. *Sustainability* 12:9335. <https://doi.org/10.3390/su12229335>
- Cowger C, Mundt C C (2002) Effects of Wheat Cultivar Mixtures on Epidemic Progression of *Septoria Tritici* Blotch and Pathogenicity of *Mycosphaerella Graminicola*. *Phytopathology®* 92:617–623. <https://doi.org/10.1094/PHYTO.2002.92.6.617>
- Crawford K M, Whitney K D (2010) Population Genetic Diversity Influences Colonization Success. *Molecular Ecology* 19:1253–1263. <https://doi.org/10.1111/j.1365-294X.2010.04550.x>
- D'Andrea R, Ostling A (2016) Challenges in Linking Trait Patterns to Niche Differentiation. *Oikos* 125:1369–1385. <https://doi.org/10.1111/oik.02979>

- Dawson J C, Goldringer I (2011) Breeding for Genetically Diverse Populations: Variety Mixtures and Evolutionary Populations. *Organic Crop Breeding* 77–98
- De La Mata R, Zas R (2023) Plasticity in Growth Is Genetically Variable and Highly Conserved across Spatial Scales in a Mediterranean Pine. *New Phytologist* 240:542–554. <https://doi.org/10.1111/nph.19158>
- De Tombeur F, Péliissier R, Shih A, et al (2023) Growth–Defence Trade-off in Rice: Fast-Growing and Acquisitive Genotypes Have Lower Expression of Genes Involved in Immunity. *Journal of Experimental Botany* 74:3094–3103. <https://doi.org/10.1093/jxb/erad071>
- Delzon S (2015) New Insight into Leaf Drought Tolerance. *Functional Ecology* 29:1247–1249. <https://doi.org/10.1111/1365-2435.12500>
- Denison R F (2012) Darwinian Agriculture: How Understanding Evolution Can Improve Agriculture. Lightning Source UK Ltd., Princeton University Press, Milton Keynes UK, Princeton, N.J.
- Donald C M (1968) The Breeding of Crop Ideotypes. *Euphytica* 17:385–403. <https://doi.org/10.1007/BF00056241>
- Doyle J, Doyle J (1987) A Rapid DNA Isolation Procedure for Small Quantities of Fresh Leaf Tissue. *Phytochemical Bulletin* 19:11–15
- Du H (first), Wu, Nai, Fu, Jing, et al (2012) A GH3 Family Member, OsGH3-2, Modulates Auxin and Abscisic Acid Levels and Differentially Affects Drought and Cold Tolerance in Rice Cold Tolerance in Rice. *Journal of Experimental Botany* 63:6467–6480. <https://doi.org/10.1093/jxb/ers300>
- Duvick D N, Smith J S C, Cooper M (2004) Long-Term Selection in a Commercial Hybrid Maize Breeding Program. *Plant Breeding Reviews: Part 2: Long-term Selection: Crops, Animals, and Bacteria* 24:109–151
- Díaz S, Kattge J, Cornelissen J H C, et al (2016) The Global Spectrum of Plant Form and Function. *Nature* 529:167–171. <https://doi.org/10.1038/nature16489>
- Ebeling A, Pompe S, Baade J, et al (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>
- Elli E F, Edwards J, Yu J, et al (2023) Maize Leaf Angle Genetic Gain Is Slowing down in the Last Decades. *Crop Science* 63:3520–3533. <https://doi.org/10.1002/csc2.21111>
- Engbersen N, Brooker R W, Stefan L, et al (2021) Temporal Differentiation of Resource Capture and Biomass Accumulation as a Driver of Yield Increase in Intercropping. *Frontiers in Plant Science* 12:668803. <https://doi.org/10.3389/fpls.2021.668803>
- Feng Y-Y, Richards R A, Jin Y, et al (2022) Yield and Water-Use Related Traits in Landrace and New Soybean Cultivars in Arid and Semi-Arid Areas of China. *Field Crops Research* 283:108559. <https://doi.org/10.1016/j.fcr.2022.108559>
- Ferrero-Serrano Á, Cantos C, Assmann S M (2019) The Role of Dwarfing Traits in Historical and Modern Agriculture with a Focus on Rice. *Cold Spring Harbor Perspectives in Biology* 11:a34645. <https://doi.org/10.1101/cshperspect.a034645>
- Finckh M (1992) Stripe Rust, Yield, and Plant Competition in Wheat Cultivar Mixtures. *Phytopathology* 82:.. <https://doi.org/10.1094/Phyto-82-905>

- Finckh M R, Gacek E S, Goyeau H, et al (2000) Cereal Variety and Species Mixtures in Practice, with Emphasis on Disease Resistance. *Agronomie* 20:813–837. <https://doi.org/10.1051/agro:2000177>
- Finckh M R, Van Bruggen A H C, Tamm L (eds) (2017) Plant Diseases and Their Management in Organic Agriculture. The American Phytopathological Society
- Finn J A, Kirwan L, Connolly J, et al (2013) Ecosystem Function Enhanced by Combining Four Functional Types of Plant Species in Intensively Managed Grassland Mixtures: A 3-Year Continental-Scale Field Experiment. *Journal of Applied Ecology* 50:365–375. <https://doi.org/10.1111/1365-2664.12041>
- Fischer R A (, Edmeades G O (2010) Breeding and Cereal Yield Progress. *Crop Science* 50:. <https://doi.org/10.2135/cropsci2009.10.0564>
- Fixt E, Hodges J L (1989) Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties
- Flor H H (1971) Current Status of the Gene-For-Gene Concept. *Annual Review of Phytopathology* 9:275–296. <https://doi.org/10.1146/annurev.py.09.090171.001423>
- Fones H N, Bebber D P, Chaloner T M, et al (2020) Threats to Global Food Security from Emerging Fungal and Oomycete Crop Pathogens. *Nature Food* 1:332–342. <https://doi.org/10.1038/s43016-020-0075-0>
- Forst E, Enjalbert J, Allard V, et al (2019) A Generalized Statistical Framework to Assess Mixing Ability from Incomplete Mixing Designs Using Binary or Higher Order Variety Mixtures and Application to Wheat. *Field Crops Research* 242:107571. <https://doi.org/10.1016/j.fcr.2019.107571>
- Frankel O H (1939) Analytical Yield Investigations on New Zealand Wheat: IV. Blending Varieties of Wheat. *The Journal of Agricultural Science* 29:249–261. <https://doi.org/10.1017/S0021859600051777>
- Franks S J (2011) Plasticity and Evolution in Drought Avoidance and Escape in the Annual Plant *Brassica rapa*. *New Phytologist* 190:249–257. <https://doi.org/10.1111/j.1469-8137.2010.03603.x>
- Franks S J, Weber J J, Aitken S N (2014) Evolutionary and Plastic Responses to Climate Change in Terrestrial Plant Populations. *Evolutionary Applications* 7:123–139. <https://doi.org/10.1111/eva.12112>
- Funk J L, Wolf A A (2016) Testing the Trait-Based Community Framework: Do Functional Traits Predict Competitive Outcomes?. *Ecology* 97:2206–2211. <https://doi.org/10.1002/ecy.1484>
- Galli G, Gezan S A, Murillo D A, Murray D (2022) ASRgwas: An R Package to Perform Complex Genome-Wide Association Studied (GWAS)
- Garrett K A, Mundt C C (1999) Epidemiology in Mixed Host Populations. *Phytopathology®* 89:984–990. <https://doi.org/10.1094/PHYTO.1999.89.11.984>
- Gaur V S, Channappa G, Chakraborti M, et al (2020) 'Green Revolution' Dwarf Gene *sdl* of Rice Has Gigantic Impact. *Briefings in Functional Genomics* 19:390–409. <https://doi.org/10.1093/bfgp/elaaa019>

- Gersani M, Brown J S, O'Brien E E, et al (2001) Tragedy of the Commons as a Result of Root Competition: Game Theory and Root Competition. *Journal of Ecology* 89:660–669. <https://doi.org/10.1046/j.0022-0477.2001.00609.x>
- Ghalambor C K, McKay J K, Carroll S P, Reznick D N (2007) Adaptive versus Non-adaptive Phenotypic Plasticity and the Potential for Contemporary Adaptation in New Environments. *Functional Ecology* 21:394–407. <https://doi.org/10.1111/j.1365-2435.2007.01283.x>
- Gibson A K (2022) Genetic Diversity and Disease: The Past, Present, and Future of an Old Idea. *Evolution* 76:20–36. <https://doi.org/10.1111/evo.14395>
- Gibson A K, Nguyen A E (2021) Does Genetic Diversity Protect Host Populations from Parasites? A Meta-Analysis across Natural and Agricultural Systems. *Evolution Letters* 5:16–32. <https://doi.org/10.1002/evl3.206>
- Godoy O, Gómez-Aparicio L, Matías L, et al (2020) An Excess of Niche Differences Maximizes Ecosystem Functioning. *Nature Communications* 11:4180. <https://doi.org/10.1038/s41467-020-17960-5>
- Golan G, Abbai R, Schnurbusch T (2023) Exploring the Trade-off between Individual Fitness and Community Performance of Wheat Crops Using Simulated Canopy Shade. *Plant, Cell & Environment* 46:3144–3157. <https://doi.org/10.1111/pce.14499>
- Gorjanc G, Jenko J, Hearne S J, Hickey J M (2016) Initiating Maize Pre-Breeding Programs Using Genomic Selection to Harness Polygenic Variation from Landrace Populations. *BMC Genomics* 17:30. <https://doi.org/10.1186/s12864-015-2345-z>
- Gottlieb R, Gruntman M (2022) Can Plants Integrate Information on Above-Ground Competition in Their Directional Responses Below Ground?. *Annals of Botany* 130:763–771. <https://doi.org/10.1093/aob/mcac106>
- Gould S J, Vrba E S (1982) Exaptation-A Missing Term in the Science of Form. *Paleobiology* 8:4–15. <https://doi.org/10.1017/S0094837300004310>
- Graham P H, Vance C P (2003) Legumes: Importance and Constraints to Greater Use. *Plant Physiology* 131:872–877. <https://doi.org/10.1104/pp.017004>
- Green-Tracewicz E, Page E R, Swanton C J (2011) Shade Avoidance in Soybean Reduces Branching and Increases Plant-to-Plant Variability in Biomass and Yield Per Plant. *Weed Science* 59:43–49. <https://doi.org/10.1614/WS-D-10-00081.1>
- Griffing B (1956) Concept of General and Specific Combining Ability in Relation to Diallel Crossing Systems. *Australian Journal of Biological Sciences* 9:463–493. <https://doi.org/10.1071/bi9560463>
- Grime J P, Pierce S (2012) The Evolutionary Strategies That Shape Ecosystems, 1st edn.. Wiley
- Gupta A, Rico-Medina A, Caño-Delgado A I (2020) The Physiology of Plant Responses to Drought. *Science* 368:266–269. <https://doi.org/10.1126/science.aaz7614>
- Hardin G (1968) The Tragedy of the Commons. *Science* 162:1243–1248. <https://doi.org/10.1126/science.162.3859.1243>
- Harlan J R (1975) Our Vanishing Genetic Resources. *Science* 188:618–621. <https://doi.org/10.1126/science.188.4188.618>

- Harper J L (1977) Population Biology of Plants. Academic Press, New York, NY
- Haubo R, Christensen B (2018) Cumulative Link Models for Ordinal Regression with the R Package ordinal. <https://api.semanticscholar.org/CorpusID:59572956>
- Hauggaard-Nielsen H, Ambus P, Jensen E S (2001) Interspecific Competition, N Use and Interference with Weeds in Pea-Barley Intercropping. *Field Crops Research* 70:101–109. [https://doi.org/10.1016/S0378-4290\(01\)00126-5](https://doi.org/10.1016/S0378-4290(01)00126-5)
- Haupt M, Schmid K (2022) Using Landscape Genomics to Infer Genomic Regions Involved in Environmental Adaptation of Soybean Genebank Accessions
- He Z, Webster S, He S Y (2022) Growth–Defense Trade-Offs in Plants. *Current Biology* 32:R634–R639. <https://doi.org/10.1016/j.cub.2022.04.070>
- Hector A, Schmid B, Beierkuhnlein C, et al (1999) Plant Diversity and Productivity Experiments in European Grasslands. *Science* 286:1123–1127. <https://doi.org/10.1126/science.286.5442.1123>
- Hedden P (2003) The Genes of the Green Revolution. *Trends in Genetics* 19:5–9. [https://doi.org/10.1016/S0168-9525\(02\)00009-4](https://doi.org/10.1016/S0168-9525(02)00009-4)
- Heschel M S, Riginos C (2005) Mechanisms of Selection for Drought Stress Tolerance and Avoidance in *Impatiens Capensis* (Balsaminaceae). *American Journal of Botany* 92:37–44. <https://doi.org/10.3732/ajb.92.1.37>
- Hess L, De Kroon H (2007) Effects of Rooting Volume and Nutrient Availability as an Alternative Explanation for Root Self / Non-Self Discrimination. *Journal of Ecology* 95:241–251. <https://doi.org/10.1111/j.1365-2745.2006.01204.x>
- Ho T K (1995) Random Decision Forests. In: Proceedings of 3rd International Conference on Document Analysis and Recognition. IEEE, pp 278–282
- Homulle Z, George T S, Karley A J (2022) Root Traits with Team Benefits: Understanding Belowground Interactions in Intercropping Systems. *Plant and Soil* 471:1–26. <https://doi.org/10.1007/s11104-021-05165-8>
- Hoogenboom M O, Connolly S R (2009) Defining Fundamental Niche Dimensions of Corals: Synergistic Effects of Colony Size, Light, and Flow. *Ecology* 90:767–780. <https://doi.org/10.1890/07-2010.1>
- Hooper D U, Chapin F S, Ewel J J, et al (2005) Effects of Biodiversity on Ecosystem Functioning: A Consensus of Current Knowledge. *Ecological Monographs* 75:3–35. <https://doi.org/10.1890/04-0922>
- Huang C, Sun Z, Wang H, et al (2012) Effects of Wheat Cultivar Mixtures on Stripe Rust: A Meta-Analysis on Field Trials. *Crop Protection* 33:52–58. <https://doi.org/10.1016/j.cropro.2011.11.020>
- Huang T, Döring T F, Zhao X, et al (2024) Cultivar Mixtures Increase Crop Yields and Temporal Yield Stability Globally. A Meta-Analysis. *Agronomy for Sustainable Development* 44:28. <https://doi.org/10.1007/s13593-024-00964-6>
- Huang Y, Chen Y, Castro-Izaguirre N, et al (2018) Impacts of Species Richness on Productivity in a Large-Scale Subtropical Forest Experiment. *Science* 362:80–83. <https://doi.org/10.1126/science.aat6405>

- Huang Y, Schuldt A, Hönig L, et al (2022) Effects of Enemy Exclusion on Biodiversity–Productivity Relationships in a Subtropical Forest Experiment. *Journal of Ecology* 110:2167–2178. <https://doi.org/10.1111/1365-2745.13940>
- Huot B, Yao J, Montgomery B L, He S Y (2014) Growth–Defense Tradeoffs in Plants: A Balancing Act to Optimize Fitness. *Molecular Plant* 7:1267–1287. <https://doi.org/10.1093/mp/ssu049>
- Hutchinson G E (1957) Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22:415–427. <https://doi.org/10.1101/SQB.1957.022.01.039>
- Hutchinson G E (1978) An Introduction to Population Ecology. Yale University Press, New Haven
- IPCC (2022) Climate Change and Land: IPCC Special Report on Climate Change, Desertification, Land Degradation, Sustainable Land Management, Food Security, and Greenhouse Gas Fluxes in Terrestrial Ecosystems, 1st edn.. Cambridge University Press
- Isbell F, Cowles J, Dee L E, et al (2018) Quantifying Effects of Biodiversity on Ecosystem Functioning across Times and Places. *Ecology Letters* 21:763–778. <https://doi.org/10.1111/ele.12928>
- Isbell F, Craven D, Connolly J, et al (2015) Biodiversity Increases the Resistance of Ecosystem Productivity to Climate Extremes. *Nature* 526:574–577. <https://doi.org/10.1038/nature15374>
- Jain H K (2010) Green Revolution: History, Impact and Future. Studium Press LLC, Houston
- Jannink J-L (2010) Dynamics of Long-Term Genomic Selection. *Genetics Selection Evolution* 42:35. <https://doi.org/10.1186/1297-9686-42-35>
- Jennings P R, Herrera R M (1968) Studies on Competition in Rice II. Competition in Segregating Populations. *Evolution* 22:332–336
- Jennings P R, De Jesus J (1968) Studies on Competition in Rice I. Competition in Mixtures of Varieties. *Evolution* 22:119. <https://doi.org/10.2307/2406656>
- Jiang N, Zhu X-G (2024) Modern Phenomics to Empower Holistic Crop Science, Agronomy, and Breeding Research. *Journal of Genetics and Genomics* 51:790–800. <https://doi.org/10.1016/j.jgg.2024.04.016>
- Jiao S, Li Q, Zhang F, et al (2024) Artificial Selection of the Green Revolution Gene *Semidwarf 1* Is Implicated in Upland Rice Breeding. *Journal of Integrative Agriculture* 23:769–780. <https://doi.org/10.1016/j.jia.2023.05.010>
- Kadam N N, Tamilselvan A, Lawas L M, et al (2017) Genetic Control of Plasticity in Root Morphology and Anatomy of Rice in Response to Water Deficit. *Plant Physiology* 174:2302–2315. <https://doi.org/10.1104/pp.17.00500>
- Kellerhals M, Mouron P, Graf B, et al (2003) Mischpflanzungen von Apfelsorten: Einfluss auf Krankheiten, Schädlinge und Wirtschaftlichkeit.. *Schweizer Zeitschrift für Obst- und Weinbau* 13:10–13
- Keuskamp D H, Sasidharan R, Pierik R (2010) Physiological Regulation and Functional Significance of Shade Avoidance Responses to Neighbors. *Plant Signaling & Behavior* 5:655–662. <https://doi.org/10.4161/psb.5.6.11401>

- Khoury C K, Brush S, Costich D E, et al (2022) Crop Genetic Erosion: Understanding and Responding to Loss of Crop Diversity. *New Phytologist* 233:84–118. <https://doi.org/10.1111/nph.17733>
- Kim K D, Kang Y, Kim C (2020) Application of Genomic Big Data in Plant Breeding: Past, Present, and Future. *Plants* 9:1454. <https://doi.org/10.3390/plants9111454>
- Kirungu J N, Magwanga R O, Lu P, et al (2019) Functional Characterization of Gh\_A08G1120 (GH3.5) Gene Reveal Their Significant Role in Enhancing Drought and Salt Stress Tolerance in Cotton. *BMC Genetics* 20:62. <https://doi.org/10.1186/s12863-019-0756-6>
- 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, Skovgaard I M, Østergård H (2012) Effects of Inter-Varietal Diversity, Biotic Stresses and Environmental Productivity on Grain Yield of Spring Barley Variety Mixtures. *Euphytica* 185:123–138. <https://doi.org/10.1007/s10681-012-0640-1>
- Kokubun M (1988) Design and Examination of Soybean Ideotypes. *Jarq-japan Agricultural Research Quarterly* 21:237–243
- Kong X, Li L, Peng P, et al (2023) Wheat Cultivar Mixtures Increase Grain Yield under Varied Climate Conditions. *Basic and Applied Ecology* 69:13–25. <https://doi.org/10.1016/j.baae.2023.03.007>
- Kooyers N J (2015) The Evolution of Drought Escape and Avoidance in Natural Herbaceous Populations. *Plant Science* 234:155–162. <https://doi.org/10.1016/j.plantsci.2015.02.012>
- Kopp E B, Niklaus P A, Wuest S E (2023) Ecological Principles to Guide the Development of Crop Variety Mixtures. *Journal of Plant Ecology* 16:rtad17. <https://doi.org/10.1093/jpe/rtad017>
- Koricheva J, Hayes D (2018) The Relative Importance of Plant Intraspecific Diversity in Structuring Arthropod Communities: A Meta-Analysis. *Functional Ecology* 32:1704–1717. <https://doi.org/10.1111/1365-2435.13062>
- Kovach M J, McCouch S R (2008) Leveraging Natural Diversity: Back through the Bottleneck. *Current Opinion in Plant Biology* 11:193–200. <https://doi.org/10.1016/j.pbi.2007.12.006>
- Kraft N J B, Godoy O, Levine J M (2015) Plant Functional Traits and the Multidimensional Nature of Species Coexistence. *Proceedings of the National Academy of Sciences* 112:797–802. <https://doi.org/10.1073/pnas.1413650112>
- Kristoffersen R, Eriksen L B, Nielsen G C, et al (2022) Management of *Septoria Tritici* Blotch Using Cultivar Mixtures. *Plant Disease* 106:1341–1349. <https://doi.org/10.1094/PDIS-01-21-0069-RE>
- Kristoffersen R, Heick T M, Müller G M, et al (2020a) The Potential of Cultivar Mixtures to Reduce Fungicide Input and Mitigate Fungicide Resistance Development. *Agronomy for Sustainable Development* 40:36. <https://doi.org/10.1007/s13593-020-00639-y>
- Kristoffersen R, Jørgensen L N, Eriksen L B, et al (2020b) Control of *Septoria Tritici* Blotch by Winter Wheat Cultivar Mixtures: Meta-analysis of 19 Years of Cultivar Trials. *Field Crops Research* 249:107696. <https://doi.org/10.1016/j.fcr.2019.107696>

- Kucharik C J, Ramankutty N (2005) Trends and Variability in U.S. Corn Yields Over the Twentieth Century. *Earth Interactions* 9:1–29. <https://doi.org/10.1175/EI098.1>
- Kuhlgert S, Austic G, Zegarac R, et al (2016) MultispeQ Beta: A Tool for Large-Scale Plant Phenotyping Connected to the Open PhotosynQ Network. *Royal Society Open Science* 3:160592. <https://doi.org/10.1098/rsos.160592>
- Kunstler G, Falster D, Coomes D A, et al (2016) Plant Functional Traits Have Globally Consistent Effects on Competition. *Nature* 529:204–207. <https://doi.org/10.1038/nature16476>
- Kunstler G, Lavergne S, Courbaud B, et al (2012) Competitive Interactions between Forest Trees Are Driven by Species' Trait Hierarchy, Not Phylogenetic or Functional Similarity: Implications for Forest Community Assembly. *Ecology Letters* 15:831–840. <https://doi.org/10.1111/j.1461-0248.2012.01803.x>
- Labarthe P, Coléno F, Enjalbert J, et al (2021) Exploration, Exploitation and Environmental Innovation in Agriculture. The Case of Variety Mixture in France and Denmark. *Technological Forecasting and Social Change* 172:121028. <https://doi.org/10.1016/j.techfore.2021.121028>
- Lande R (2009) Adaptation to an Extraordinary Environment by Evolution of Phenotypic Plasticity and Genetic Assimilation. *Journal of Evolutionary Biology* 22:1435–1446. <https://doi.org/10.1111/j.1420-9101.2009.01754.x>
- Lee J, Chin J H, Ahn S N, Koh H-J (2015) Brief History and Perspectives on Plant Breeding. *Current Technologies in Plant Molecular Breeding* 1–14
- Lepik A, Abakumova M, Davison J, et al (2021) Spatial Mapping of Root Systems Reveals Diverse Strategies of Soil Exploration and Resource Contest in Grassland Plants. *Journal of Ecology* 109:652–663. <https://doi.org/10.1111/1365-2745.13535>
- Levitt J (1980) Responses of Plants to Environmental Stresses, 2nd ed. Academic press, Orlando San Diego New York
- Li B, Li Y-Y, Wu H-M, et al (2016) Root Exudates Drive Interspecific Facilitation by Enhancing Nodulation and N<sup>2</sup> Fixation. *Proceedings of the National Academy of Sciences* 113:6496–6501. <https://doi.org/10.1073/pnas.1523580113>
- Li C, Guan H, Jing X, et al (2022) Genomic Insights into Historical Improvement of Heterotic Groups during Modern Hybrid Maize Breeding. *Nature Plants* 8:750–763. <https://doi.org/10.1038/s41477-022-01190-2>
- Li C, Hoffland E, Kuyper T W, et al (2020) Syndromes of Production in Intercropping Impact Yield Gains. *Nature Plants* 6:653–660. <https://doi.org/10.1038/s41477-020-0680-9>
- Li J H, Mazur C A, Berisa T, Pickrell J K (2021a) Low-Pass Sequencing Increases the Power of GWAS and Decreases Measurement Error of Polygenic Risk Scores Compared to Genotyping Arrays. *Genome Research* 31:529–537. <https://doi.org/10.1101/gr.266486.120>
- Li L, Li S-M, Sun J-H, et al (2007) Diversity Enhances Agricultural Productivity via Rhizosphere Phosphorus Facilitation on Phosphorus-Deficient Soils. *Proceedings of the National Academy of Sciences* 104:11192–11196. <https://doi.org/10.1073/pnas.0704591104>
- Li L, Tilman D, Lambers H, Zhang F-S (2014) Plant Diversity and Overyielding: Insights from Belowground Facilitation of Intercropping in Agriculture. *New Phytologist* 203:63–69. <https://doi.org/10.1111/nph.12778>

- Li X-F, Wang Z-G, Bao X-G, et al (2021b) Long-Term Increased Grain Yield and Soil Fertility from Intercropping. *Nature Sustainability* 4:943–950. <https://doi.org/10.1038/s41893-021-00767-7>
- Litrico I, Violle C (2015) Diversity in Plant Breeding: A New Conceptual Framework. *Trends in Plant Science* 20:604–613. <https://doi.org/10.1016/j.tplants.2015.07.007>
- Liu S, Zhang M, Feng F, Tian Z (2020) Toward a “Green Revolution” for Soybean. *Molecular Plant* 13:688–697. <https://doi.org/10.1016/j.molp.2020.03.002>
- Louarn G, Barillot R, Combes D, Escobar-Gutiérrez A (2020) Towards Intercrop Ideotypes: Non-Random Trait Assembly Can Promote Overyielding and Stability of Species Proportion in Simulated Legume-Based Mixtures. *Annals of Botany* 126:671–685. <https://doi.org/10.1093/aob/mcaa014>
- Loukas A, Vasiliades L, Tzabiras J (2008) Climate Change Effects on Drought Severity. *Advances in Geosciences* 17:23–29. <https://doi.org/10.5194/adgeo-17-23-2008>
- Louwaars N P (2018) Plant Breeding and Diversity: A Troubled Relationship?. *Euphytica* 214:114. <https://doi.org/10.1007/s10681-018-2192-5>
- Ludlow M (1989) Strategies of Response to Water Stress. SPB Academic, The Hague
- Lynch J, Cain M, Frame D, Pierrehumbert R (2021) Agriculture's Contribution to Climate Change and Role in Mitigation Is Distinct From Predominantly Fossil CO<sub>2</sub>-Emitting Sectors. *Frontiers in Sustainable Food Systems* 4:518039. <https://doi.org/10.3389/fsufs.2020.518039>
- Lynch J P (2013) Steep, Cheap and Deep: An Ideotype to Optimize Water and N Acquisition by Maize Root Systems. *Annals of Botany* 112:347–357. <https://doi.org/10.1093/aob/mcs293>
- Lyu X, Mu R, Liu B (2023) Shade Avoidance Syndrome in Soybean and Ideotype toward Shade Tolerance. *Molecular Breeding* 43:31. <https://doi.org/10.1007/s11032-023-01375-3>
- López Pereira M, Sadras V O, Batista W, et al (2017) Light-Mediated Self-Organization of Sunflower Stands Increases Oil Yield in the Field. *Proceedings of the National Academy of Sciences* 114:7975–7980. <https://doi.org/10.1073/pnas.1618990114>
- MacArthur R H (1958) Population Ecology of Some Warblers of Northeastern Coniferous Forests. *Ecology* 39:599–619. <https://doi.org/10.2307/1931600>
- Maccaferri M, Harris N S, Twardziok S O, et al (2019) Durum Wheat Genome Highlights Past Domestication Signatures and Future Improvement Targets. *Nature Genetics* 51:885–895. <https://doi.org/10.1038/s41588-019-0381-3>
- Mahaut L, Violle C, Shihani A, et al (2023) Beyond Trait Distances: Functional Distinctiveness Captures the Outcome of Plant Competition. *Functional Ecology* 37:2399–2412. <https://doi.org/10.1111/1365-2435.14397>
- Maina G G, Brown J S, Gersani M (2002) Intra-Plant versus Inter-plant Root Competition in Beans: Avoidance, Resource Matching or Tragedy of the Commons. *Plant Ecology* 160:235–247
- Makino A (2021) Photosynthesis Improvement for Enhancing Productivity in Rice. *Soil Science and Plant Nutrition* 67:513–519. <https://doi.org/10.1080/00380768.2021.1966290>

- Mangin B, Casadebaig P, Cadic E, et al (2017) Genetic Control of Plasticity of Oil Yield for Combined Abiotic Stresses Using a Joint Approach of Crop Modelling and Genome-wide Association. *Plant, Cell & Environment* 40:2276–2291. <https://doi.org/10.1111/pce.12961>
- Maron J L, Marler M, Klironomos J N, Cleveland C C (2011) Soil Fungal Pathogens and the Relationship between Plant Diversity and Productivity. *Ecology Letters* 14:36–41. <https://doi.org/10.1111/j.1461-0248.2010.01547.x>
- McDonald B A, Allard R W, Webster R K (1988) Responses of Two-, Three-, and Four-Component Barley Mixtures to a Variable Pathogen Population. *Crop Science* 28:447–452. <https://doi.org/10.2135/cropsci1988.0011183X002800030003x>
- McDonald B A, Linde C (2002) Pathogen Population Genetics, Evolutionary Potential, and Durable Resistance. *Annual Review of Phytopathology* 40:349–379. <https://doi.org/10.1146/annurev.phyto.40.120501.101443>
- McKane R B, Johnson L C, Shaver G R, et al (2002) Resource-Based Niches Provide a Basis for Plant Species Diversity and Dominance in Arctic Tundra. *Nature* 415:68–71. <https://doi.org/10.1038/415068a>
- McNickle G G (2020) Interpreting Plant Root Responses to Nutrients, Neighbours and Pot Volume Depends on Researchers' Assumptions. *Functional Ecology* 34:2199–2209. <https://doi.org/10.1111/1365-2435.13517>
- Midgley G, Moll E (1993) Gas Exchange in Arid-Adapted Shrubs: When Is Efficient Water Use a Disadvantage?. *South African Journal of Botany* 59:491–495. [https://doi.org/10.1016/S0254-6299\(16\)30693-7](https://doi.org/10.1016/S0254-6299(16)30693-7)
- Mikaberidze A, McDonald B A, Bonhoeffer S (2015) Developing Smarter Host Mixtures to Control Plant Disease. *Plant Pathology* 64:996–1004. <https://doi.org/10.1111/ppa.12321>
- Mobley M L, Kruse A S, McNickle G G (2022) *Pisum Sativum* Has No Competitive Responses to Neighbors: A Case Study in (Non)Reproducible Plant Biology. *Plant Direct* 6:e411. <https://doi.org/10.1002/pld3.411>
- Montazeaud G, Flutre T, Ballini E, et al (2022) From Cultivar Mixtures to Allelic Mixtures: Opposite Effects of Allelic Richness between Genotypes and Genotype Richness in Wheat. *New Phytologist* 233:2573–2584. <https://doi.org/10.1111/nph.17915>
- Montazeaud G, Helleu Q, Wuest S E, Keller L (2023) Indirect Genetic Effects Are Shaped by Demographic History and Ecology in *Arabidopsis thaliana*. *Nature Ecology & Evolution* 7:1878–1891. <https://doi.org/10.1038/s41559-023-02189-4>
- Montazeaud G, Violle C, Fréville H, et al (2018) Crop Mixtures: Does Niche Complementarity Hold for Belowground Resources? An Experimental Test Using Rice Genotypic Pairs. *Plant and Soil* 424:187–202. <https://doi.org/10.1007/s11104-017-3496-2>
- Montazeaud G, Violle C, Roumet P, et al (2020) Multifaceted Functional Diversity for Multi-faceted Crop Yield: Towards Ecological Assembly Rules for Varietal Mixtures. *Journal of Applied Ecology* 57:2285–2295. <https://doi.org/10.1111/1365-2664.13735>
- Morin X, Fahse L, de Mazancourt C, et al (2014) Temporal Stability in Forest Productivity Increases with Tree Diversity Due to Asynchrony in Species Dynamics. *Ecology Letters* 17:1526–1535. <https://doi.org/10.1111/ele.12357>

- 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>
- Mundt C C (2014) Durable Resistance: A Key to Sustainable Management of Pathogens and Pests. Infection, Genetics and Evolution 27:446–455. <https://doi.org/10.1016/j.meegid.2014.01.011>
- Nakazawa M, Yabe N, Ichikawa T, et al (2001) *DFL1*, an Auxin-responsive *GH3* Gene Homologue, Negatively Regulates Shoot Cell Elongation and Lateral Root Formation, and Positively Regulates the Light Response of Hypocotyl Length. The Plant Journal 25:213–221. <https://doi.org/10.1111/j.1365-313X.2001.00957.x>
- Newton A, Begg G, Swanston J (2009) Deployment of Diversity for Enhanced Crop Function. Annals of Applied Biology 154:309–322. <https://doi.org/10.1111/j.1744-7348.2008.00303.x>
- Norrrington-Davies J (1967) Application of Diallel Analysis to Experiments in Plant Competition. Euphytica 16:391–406. <https://doi.org/10.1007/BF00028948>
- Ouédraogo D-Y, Mortier F, Gourlet-Fleury S, et al (2013) Slow-growing Species Cope Best with Drought: Evidence from Long-term Measurements in a Tropical Semi-deciduous Moist Forest of Central Africa. Journal of Ecology 101:1459–1470. <https://doi.org/10.1111/1365-2745.12165>
- Overton M (1996) Agricultural Revolution in England: The Transformation of the Agrarian Economy, 1500–1850. Cambridge University Press, Cambridge ; New York
- Palta J A, Turner N C (2019) Crop Root System Traits Cannot Be Seen as a Silver Bullet Delivering Drought Resistance. Plant and Soil 439:31–43. <https://doi.org/10.1007/s11104-018-3864-6>
- Pantazopoulou C K, Bongers F J, Pierik R (2021) Reducing Shade Avoidance Can Improve *Arabidopsis* Canopy Performance against Competitors. Plant, Cell & Environment 44:1130–1141. <https://doi.org/10.1111/pce.13905>
- Park J-E, Park J-Y, Kim Y-S, et al (2007) GH3-mediated Auxin Homeostasis Links Growth Regulation with Stress Adaptation Response in *Arabidopsis*. Journal of Biological Chemistry 282:10036–10046. <https://doi.org/10.1074/jbc.M610524200>
- Pendergrass A G, Knutti R, Lehner F, et al (2017) Precipitation Variability Increases in a Warmer Climate. Scientific Reports 7:17966. <https://doi.org/10.1038/s41598-017-17966-y>
- Peoples M B, Brockwell J, Herridge D F, et al (2009) The Contributions of Nitrogen-Fixing Crop Legumes to the Productivity of Agricultural Systems. Symbiosis 48:1–17. <https://doi.org/10.1007/BF03179980>
- Pianka E R (1973) The Structure of Lizard Communities. Annual Review of Ecology, Evolution, and Systematics 4:53–74. <https://doi.org/10.1146/annurev.es.04.110173.000413>
- Pianka E R (1974) Niche Overlap and Diffuse Competition. Proceedings of the National Academy of Sciences 71:2141–2145. <https://doi.org/10.1073/pnas.71.5.2141>
- Pierik R, De Wit M (2014) Shade Avoidance: Phytochrome Signalling and Other Aboveground Neighbour Detection Cues. Journal of Experimental Botany 65:2815–2824. <https://doi.org/10.1093/jxb/ert389>

- Plas F van der, Schröder-Georgi T, Weigelt A, et al (2020) Plant Traits Alone Are Poor Predictors of Ecosystem Properties and Long-Term Ecosystem Functioning. *Nature Ecology & Evolution* 4:1602–1611. <https://doi.org/10.1038/s41559-020-01316-9>
- Poorter H, Sack L (2012) Pitfalls and Possibilities in the Analysis of Biomass Allocation Patterns in Plants. *Frontiers in Plant Science* 3:. <https://doi.org/10.3389/fpls.2012.00259>
- Porter W P, Kearney M (2009) Size, Shape, and the Thermal Niche of Endotherms. *Proceedings of the National Academy of Sciences* 106:19666–19672. <https://doi.org/10.1073/pnas.0907321106>
- Postiglione A E, Muday G K (2023) Abscisic Acid Increases Hydrogen Peroxide in Mitochondria to Facilitate Stomatal Closure. *Plant Physiology* 192:469–487. <https://doi.org/10.1093/plphys/kiac601>
- Proulx R, Wirth C, Voigt W, et al (2010) Diversity Promotes Temporal Stability across Levels of Ecosystem Organization in Experimental Grasslands. *PLOS ONE* 5:e13382. <https://doi.org/10.1371/journal.pone.0013382>
- Pélissier R, Buendia L, Brousse A, et al (2021) Plant Neighbour-Modulated Susceptibility to Pathogens in Intraspecific Mixtures. *Journal of Experimental Botany* 72:6570–6580. <https://doi.org/10.1093/jxb/erab277>
- Qaim M (2016) Plant Breeding and Agricultural Development. *Genetically Modified Crops and Agricultural Development* 15–38
- R Core Team. (2021) R: A Language and Environment for Statistical Computing
- Rankin D J, Bargum K, Kokko H (2007) The Tragedy of the Commons in Evolutionary Biology. *Trends in Ecology & Evolution* 22:643–651. <https://doi.org/10.1016/j.tree.2007.07.009>
- Reif J C, Zhang P, Dreisigacker S, et al (2005) Wheat Genetic Diversity Trends during Domestication and Breeding. *Theoretical and Applied Genetics* 110:859–864. <https://doi.org/10.1007/s00122-004-1881-8>
- Reiss E R, Drinkwater L E (2018) Cultivar Mixtures: A Meta-analysis of the Effect of Intraspecific Diversity on Crop Yield. *Ecological Applications* 28:62–77. <https://doi.org/10.1002/eap.1629>
- Renard D, Tilman D (2019) National Food Production Stabilized by Crop Diversity. *Nature* 571:257–260. <https://doi.org/10.1038/s41586-019-1316-y>
- Rimbaud L, Fabre F, Papaïx J, et al (2021) Models of Plant Resistance Deployment. *Annual Review of Phytopathology* 59:125–152. <https://doi.org/10.1146/annurev-phyto-020620-122134>
- Rimbaud L, Papaïx J, Barrett L G, et al (2018) Mosaics, Mixtures, Rotations or Pyramiding: What Is the Optimal Strategy to Deploy Major Gene Resistance?. *Evolutionary Applications* 11:1791–1810. <https://doi.org/10.1111/eva.12681>
- Rinehart B, Borras L, Salmeron M, et al (2024) Commercial Maize Hybrids Have Smaller Root Systems after 80 Years of Breeding. *Rhizosphere* 30:100915. <https://doi.org/10.1016/j.rhisph.2024.100915>
- Roberts D R, Bahn V, Ciuti S, et al (2017) Cross-validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure. *Ecography* 40:913–929. <https://doi.org/10.1111/ecog.02881>

- Rodrigues J, Inzé D, Nelissen H, Saibo N J (2019) Source–Sink Regulation in Crops under Water Deficit. *Trends in Plant Science* 24:652–663. <https://doi.org/10.1016/j.tplants.2019.04.005>
- Roscher C, Schumacher J, Gubsch M, et al (2012) Using Plant Functional Traits to Explain Diversity–Productivity Relationships. *PLoS ONE* 7:e36760. <https://doi.org/10.1371/journal.pone.0036760>
- Roth L, Barendregt C, Bétrix C-A, et al (2022) High-Throughput Field Phenotyping of Soybean: Spotting an Ideotype. *Remote Sensing of Environment* 269:112797. <https://doi.org/10.1016/j.rse.2021.112797>
- Sadeghi B, Poom Chiarawongse, Squire K, et al (2022) DecisionTree.Jl - A Julia Implementation of the CART Decision Tree and Random Forest Algorithms
- Salinger M J (2005) Climate Variability and Change: Past, Present and Future – An Overview. *Climatic Change* 70:9–29. <https://doi.org/10.1007/s10584-005-5936-x>
- Salles J F, Poly F, Schmid B, Roux X L (2009) Community Niche Predicts the Functioning of Denitrifying Bacterial Assemblages. *Ecology* 90:3324–3332. <https://doi.org/10.1890/09-0188.1>
- Sambatti J B M, Caylor K K (2007) When Is Breeding for Drought Tolerance Optimal If Drought Is Random?. *New Phytologist* 175:70–80. <https://doi.org/10.1111/j.1469-8137.2007.02067.x>
- Sansaloni C, Franco J, Santos B, et al (2020) Diversity Analysis of 80,000 Wheat Accessions Reveals Consequences and Opportunities of Selection Footprints. *Nature Communications* 11:4572. <https://doi.org/10.1038/s41467-020-18404-w>
- Sato Y, Yamamoto E, Shimizu K K, Nagano A J (2021) Neighbor GWAS: Incorporating Neighbor Genotypic Identity into Genome-Wide Association Studies of Field Herbivory. *Heredity* 126:597–614. <https://doi.org/10.1038/s41437-020-00401-w>
- Schellenberger Costa D, Gerschlauer F, Pabst H, et al (2017) Community-Weighted Means and Functional Dispersion of Plant Functional Traits along Environmental Gradients on Mount Kilimanjaro. *Journal of Vegetation Science* 28:684–695. <https://doi.org/10.1111/jvs.12542>
- Schmid B, Hector A, Huston M A, et al (2002) The Design and Analysis of Biodiversity Experiments. *Biodiversity and Ecosystem Functioning* 61–76
- Schmid B (1994) Effects of Genetic Diversity in Experimental Stands of *Solidago Altissima* – Evidence for the Potential Role of Pathogens as Selective Agents in Plant Populations. *Journal of Ecology* 82:165–175. <https://doi.org/10.2307/2261395>
- Schmid C, Bauer S, Bartelheimer M (2015) Should I Stay or Should I Go? Roots Segregate in Response to Competition Intensity. *Plant and Soil* 391:283–291. <https://doi.org/10.1007/s11104-015-2419-3>
- Schmitt J, McCormac A C, Smith H (1995) A Test of the Adaptive Plasticity Hypothesis Using Transgenic and Mutant Plants Disabled in Phytochrome-Mediated Elongation Responses to Neighbors. *The American Naturalist* 146:937–953. <https://doi.org/10.1086/285832>
- Schnitzer S A, Klironomos J N, HilleRisLambers J, et al (2011) Soil Microbes Drive the Classic Plant Diversity–Productivity Pattern. *Ecology* 92:296–303. <https://doi.org/10.1890/10-0773.1>

- Schori A, Charles S, Peter D (2003) Soja: sélection, agronomie et production en Suisse. Revue suisse Agric 35:69–76
- Schwinning S, Weiner J (1998) Mechanisms Determining the Degree of Size Asymmetry in Competition among Plants. *Oecologia* 113:447–455. <https://doi.org/10.1007/s004420050397>
- Schöb C, Engbersen N, López-Angulo J, et al (2023) Crop Diversity Experiment: Towards a Mechanistic Understanding of the Benefits of Species Diversity in Annual Crop Systems. *Journal of Plant Ecology* 16:rtad16. <https://doi.org/10.1093/jpe/rtad016>
- Sedivy E J, Wu F, Hanzawa Y (2017) Soybean Domestication: The Origin, Genetic Architecture and Molecular Bases. *New Phytologist* 214:539–553. <https://doi.org/10.1111/nph.14418>
- Semba R D, Ramsing R, Rahman N, et al (2021) Legumes as a Sustainable Source of Protein in Human Diets. *Global Food Security* 28:100520. <https://doi.org/10.1016/j.gfs.2021.100520>
- Semchenko M, Hutchings M J, John E A (2007) Challenging the Tragedy of the Commons in Root Competition: Confounding Effects of Neighbour Presence and Substrate Volume. *Journal of Ecology* 95:252–260. <https://doi.org/10.1111/j.1365-2745.2007.01210.x>
- Shavrukov Y, Kurishbayev A, Jataev S, et al (2017) Early Flowering as a Drought Escape Mechanism in Plants: How Can It Aid Wheat Production?. *Frontiers in Plant Science* 8:. <https://doi.org/10.3389/fpls.2017.01950>
- Shipley B, De Bello F, Cornelissen J H C, et al (2016) Reinforcing Loose Foundation Stones in Trait-Based Plant Ecology. *Oecologia* 180:923–931. <https://doi.org/10.1007/s00442-016-3549-x>
- Shmueli G (2010) To Explain or to Predict?. *Statistical Science* 25:. <https://doi.org/10.1214/10-STS330>
- Silvertown J, Dodd M E, Gowing D J G, Mountford J O (1999) Hydrologically Defined Niches Reveal a Basis for Species Richness in Plant Communities. *Nature* 400:61–63. <https://doi.org/10.1038/21877>
- Skirycz A (2011) Survival and Growth of *Arabidopsis* Plants given Limited Water Are Not Equal. *Nature Biotechnology* 29:. <https://doi.org/10.1038/nbt.1800>
- Smithson J B, Lenné J M (1996) Varietal Mixtures: A Viable Strategy for Sustainable Productivity in Subsistence Agriculture. *Annals of Applied Biology* 128:127–158. <https://doi.org/10.1111/j.1744-7348.1996.tb07096.x>
- Spehn E M, Scherer-Lorenzen M, Schmid B, et al (2002) The Role of Legumes as a Component of Biodiversity in a Cross-European Study of Grassland Biomass Nitrogen. *Oikos* 98:205–218. <https://doi.org/10.1034/j.1600-0706.2002.980203.x>
- Suhre J J, Weidenbenner N H, Rowntree S C, et al (2014) Soybean Yield Partitioning Changes Revealed by Genetic Gain and Seeding Rate Interactions. *Agronomy Journal* 106:1631–1642. <https://doi.org/10.2134/agronj14.0003>
- Technow F, Schrag T A, Schipprack W, et al (2014) Genome Properties and Prospects of Genomic Prediction of Hybrid Performance in a Breeding Program of Maize. *Genetics* 197:1343–1355. <https://doi.org/10.1534/genetics.114.165860>
- Tester M, Langridge P (2010) Breeding Technologies to Increase Crop Production in a Changing World. *Science* 327:818–822. <https://doi.org/10.1126/science.1183700>

- Tian J, Wang C, Chen F, et al (2024) Maize Smart-Canopy Architecture Enhances Yield at High Densities. *Nature* 632:576–584. <https://doi.org/10.1038/s41586-024-07669-6>
- Tilman D, Reich P B, Isbell F (2012) Biodiversity Impacts Ecosystem Productivity as Much as Resources, Disturbance, or Herbivory. *Proceedings of the National Academy of Sciences* 109:10394–10397. <https://doi.org/10.1073/pnas.1208240109>
- Tooker J F, Frank S D (2012) Genotypically Diverse Cultivar Mixtures for Insect Pest Management and Increased Crop Yields. *Journal of Applied Ecology* 49:974–985. <https://doi.org/10.1111/j.1365-2664.2012.02173.x>
- Turnbull L A, Isbell F, Purves D W, et al (2016) Understanding the Value of Plant Diversity for Ecosystem Functioning through Niche Theory. *Proceedings of the Royal Society B: Biological Sciences* 283:20160536. <https://doi.org/10.1098/rspb.2016.0536>
- United Nations (2024) World Population Prospects 2024: Ten Key Messages
- Valladares F, Matesanz S, Guilhaumon F, et al (2014) The Effects of Phenotypic Plasticity and Local Adaptation on Forecasts of Species Range Shifts under Climate Change. *Ecology Letters* 17:1351–1364. <https://doi.org/10.1111/ele.12348>
- Vestergaard N F, Jørgensen L N (2024) Variety Mixtures of Winter Wheat: A General Status and National Case Study. *Journal of Plant Diseases and Protection* 131:1127–1136. <https://doi.org/10.1007/s41348-023-00856-z>
- Violle C, Jiang L (2009) Towards a Trait-Based Quantification of Species Niche. *Journal of Plant Ecology* 2:87–93. <https://doi.org/10.1093/jpe/rtp007>
- Violle C, Navas M-L, Vile D, et al (2007) Let the Concept of Trait Be Functional!. *Oikos* 116:882–892. <https://doi.org/10.1111/j.0030-1299.2007.15559.x>
- von Felten S, Niklaus P A, Scherer-Lorenzen M, et al (2012) Do Grassland Plant Communities Profit from N Partitioning by Soil Depth?. *Ecology* 93:2386–2396. <https://doi.org/10.1890/11-1439.1>
- Voss-Fels K P, Stahl A, Wittkop B, et al (2019) Breeding Improves Wheat Productivity under Contrasting Agrochemical Input Levels. *Nature Plants* 5:706–714. <https://doi.org/10.1038/s41477-019-0445-5>
- The VSNI Team (2023) Asreml: Fits Linear Mixed Models Using REML
- Wagg C, Ebeling A, Roscher C, et al (2017) Functional Trait Dissimilarity Drives Both Species Complementarity and Competitive Disparity. *Functional Ecology* 31:2320–2329. <https://doi.org/10.1111/1365-2435.12945>
- Wang S, Callaway R M (2021) Plasticity in Response to Plant–Plant Interactions and Water Availability. *Ecology* 102:e3361. <https://doi.org/10.1002/ecy.3361>
- Weiner J (2003) Ecology – the Science of Agriculture in the 21st Century. *The Journal of Agricultural Science* 141:371–377. <https://doi.org/10.1017/S0021859603003605>
- Weiner J (2019) Looking in the Wrong Direction for Higher-Yielding Crop Genotypes. *Trends in Plant Science* 24:927–933. <https://doi.org/10.1016/j.tplants.2019.07.001>
- Weiner J (2017) Applying Plant Ecological Knowledge to Increase Agricultural Sustainability. *Journal of Ecology* 105:865–870. <https://doi.org/10.1111/1365-2745.12792>

- Weiner J (1990) Asymmetric Competition in Plant Populations. *Trends in Ecology & Evolution* 5:360–364. [https://doi.org/10.1016/0169-5347\(90\)90095-U](https://doi.org/10.1016/0169-5347(90)90095-U)
- Weiner J, Andersen S B, Wille W K-M, et al (2010) Evolutionary Agroecology: The Potential for Cooperative, High Density, Weed-Suppressing Cereals. *Evolutionary Applications* 3:473–479. <https://doi.org/10.1111/j.1752-4571.2010.00144.x>
- Weiner J, Du Y-L, Zhang C, et al (2017) Evolutionary Agroecology: Individual Fitness and Population Yield in Wheat (*Triticum aestivum*). *Ecology* 98:2261–2266. <https://doi.org/10.1002/ecy.1934>
- Wille W, Pipper C B, Rosenqvist E, et al (2017) Reducing Shade Avoidance Responses in a Cereal Crop. *AoB PLANTS* 9:. <https://doi.org/10.1093/aobpla/plx039>
- 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. <https://doi.org/10.1146/annurev.py.23.090185.001343>
- Wolfe M S (1980) Can We Lead the Pathogen. *Plant Disease* 64:148. <https://doi.org/10.1094/PD-64-148>
- Wright A J, Wardle D A, Callaway R, Gaxiola A (2017) The Overlooked Role of Facilitation in Biodiversity Experiments. *Trends in Ecology & Evolution* 32:383–390. <https://doi.org/10.1016/j.tree.2017.02.011>
- Wright I J, Reich P B, Westoby M, et al (2004) The Worldwide Leaf Economics Spectrum. *Nature* 428:821–827. <https://doi.org/10.1038/nature02403>
- Wu C-H, Derevnina L, Kamoun S (2018) Receptor Networks Underpin Plant Immunity. *Science* 360:1300–1301. <https://doi.org/10.1126/science.aat2623>
- Wuest S E, Niklaus P A (2018) A Plant Biodiversity Effect Resolved to a Single Chromosomal Region. *Nature Ecology & Evolution* 2:1933–1939. <https://doi.org/10.1038/s41559-018-0708-y>
- Wuest S E, Peter R, Niklaus P A (2021) Ecological and Evolutionary Approaches to Improving Crop Variety Mixtures. *Nature Ecology & Evolution* 5:1068–1077. <https://doi.org/10.1038/s41559-021-01497-x>
- Wuest S E, Pires N D, Luo S, et al (2022) Increasing Plant Group Productivity through Latent Genetic Variation for Cooperation. *PLOS Biology* 20:e3001842. <https://doi.org/10.1371/journal.pbio.3001842>
- Wuest S E, Schulz L, Rana S, et al (2023) Single-Gene Resolution of Diversity-Driven Overyielding in Plant Genotype Mixtures. *Nature Communications* 14:3379. <https://doi.org/10.1038/s41467-023-39130-z>
- Xi Y, Wang D, Weiner J, et al (2023) Time to Onset of Flowering, Water Use, and Yield in Wheat. *Agronomy* 13:1217. <https://doi.org/10.3390/agronomy13051217>
- Xiong W, Reynolds M P, Crossa J, et al (2021) Increased Ranking Change in Wheat Breeding under Climate Change. *Nature Plants* 7:1207–1212. <https://doi.org/10.1038/s41477-021-00988-w>
- Xu L, Yuan S, Man J (2020) Changes in Rice Yield and Yield Stability in China during the Past Six Decades. *Journal of the Science of Food and Agriculture* 100:3560–3569. <https://doi.org/10.1002/jsfa.10385>

- Yang L-N, Pan Z-C, Zhu W, et al (2019) Enhanced Agricultural Sustainability through Within-Species Diversification. *Nature Sustainability* 2:46–52. <https://doi.org/10.1038/s41893-018-0201-2>
- Yang M-H, Jahufer M Z, He J, et al (2020) Effect of Traditional Soybean Breeding on Water Use Strategy in Arid and Semi-Arid Areas. *European Journal of Agronomy* 120:126128. <https://doi.org/10.1016/j.eja.2020.126128>
- Yang Y, Tilman D, Lehman C, Trost J J (2018) Sustainable Intensification of High-Diversity Biomass Production for Optimal Biofuel Benefits. *Nature Sustainability* 1:686–692. <https://doi.org/10.1038/s41893-018-0166-1>
- Yates L A, Aandahl Z, Richards S A, Brook B W (2023) Cross Validation for Model Selection: A Review with Examples from Ecology. *Ecological Monographs* 93:e1557. <https://doi.org/10.1002/ecm.1557>
- Zea-Cabrera E, Iwasa Y, Levin S, Rodríguez-Iturbe I (2006) Tragedy of the Commons in Plant Water Use. *Water Resources Research* 42:. <https://doi.org/10.1029/2005WR004514>
- Zenes N, Kerr K L, Trugman A T, Anderegg W R L (2020) Competition and Drought Alter Optimal Stomatal Strategy in Tree Seedlings. *Frontiers in Plant Science* 11:478. <https://doi.org/10.3389/fpls.2020.00478>
- Zhai Y, Lv Y, Li X, et al (2014) A Synthetic Framework for Modeling the Genetic Basis of Phenotypic Plasticity and Its Costs. *New Phytologist* 201:357–365. <https://doi.org/10.1111/nph.12458>
- Zhang D-Y, Sun G-J, Jiang X-H (1999) Donald's Ideotype and Growth Redundancy: A Game Theoretical Analysis. *Field Crops Research* 61:179–187. [https://doi.org/10.1016/S0378-4290\(98\)00156-7](https://doi.org/10.1016/S0378-4290(98)00156-7)
- Zhao Y, Li Z, Liu G, et al (2015) Genome-Based Establishment of a High-Yielding Heterotic Pattern for Hybrid Wheat Breeding. *Proceedings of the National Academy of Sciences* 112:15624–15629. <https://doi.org/10.1073/pnas.1514547112>
- Zhou Y, Kusmec A, Schnable P S (2024) Genetic Regulation of Self-Organizing Azimuthal Canopy Orientations and Their Impacts on Light Interception in Maize. *The Plant Cell* 36:1600–1621. <https://doi.org/10.1093/plcell/koae007>
- Zhu Y-H, Weiner J, Li F-M (2019b) Root Proliferation in Response to Neighbouring Roots in Wheat (*Triticum aestivum*). *Basic and Applied Ecology* 39:10–14. <https://doi.org/10.1016/j.baae.2019.07.001>
- Zhu Y-H, Weiner J, Yu M-X, Li F-M (2019a) Evolutionary Agroecology: Trends in Root Architecture during Wheat Breeding. *Evolutionary Applications* 12:733–743. <https://doi.org/10.1111/eva.12749>
- Zhu Y, Chen H, Fan J, et al (2000) Genetic Diversity and Disease Control in Rice. *Nature* 406:718–722. <https://doi.org/10.1038/35021046>
- Zimmer S, Messmer M, Haase T, et al (2016) Effects of Soybean Variety and Bradyrhizobium Strains on Yield, Protein Content and Biological Nitrogen Fixation under Cool Growing Conditions in Germany. *European Journal of Agronomy* 72:38–46. <https://doi.org/10.1016/j.eja.2015.09.008>