Both selection and plasticity drive niche differentiation in experimental grasslands

Julien Meilhac[1], Lucas Deschamps[2], Vincent Maire[2], Sandrine Flajoulot[3], and Isabelle Litrico[1*]

The way species avoid each other in a community by using resources differently across space and time is one of the main drivers of species coexistence in nature[1,2]. This mechanism, known as niche differentiation, has been widely examined theoretically but still lacks thorough experimental validation in plants. To shape niche differences over time, species within communities can reduce the overlap between their niches or find unexploited environmental space[3]. Selection and phenotypic plasticity have been advanced as two candidate processes driving niche differentiation[4,5], but their respective role remains to be quantified[6]. Here we tracked changes in plant height, as a candidate trait for light capture[7], in 5-year multispecies sown grasslands. We found increasing among-species height differences over time. Phenotypic plasticity promotes this change, which explains the rapid setting of differentiation in our system. Through the inspection of changes in genetic structure, we also highlighted the contribution of selection. Altogether, we experimentally demonstrated the occurrence of species niche differentiation within artificial grassland communities over a short time scale through the joined action of both plasticity and selection.

Niche differentiation (ND) occupies a central place in community and ecosystem ecology, explaining the maintenance of biodiversity[1,2,8] and its positive effects on ecosystem functioning, through complementary of resource use[9,10]. However, the mechanisms underlying ND dynamics remain unclear, particularly those explaining how each species shapes its niche through community assembling. A species niche can be described by traits that indicate how resources are strategically acquired from the environment[11-13]. Niche differentiation is then expected when the dispersion of trait values for each species do not overlap to avoid direct competition between similar plant strategies [14-16]. This occurs either, in the case of a fixed community trait range, by constraining the average dispersion of trait values expressed by each species[17], or by modifying the trait values outside the dispersion of community trait values to discover a previously unexploited environmental space[18]. As different species within a given community can follow both routes, the study of ND dynamics requires consideration of both the trajectory and the deformation of the niche, i.e. the temporal change in the position and the

[1]P3F UR 004 - INRA - Le Chêne RD150, F-86000 Lusignan, BP 86006, France. [2]Département des Sciences de l'Environnement - UQTR - QC G9A 5H7 Trois-Rivières, Canada. [3]Jouffray-Drillaud La Litière, Saint Sauvant 86600, France. *corresponding author isabelle.litrico-chiarelli@inra.fr

breadth of the species niche. To our knowledge, this has not previously been done experimentally. Niche differentiation dynamics may result first from the competitive exclusion of phenotypes located in the overlapping part of the niche. This is expected when plasticity is not sufficient to avoid species overlap[19]. Niche differentiation may also result from the selection of phenotypes that plasticity placed in unexploited environmental space[19]. Although genetic mechanisms play key roles in the process of ND, studies that focus on it are rare[20,21]. The time scale at which the process occurs remains particularly unclear, ranging from a single growing season and stretching out to evolutionary time. Time scale likely reflects a gradient between the ecological and the genetic shaping of the niche.

Our objective was to display ND dynamics by monitoring species trait variation over five years in multispecies sown grasslands. Using seven species commonly used in sown grasslands, five mixtures were assembled and sown in field trials as described in Meilhac, et al. [22]. The five mixtures (M-1 to M-5) varied only in within-species genetic diversity (i.e. the numbers of cultivars represented per species). There were three genetic structure levels: simple structure (M-1 to M-3), intermediate structure (M-4) and complex structure (M-5). Species were chosen to be of contrasting strategies in relation to light capture in order to promote a ND among vertical space between them: (i) a classical vertical-elongation strategy to maximise access to the above-canopy light (*Lotus corniculatus*, *Dactylis glomerata*, *Festuca arundinacea*, *Trifolium pratense* and *Medicago sativa*) and (ii) a horizontal-spreading strategy to maximise access to the light penetrating vegetation gaps (skylights)[23,24] (*Lolium perenne* and *Trifolium repens*).

To create a measurable selection opportunity, the within-species phenotypic variability of competitive traits was similar between mixtures (similar range trait variability) but the within-species genetic structure was dissimilar (see Table S1, Fig. S1 and S2 in supplementary information), this genetic diversity is not necessarily synonymous with phenotypic diversity. Using molecular tools, we followed the dynamics of cultivar abundance of each species during community assembling by a molecular method of cultivar fingerprinting (see Methods section). In this experiment, the effects of cultivar frequency change (selection estimation) on ND were tested. In parallel, the heights of each cultivar was characterised without competition in a common garden in order to relate cultivar selection to relate species strategy in the ND process (see Methods section).

Species distribution of vegetative height was monitored over time within each mixture to characterise the shaping of the species niche[25,26]. Vegetative height is a key plant feature widely used in the literature to characterise the light-acquisition strategy[7] to express the plant competitive ability over neighbours[27] and to explain a species ability to perform in a grassland community[7,28-31]. It is also associated with the competition – colonization trade-off for clonal plants opposing a strategy competing for vertical space to a strategy competing for horizontal space[32,33]. As observed in literature[34,35] and in our dataset (Fig. S5), plant height covaries negatively with tiller density and lateral spread at the interspecific level. Hence, we assume plant height is the key candidate trait to characterise a dimension of the species niche[36]. We modelled the temporal dynamics of height niches of all species using Bayesian Generalised Additive Distributional Models (GAM). We determined by model selection procedure based on the Watanabe-Akaike-Information Criterion (an unbiased measure of the log posterior predictive density of models, see *Methods*) if the niche dynamics differed for each species in terms of both mean trajectory and of niche deformation (dilatation or contraction).

From an overlapped distribution of species height for all mixtures studied (Fig. 1a - year 1), three distinct trajectories appeared over time (Fig. 1b), which were linked with known strategies of light capture (Fig. 1b), but were independent of mixture identity (Table 1, mod_{div} does not improve significantly the predictive power). Our results show that the partitioning of vegetative height distribution was strongly species-specific, the model describing a trajectory for each species ($Mod_{t,\mu}$) providing a far higher predictive power than Mod_t estimating a common trajectory for all species (Fig. 1b, Table 1, ΔWAIC Mod_t-$Mod_{t,\mu}$ = 1851±80 ; Table S6). *L. corniculatus* and *M. sativa* rose above the canopy between years one and five. Conversely, *T. repens* and *L. perenne* clearly adopted a different strategy, spreading at a low canopy level seeking for skylights. Meanwhile, *F. arundinacea* and *D. glomerata* remained at canopy level during the five-year experiment, exploiting the light by their position relative to the canopy (as having the highest biomass proportion in the community[22]). It is worth noting that *T. pratense* stayed at canopy level but was excluded by year five in line with its lower lifetime character.

Niche trajectories were tightly linked with niche deformation. The best model describing the distribution of plant height was the one where both mean and dispersion of each species depended upon a smooth time-function (Fig. 1, Table 1, ΔWAIC $Mod_{t,\mu}$-$Mod_{t,\mu,\sigma}$ = 230±33; Table S6). Species niche deformation through time clearly reduced the general overlap of height distribution and structured the exploitation of light through the canopy (Fig. S6). The species that climbed above the canopy (*L. corniculatus* and *M. sativa*), exhibited loss of small

97 individuals and a clear increase in dispersion of height values with time (Fig. 2a). This niche
98 dilatation showed a relaxation of competition intensity for the species able to dominate others
99 and bypass light reduction by other species. Conversely, the diversity of height values of the
100 species spreading at ground levels for skylights (*T. repens* and *L. perenne*), drastically
101 decreased with time, indicating a contraction of their niche for species known to be poor
102 competitors for light in the presence of other species[24]. Competitive pressures of species could
103 generate a displacement of niches either by plasticity or by selection in order to reduce light
104 competition by a displacement towards tallest values for species with vertical elongation
105 strategies and conversely towards small values for species with horizontal spreading strategies,
106 potentially linked to a resource allocation strategy[34]. For instance, *T. repense* favours stolon
107 extension when under light competition with grass species[37].

108 From both the trajectory and the deformation of species height niches, we conclude that ND
109 has occurred between species during the assemblage of our temporary grassland communities
110 (Fig 1). Niche contraction occurred for some species while niche dilation occurred for others
111 along a unique resource dimension of species niches. Interestingly, the dynamics of species
112 niche, paralleled their dynamics in abundance in the community (Fig. 3), suggesting the
113 important role of ND during the community assemblage, as observed in other studies[20]. Unlike
114 what is often expected when ND is increased between species, we did not observe a general
115 reduction of trait dispersion within species, explained by increases in total community variance
116 that strongly separate species niches without requiring their contraction. Moreover, the overlap
117 between species having similar (previously-known) light strategies was almost total.

118 Phenotypic plasticity (including ontogeny) and selection are major adaptive mechanisms that
119 contribute to niche shaping[38,39] but their identification and relative importance in ND is unclear.
120 Using a Bayesian GAM approach, we tested if the addition of cultivar relative frequencies
121 (proportion of each cultivar) into our previous best model describing the distribution of plant
122 height (Table 1 – $Mod_{t,\mu,\sigma}$; Table S6) improved the dynamics of species niche shaping.

123 Over the five-year period, cultivar dynamics within a grassland community contributed
124 significantly to the shaping of species niches. When changes in cultivar abundances were
125 included in the GAM model to modulate the mean and dispersion of plant height distribution,
126 the performance of the model was significantly improved (Table 1 ΔWAIC $Mod_{t,\mu,\sigma}$-$Mod_{t,prop}$,
127 = 82±21; Table S6). Interestingly, we observed the changes in cultivar frequencies and the
128 disappearance of height values of the initial niche of species were related to the light strategies

of the species. The cultivar frequencies showed significant evolution (Fig. 2) for most species over the five-year period, mainly due to differential mortality between cultivars and vegetative multiplication (sexual reproduction was limited in this experimental design – see the proportion of individuals no-assigned in Table S2). Species with a horizontal spreading strategies and showing niche contraction towards lower height values were characterised by decreases in the abundance of the highest stature cultivars (V3 for *T. repens* and V5 and V6 for *L. perenne*) and increases in the lowest stature cultivars (V2 for *T. repens* and V1-V2 for *L. perenne*). Conversely, *M. sativa*, which has an elongation strategy, showed niche dilatation towards greater height values and was characterised by the opposite response. Although contrasting trends were observed (V1 for *T. repens* was expected to increase and conversely V1 for *M. sativa* was expected to decrease but was not observed), they remained exceptions. Similar results were obtained with the mixture characterised by medium genetic-complexity (see Fig. S3 in Supplementary Information). This result suggests selection exists at cultivar level for most species and this contributed to ND, although no change was observed for *D. glomerata* or *T. pratense*. However, it could be that our selection estimation was not fine enough, as selection within a cultivar may well exist.

While temporal changes in niche position and breadth, i.e. ND, were mainly the result of phenotypic plasticity, we show here that selection also acted to shape the species niches along the light-capture dimension. Strong similarities in species trajectories between mixtures excluded genetic drift as driver of ND (see Figure S7 in supplementary information). Selection was observed over the five years of experimentation. This is a very short period in relation to evolutionary time and it would be interesting to study the impacts of this selection over longer time scales. As the niche of species is multidimensional, other dimensions of the niche also need to be investigated to better understand the cultivar selection of these species. For instance, plant precocity and root preference for nitrate vs ammonium as the nitrogen resource has previously been shown to structure grassland communities[35]. Although our study focused on ND at species level, similar processes could be observed at within-species level under the pressure competition between genotypes belong to the same species.

Originally, our results experimentally demonstrated that genetic selection is related with ND over a very short time scale. This offers a highly promising avenue for community ecology showing for the first time that ND is truly associated with evolutionary fitness. In addition, these results could be important in agro-ecology, especially to assemblages in communities in the context of species diversification of agro-systems. If the limiting resources and plant traits

linked to the capture of this resource are known, the species and genetic composition of sown grassland can be thought of as favouring niche differences between species. As ND should be linked with the production stability of ecosystems[22], having a good understanding of these processes is key to improving the management of these[40].

Methods

Experimental design. The experimental design consisted of five different multispecies-grassland seed mixtures, which were established in the field in September 2011. This design was a subset, randomly selected, of the one in Meilhac, et al. [22] to limit logistic and financial load associated with fingerprint. Seeds were sown in plots (5×1.3 m) in once time with two replicates (five seed mixtures x two replicates = 10 plots). The soil was a clay-limestone, located at the Jouffray Drillaud Station, Saint Sauvant, France (46° 21′ 37″ North, 0° 03′ 25″ East). Plots were exposed to the local temperate climate with an average annual rainfall of 730 mm. No irrigation or nitrogen were added during the five-year experimental period. The low weed biomass during experiment (see figure S4 in supplementary information) and species nature (Crepis sancta was a rosette plant form type and not in competition for light) no justified a particular management. A plot comprised eight 5-m-long rows, each containing the same seed mixture. Each mixture contained seven perennial species, all of them in common use in temporary grasslands and are known to naturally exhibit high level of phenotypic plasticity - *Lotus corniculatus*, *Dactylis glomerata*, *Festuca arundinacea*, *Trifolium pratense*, *Medicago sativa*, *Lolium perenne* and *Trifolium repens*. The five mixtures differed in within-species genetic diversity. This was achieved by varying the number of cultivars per species (see Table S1, supplementary information). The range of variability of phenotypic traits was similar between mixtures but the genetic structure was different. Three mixtures (M-1, M-2 and M-3) were of simple genetic structure within each species (just one cultivar per species), one mixture (M-4) had an intermediate genetic structure within each species (just two or three cultivars per species) and one mixture (M-5) had a more complex genetic structure within each species (up to six cultivars per species and containing all the cultivars used in this experiment). Each seed mixture was sown with the same total seed weight and species proportions in each plot (Table S1, supplementary information). Each mixture plot was replicated twice from the same seed lots and plots were distributed randomly within two blocks of five plots each. Plots were maintained over five years and harvested three times each year. Species were chosen to have contrasting response strategies to light competition and space occupation. The first group

contained species having an elongation strategy - *Lotus corniculatus*, *Dactylis glomerata*, *Festuca arundinacea*, *Trifolium pratense* and *Medicago sativa* – these all seek light by extending their leaves or stems upwards to reach the top of the canopy. The second group contained species having a spreading strategy i.e. they seek skylights in the community by spreading horizontally - *Trifolium repens* (stolons) and *Lolium perenne* (tillers). The vegetative height mean and variance of each cultivar was measured without competition in a common garden at the INRA station in Lusignan (near Saint Sauvant), France (see cultivar characterisation section).

Phenotypic diversity in the mixture. Vegetative height was measured *in situ* on 20 individuals per species in each plot and each block. To limit the measure individuals from the same seed, each plot was divided into four subplots. In each of its subplots, five most visible individuals and most separated were measured. These measurements were carried out several times – during the first, second, third and fifth years after sowing to quantify changes in phenotypic diversity over time.

Species biomass in a mixture. For the five years of the experiment (2012 to 2016) the whole canopy of each plot was harvested three times each year (spring, summer and autumn), limiting sexual reproduction and by consequence the recruitment from seeds. Vegetative multiplication was possible for some species, especially for *T. repens* and *F. arundinacea*. The annual species biomass data came from the study of Meilhac, et al. [22] where all details, including harvesting dates, methods and treatment procedure of samples, are described. All plots were cut at the same time at 5 cm above ground level. Harvest date was decided based on a visual assessment of the aboveground standing biomass. Each harvest from each plot was weighed fresh and a sample of each was dried to constant weight at 60°C for 72 h. At each harvest, four quadrats (0.33 x 0.15 m) were placed randomly in each plot and the species biomasses were separated. These samples were dried and weighed to measure the proportion each species in the total dry biomass. For each plot and each year, the annual total biomass (for all species) and annual species biomass (for each species) were calculated by summation of the three biomass measurements (spring, summer and autumn).

Cultivar characterisation. Simultaneously, each cultivar used in the mixtures was planted in an isolated-plants nursery at INRA (2014 September). The seed was taken from the same lot as used in the mixtures. Each cultivar was represented by 30 individuals and each individual was cloned to realise three genetically-identical replicates. A total of 32 cultivars of the seven

227 species was planted. Each individual was harvested three times each year (spring, summer and
228 autumn). Vegetative height was measured twice a year over three years for each plant to
229 characterise each cultivar (mean height and variance) under conditions of no competition.

230 **Selection estimation by cultivar fingerprinting.** Individuals within each species were
231 sampled three times during the experiment in the mixtures (M-4 and M-5) containing within-
232 species genetic diversity - at six months, and at three and five years after sowing. The number
233 of individuals sampled per species varied according to genetic structure (the number of cultivars
234 per species in the mixture at sowing). A total of 16 individuals were collected per cultivar used,
235 i.e. 32 individuals for the species with two cultivars per species and 96 individuals for species
236 with six cultivars per species (Table S2, supplementary information). Individuals were assigned
237 to cultivars by genome fingerprinting in two steps, with a first step constructing a reference
238 source and second step comparing it to individuals profile sampled *in situ*. This reference
239 database was constructed from 96 individuals per cultivar taken from the same seed batches of
240 cultivars as used in the plots. Seeds were sowed in germination plates and individuals were
241 sampled from first leaves. Extraction of DNA employed CTAB and chloroform purification
242 (CYMMIT, 2005). The quality and concentration of each DNA sample was checked by 1%
243 agarose gel electrophoresis. The choice of the type of markers has been dependent on the
244 species ploidy level, the available markers and the polymorphism of markers. SSR markers
245 were used for *Lolium perenne* (diploid) and we developed and used AFLP markers for the other
246 species (tetraploid). For SSR markers, DNA samples were amplified by polymerase chain
247 reaction (PCR) for the six loci used to discriminate the six *Lolium perenne* cultivars used in the
248 mixture (Table S3, supplementary information). PCR was carried out in a final reaction volume
249 of 10 µl containing 1X polymerase buffer, 0.325 U of MP Biomedicals polymerase, 0.2 mM of
250 dNTP (Invitro-gen), 0.1 µM of forward primer with a M13 tail, 0.2 µM of reverse primer, 0.1
251 µM of M13 tail IRD700 or IRD800 labelled primer and 20 ng of DNA. The PCR reactions were
252 carried out in a DNA Engine Tetrad2 thermocycler (Biorad). A denaturation period of 4 min at
253 94°C was followed by 35 cycles of 30 s at 94°C, 1 min at Tm (65°C – 1°C/cycle) and 1 min at
254 72°C and then 10 min at 72°C for final extension. For AFLP markers, the protocol described
255 by Vos, et al. [41] was used. The selective amplification was carried out on the basis of the primer
256 pair generating maximum polymorphism between cultivars. The number of specific primer
257 pairs used varied between species from 2 to 10 pairs (Table S4, supplementary information). A
258 Li-Cor IR2 (Li-Cor Inc) sequencer was used to separate the labelled, amplified DNA fragments
259 on a 6.5% acrylamide gel. Marker segregation was scored using SAGA Generation 2 software

(Li-Cor Inc) by two different persons and the results were compared. The number of bands scored to provide a cultivar assignment with an error rate of less than 5%, varied with species from 78 to 380 scored bands (Table S4, supplementary information). The same protocol (SSR or AFLP according to species) was carried out on individuals sampled *in situ* (multi-cultivar mixture). Assignation of individuals sampled *in situ* was done by comparison of their genetic profile to it of cultivars from database source. For *Lolium perenne,* individuals sampled *in situ* were attributed to cultivars with GeneClass2 software[42] from reference source and according to Bayesian method of Rannala and Mountain [43] and with the Distance method developed by Nei, et al. [44]. For other species, to analyse AFLP data, we used Structure software[45], based on methods for ambiguous genotype data such as dominant markers[46] (Version 2.3.4) with a systematic Bayesian clustering approach applying Markov Chain Monte Carlo estimation. According to species, we used admixture model with the length of burn-in period was 2.10^4 iterations and the number of MCMC after burn-in was 5.10^4 iterations for all species except *Festuca arundinacea* (10^4 iterations) and *Medicago sativa* (2.10^4 iterations). In order to assign individuals to cultivars, the means of the permuted matrices across replicates were computed using CLUMPP software[47]. From these assignment, the cultivar proportion was then calculated for each species in each block for the medium and high-diversity mixture (M-4 and M-5) and excluding ambiguous genotypes (i.e. genotypes that not clearly assigned to a cultivar - Table S2 and S5) to the calculations.

Statistical analyses. Species trait distributions were modelled explicitly using a distributional modelling framework (similar to the GAMLSS approach). This allowed modelling of each parameter of a parametric distribution as the result of an equation containing hierarchical parameters and smoothing functions. Because the complexity of such models can hardly be captured by classical maximum-likelihood methods, parameters were estimated in a Bayesian framework using Hamiltonian Monte-Carlo, as implemented in the *No-U-Turn* sampler (NUTS) of the *Stan* software. We described the variation of species trait distribution with models of increasing complexity, including time and genetic selection effects successively. We compared models using Watanabe-Akaike-Information-Criterion (WAIC) (see below) to identify the niches characteristics and the covariables that had an effect of them.

Species trait distributions were described using gamma distributions parameterised in terms of two independent parameters describing mean (μ) and dispersion (φ). Gamma distributions have the advantage of being able to describe a skewed distribution, which often arise in the case of

strictly positive random variables close to 0. The gamma distribution from which the height value y of individual i is sampled followed the general formulation

$$y_i \sim Gamma(\mu_i^2 \phi_i, \mu_i \phi_i)$$
$$log(\mu_i) = \beta_f X_i + \beta_r X_i + f_1(t_i)$$
$$log(\phi_i) = \gamma_f X_i + \gamma_r X_i + f_2(t_i)$$

With X_i being a design matrix of covariates describing the environment of observation i, while β_f and γ_f are fixed parameters, such as intercepts for block identities (all models) for mean (β_f) and dispersion (γ_f). Similarly, β_r and γ_r represent hierarchical population parameters described by a normal distribution with estimated standard deviation, for mean (β_r) and dispersion (γ_r,). In our case, the hierarchical parameters describe varying intercepts of mixture (all models), species (all models except Mod_{div}) or species at different levels of genetic diversity (Mod_{div}), but also varying slopes between cultivar proportion in each mixture and height distribution ($Mod_{t,prop'}$). f_{1s} and f_{2s}, describe cubic splines linking years to mean and dispersion of each species trait distribution. Mod_t estimates a common spline for all species for both mean and dispersion, while $Mod_{t,\mu}$ estimates a spline per species for mean only. $Mod_{t,\mu,\sigma}$, $Mod_{t,\mu,\sigma'}$ and $Mod_{t,prop'}$ estimate a spline per species for both mean and dispersion. Mod_{div} estimates a spline per species for each level of within species genetic diversity (one cultivar per species, up to three cultivars per species or up to six cultivars per species). Data of year 2 was excluded for models $Mod_{0'}$, $Mod_{t,\mu,\sigma'}$ and $Mod_{t,prop'}$ because no cultivar proportion was estimated this year. We used regularising priors to optimise estimation and limit overfitting, based on student-t distributions with three degrees of freedom and adjusted mean and scale.

The posterior of each model was sampled using the NUTS algorithm through the *brms* R package (version 2.2.3), which allows easily computed predictions and provides links to the model comparison package *loo* (version 2.0.0). The Stan Bayesian software provides diagnostics for sampling abnormal behaviour and chain mixing. We took care to avoid divergent transitions. We ensured convergence, by verifying that the scale reduction factor of each parameter (indicating chain convergence), did not exceed 1.1. Posterior predictive checks were carried out to ensure the model captured the key features of the data satisfactorily.

Model comparisons were carried out using Watanabe-Akaike-Information-Criterion (WAIC) implemented in the *loo* package. This criterion can be interpreted as the Akaike-Information-Criterion, with the best models are those with the lowest criterion. Both criteria provide an

unbiased measure of the expected log posterior predictive density. However, WAIC estimate the effective number of parameters estimated during model fitting, which makes its use accurate for hierarchical models. It also incorporates uncertainty arising from parameters estimation, and is thus described by an approximately normal distribution with estimated standard-error[48]. We interpreted the magnitude of the difference of the WAIC of different models (ΔWAIC), but also the uncertainty about this difference. During comparison of nested models of increasing complexity, a Δ_{WAIC} that was lower than twice its standard-error (approximately equivalent to the 95% confidence interval) was considered to provide little support for the more complex model. All computations were carried out using *R* version 3.4.4.

References

1 Levine, J. M. & HilleRisLambers, J. The importance of niches for the maintenance of species diversity. *Nature* **461**, 254-U130, doi:10.1038/nature08251 (2009).

2 Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31**, 343-366, doi:10.1146/annurev.ecolsys.31.1.343 (2000).

3 Von Felten, S. *et al.* Belowground nitrogen partitioning in experimental grassland plant communities of varying species richness. *Ecology* **90**, 1389-1399, doi:10.1890/08-0802.1 (2009).

4 Vellend, M. The consequences of genetic diversity in competitive communities. *Ecology* **87**, 304-311, doi:10.1890/05-0173 (2006).

5 Roscher, C., Schumacher, J., Schmid, B. & Schulze, E. D. Contrasting Effects of Intraspecific Trait Variation on Trait-Based Niches and Performance of Legumes in Plant Mixtures. *PLoS One* **10**, 18, doi:10.1371/journal.pone.0119786 (2015).

6 Niklaus, P. A., Baruffol, M., He, J. S., Ma, K. P. & Schmid, B. Can niche plasticity promote biodiversity-productivity relationships through increased complementarity? *Ecology* **98**, 1104-1116 (2017).

7 Westoby, M. A leaf-height-seed (LHS) plant ecology strategy scheme. *Plant Soil* **199**, 213-227, doi:10.1023/a:1004327224729 (1998).

8 Turnbull, L. A., Levine, J. M., Loreau, M. & Hector, A. Coexistence, niches and biodiversity effects on ecosystem functioning. *Ecology Letters* **16**, 116-127, doi:10.1111/ele.12056 (2013).

9 Cardinale, B. J. *et al.* Impacts of plant diversity on biomass production increase through time because of species complementarity. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 18123-18128, doi:10.1073/pnas.0709069104 (2007).

10 Tilman, D., Isbell, F. & Cowles, J. M. in *Annual Review of Ecology, Evolution, and Systematics, Vol 45* Vol. 45 *Annual Review of Ecology Evolution and Systematics* (ed D. J. Futuyma) 471-493 (Annual Reviews, 2014).

11 Navas, M. L. & Violle, C. Plant traits related to competition: how do they shape the functional diversity of communities? *Community Ecol.* **10**, 131-137, doi:10.1556/ComEc.10.2009.1.15 (2009).

12 D'Andrea, R. & Ostling, A. Challenges in linking trait patterns to niche differentiation. *Oikos* **125**, 1369-1385, doi:10.1111/oik.02979 (2016).

13 Roughgarden, J. Resource partitioning among competing species—A coevolutionary approach. *Theor. Popul. Biol.* **9**, 388-424, doi:10.1016/0040-5809(76)90054-X (1976).

14 Loreau, M. & de Mazancourt, C. Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecology Letters* **16**, 106-115, doi:10.1111/ele.12073 (2013).

15 Hart, S. P., Schreiber, S. J. & Levine, J. M. How variation between individuals affects species coexistence. *Ecology Letters* **19**, 825-838, doi:10.1111/ele.12618 (2016).

16 Schoener, T. W. Resource Partitioning in Ecological Communities. *Science* **185**, 27-39, doi:10.1126/science.185.4145.27 (1974).

17 Scheele, B. C., Foster, C. N., Banks, S. C. & Lindenmayer, D. B. Niche Contractions in Declining Species: Mechanisms and Consequences. *Trends in Ecology & Evolution* **32**, 346-355, doi:10.1016/j.tree.2017.02.013 (2017).

18 Berg, M. P. & Ellers, J. Trait plasticity in species interactions: a driving force of community dynamics. *Evol. Ecol.* **24**, 617-629, doi:10.1007/s10682-009-9347-8 (2010).

19 Grenier, S., Barre, P. & Litrico, I. Phenotypic Plasticity and Selection: Nonexclusive Mechanisms of Adaptation. *Scientifica*, 9, doi:10.1155/2016/7021701 (2016).

20 Zuppinger-Dingley, D. *et al.* Selection for niche differentiation in plant communities increases biodiversity effects. *Nature* **515**, 108-+, doi:10.1038/nature13869 (2014).

21 van Moorsel, S. J. *et al.* Community evolution increases plant productivity at low diversity. *Ecology Letters* **21**, 128-137, doi:10.1111/ele.12879 (2018).

22 Meilhac, J., Durand, J. L., Beguier, V. & Litrico, I. Increasing the benefits of species diversity in multispecies temporary grasslands by increasing within-species diversity. *Ann. Bot.* **[e-first]**, 1-10 (2019).

23 Gautier, H., Varlet-Grancher, C. & Baudry, N. Comparison of horizontal spread of white clover (Trifolium repens L.) grown under two artificial light sources differing in their content of blue light. *Ann. Bot.* **82**, 41-48, doi:10.1006/anbo.1998.0643 (1998).

24 Roscher, C., Kutsch, W. L. & Schulze, E. D. Light and nitrogen competition limit Lolium perenne in experimental grasslands of increasing plant diversity. *Plant Biol.* **13**, 134-144, doi:10.1111/j.1438-8677.2010.00338.x (2011).

25 McGill, B. J., Enquist, B. J., Weiher, E. & Westoby, M. Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution* **21**, 178-185, doi:10.1016/j.tree.2006.02.002 (2006).

26 Violle, C. & Jiang, L. Towards a trait-based quantification of species niche. *J. Plant Ecol.* **2**, 87-93, doi:10.1093/jpe/rtp007 (2009).

27 Gommers, C. M. M., Visser, E. J. W., St Onge, K. R., Voesenek, L. & Pierik, R. Shade tolerance: when growing tall is not an option. *Trends in Plant Science* **18**, 65-71, doi:10.1016/j.tplants.2012.09.008 (2013).

28 Pontes, L. D., Maire, V., Louault, F., Soussana, J. F. & Carrere, P. Impacts of species interactions on grass community productivity under contrasting management regimes. *Oecologia* **168**, 761-771, doi:10.1007/s00442-011-2129-3 (2012).

29 Schippers, P. & Kropff, M. J. Competition for light and nitrogen among grassland species: a simulation analysis. *Funct. Ecol.* **15**, 155-164 (2001).

30 Weiner, J. & Thomas, S. C. Size variability and competition in plant monocultures. *Oikos* **47**, 211-222, doi:10.2307/3566048 (1986).

31 Falster, D. S. & Westoby, M. Plant height and evolutionary games. *Trends in Ecology & Evolution* **18**, 337-343, doi:10.1016/s0169-5347(03)00061-2 (2003).

32 Bittebiere, A. K., Saiz, H. & Mony, C. New insights from multidimensional trait space responses to competition in two clonal plant species. *Funct. Ecol.* **33**, 297-307, doi:10.1111/1365-2435.13220 (2019).

33 Maire, V. *et al.* Disentangling Coordination among Functional Traits Using an Individual-Centred Model: Impact on Plant Performance at Intra- and Inter-Specific Levels. *PLoS One* **8**, e77372, doi:10.1371/journal.pone.0077372 (2013).

34 Nelson, C. J. in *Grassland ecophysiology and grazing ecology.* (ed University Press Cambridge) Ch. 6, 101 (2000).

35 Maire, V. *et al.* Habitat filtering and niche differentiation jointly explain species relative abundance within grassland communities along fertility and disturbance gradients. *New Phytol.* **196**, 497-509, doi:10.1111/j.1469-8137.2012.04287.x (2012).

36 Adler, P. B., Fajardo, A., Kleinhesselink, A. R. & Kraft, N. J. B. Trait-based tests of coexistence mechanisms. *Ecology Letters* **16**, 1294-1306, doi:10.1111/ele.12157 (2013).

37 Marriott, C. A., Bolton, G. R. & Duff, E. I. Factors affecting the stolon growth of white clover in ryegrass/clover patches. *Grass Forage Sci.* **52**, 147-155, doi:10.1111/j.1365-2494.1997.tb02345.x (1997).

38 Turcotte, M. M. & Levine, J. M. Phenotypic Plasticity and Species Coexistence. *Trends in Ecology & Evolution* **31**, 803-813, doi:10.1016/j.tree.2016.07.013 (2016).

39 McPeek, M. A. The Ecological Dynamics of Natural Selection: Traits and the Coevolution of Community Structure. *Am. Nat.* **189**, E91-E117, doi:10.1086/691101 (2017).

40 Chacón-Labella, J., García Palacios, P., Matesanz, S., Schöb, C. & Milla, R. Plant domestication disrupts biodiversity effects across major crop types. *Ecology Letters* **22**, 1472-1482, doi:10.1111/ele.13336 (2019).

41 Vos, P. *et al.* AFLP - A NEW TECHNIQUE FOR DNA-FINGERPRINTING. *Nucleic Acids Res.* **23**, 4407-4414, doi:10.1093/nar/23.21.4407 (1995).

42 Piry, S. *et al.* GENECLASS2: A software for genetic assignment and first-generation migrant detection. *J. Hered.* **95**, 536-539, doi:10.1093/jhered/esh074 (2004).
43 Rannala, B. & Mountain, J. L. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences* **94**, 9197-9201, doi:10.1073/pnas.94.17.9197 (1997).
44 Nei, M., Tajima, F. & Tateno, Y. Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**, 153-170, doi:10.1007/bf02300753 (1983).
45 Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**, 945-959 (2000).
46 Falush, D., Stephens, M. & Pritchard, J. K. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* **7**, 574-578, doi:10.1111/j.1471-8286.2007.01758.x (2007).
47 Jakobsson, M. & Rosenberg, N. A. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801-1806, doi:10.1093/bioinformatics/btm233 (2007).
48 Gelman, A., Hwang, J. & Vehtari, A. Understanding predictive information criteria for Bayesian models. *Statistics and computing* **24**, 997-1016, doi:10.1007/s11222-013-9416-2 (2014).

Acknowledgements

The authors thank, Vincent Beguier, head of research at the Jouffray Drillaud Company (Saint Sauvant, France) and his technical team for the installation of the experiment, the URP3F technical team and particularly Dominique Dénoue and Brigitte Bonneau who provided important experimental assistance. The authors thank four anonymous referees and the Associate Editor for their relevant comments on the manuscript. The Agence Nationale de la Recherche, France (PRAISE, ANR-13- ADAP-0015) funded this work. Julien Meilhac was supported by the Region Poitou Charente and INRA (BAP division and MP EcoServ) for his PhD salary. VM and LD were supported by the Natural Sciences and Engineering Research Council of Canada (NSERC-Discovery-2016-05716).

Author contributions

I.L. initiated the research question, obtained the funds and led the working group. I.L., J.M., and S.F. collected the data and J.M. and S.F. organised the dataset and all authors coordinated the analyses. J.M., I.L., V.M. and L.D. drafted the manuscript and all authors contributed to the final manuscript.

Author Information

Supplementary information is available online. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to I.L. (isabelle.litrico-chiarelli@inra.fr).

Table 1 | M**odels describing the dynamics of species-niches, trough time and genetic selection**

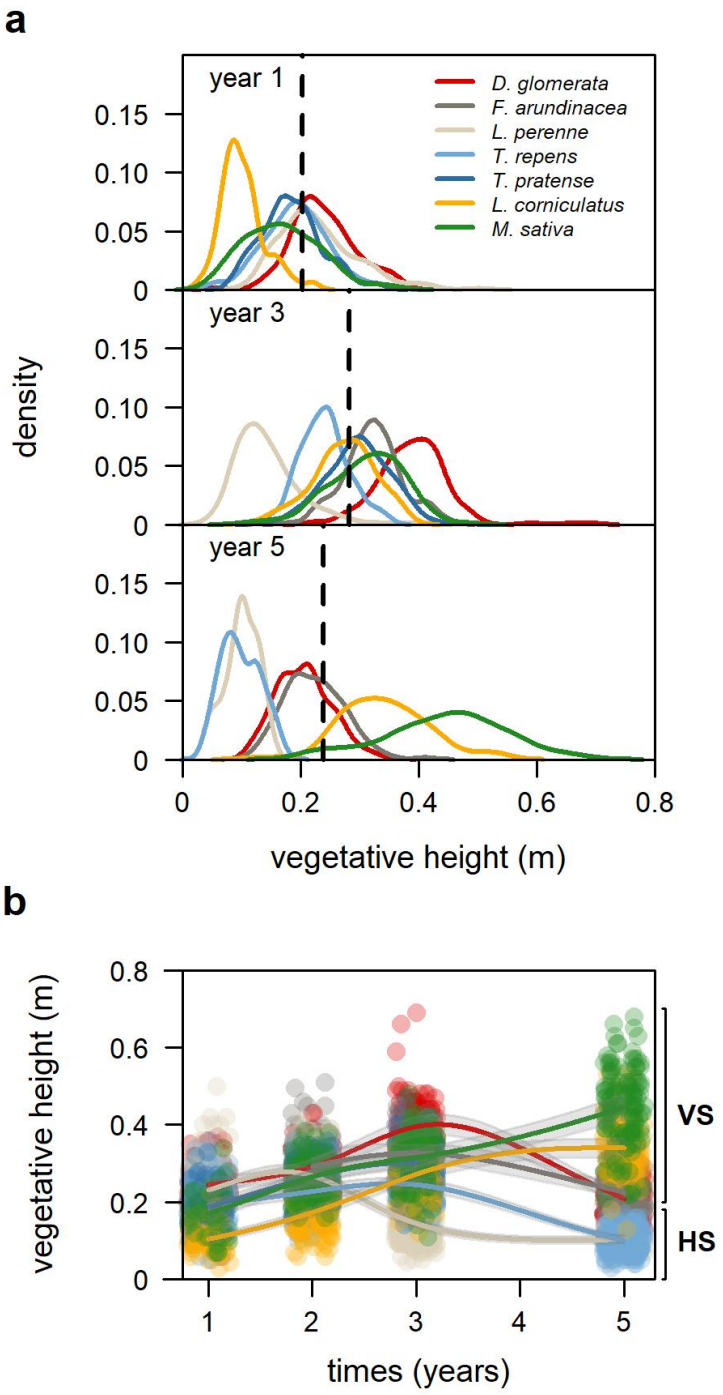
Models	Specific mean (μ)	Specific dispersion (σ)	Time (t)	Genetic Diversity (div)	Cultivar proportion (prop)	WAIC (± s.e)	Δ_{WAIC} (± s.e.)
Model selection - niches							
Mod_0						27452 (±84)	3392 (±96)
Mod_t			X			26142 (±105)	2081 (±80)
$Mod_{t,\mu}$	X		X			24290 (±108)	230 (±33)
$Mod_{t,\mu,\sigma}$	X	X	X			24060 (±101)	45 (±33)
Mod_{div}	X	X	X	X		24014 (±101)	0
Model selection - genetic selection							
$Mod_{0'}$						42178 (±100)	5871 (±119)
Mod_{prop}					X	42001 (±102)	5694 (±119)
$Mod_{t,\mu,\sigma'}$	X	X	X			36389 (±133)	82 (±21)
$Mod_{t,prop}$	X	X	X		X	36306 (±133)	0

Legends of figures

Figure 1 | Niche differentiation. a, Density distribution of vegetative heights between species at respectively one, three, and five years after sowing the mixtures. Dotted line represents average height of mixtures giving the approximate canopy height (average height of all species combined). One species (Festuca arundinacea) is missing at the beginning as its establishment took more time and Trifolium pratense had disappeared by the fifth year. **b,** Modelled mean height of each species, with shaded area representing 95% credible interval around the mean after accounting for variability between mixtures. HS indicates species with a horizontal strategy and VS with a vertical strategy.

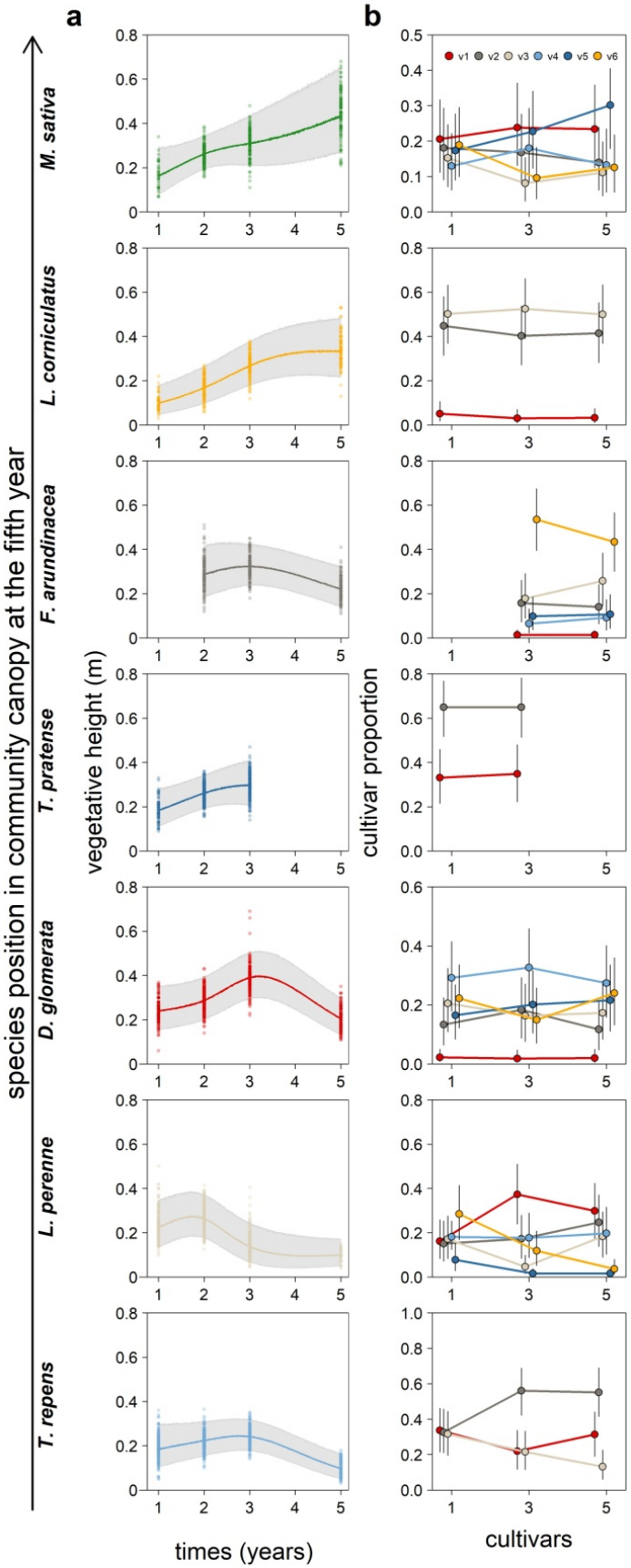
Figure 2 | Phenotypic and genotypic dynamics over the five years of experimentation. a, Evolution of species trait distribution over time. Solid line represents the mean of each species trait distribution and the shaded area the 95% predictive interval, after accounting for variability between mixtures. **b,** Evolution of cultivar proportions over times for each species for the mixture with high complexity of genetic structure (M-5 see supplementary method). Species are sorted by their height at the fifth year and cultivars are ranked along the height values of isolated plants (from shortest with v1 to tallest with v6). *Trifolium pratense* disappeared by the fifth year but was still present in the third year. *Festuca arundinacea* was missing in the first year because its establishment took longer than the other species. Cultivar proportion was computed from individual samples which taken several weeks after phenotypic measurements (*F. arundinacea* was present). Values are median ± error (90% credible interval).

Figure 3 | Species abundance. Percentage of species dry mass per year with *Trifolium repens* (white with black dots), *Lolium perenne* (light grey with black dots), *Dactylis glomerata* (dark grey hatched), *Trifolium pratense* (black), *Festuca arundinacea* (light grey hatched), *Lotus corniculatus* (white) and *Medicago sativa* (dark grey). Species are ranked by their position in the canopy in the fifth year.



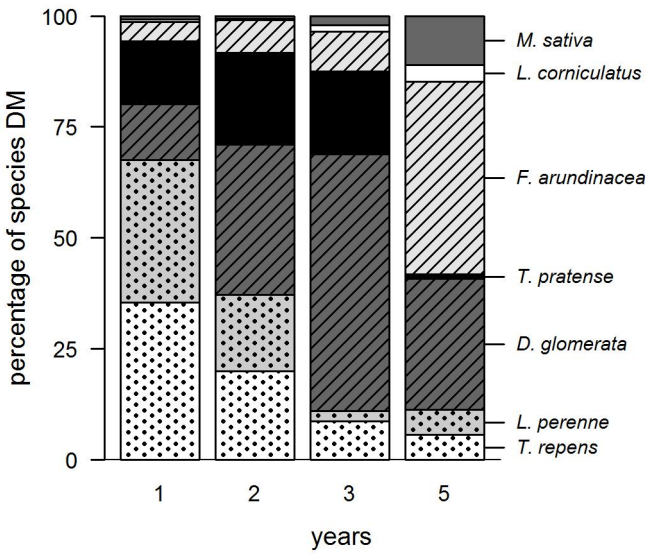
499

500 **Figure 1 | Niche differentiation.**



501

502 **Figure 2 | Phenotypic and genotypic dynamics over the five years of experimentation.**



503

504 **Figure 3 | Species abundance**

Table S1. Proportions of species and cultivars sown for mixtures

Species	Cutlivar	Designation	Proportions (seed-mass basis)					Height cultivars (cm)	
			M-1	M-2	M-3	M-4	M-5	Mean ± s.e.	Tukey test
Dactylis glomerata			**0.115**	**0.115**	**0.115**	**0.115**	**0.115**		
	E1V5	V1	0.000	0.000	0.000	0.000	0.166	17.4 ± 0.5	a
	Accord	V2	0.000	0.000	0.000	0.333	0.166	22.8 ± 0.4	b
	Lucullus	V3	0.000	0.000	0.000	0.333	0.166	24.4 ± 0.5	bc
	Vaillant	V4	1.000	0.000	0.000	0.333	0.166	26.3 ± 0.9	c
	Otop	V5	0.000	0.000	1.000	0.000	0.166	26.9 ± 0.7	c
	President	V6	0.000	1.000	0.000	0.000	0.166	29.8 ± 0.5	d
Festuca arundinacea			**0.231**	**0.231**	**0.231**	**0.231**	**0.231**		
	E3V5	V1	0.000	0.000	1.000	0.000	0.250	24.0 ± 0.4	a
	Soni	V2	1.000	0.000	0.000	0.500	0.125	24.4 ± 0.8	a
	Mariellendo	V3	0.000	0.000	0.000	0.000	0.125	24.5 ± 0.4	a
	Elodie	V4	0.000	0.000	0.000	0.250	0.125	25.4 ± 0.4	ab
	Noria	V5	0.000	0.000	0.000	0.250	0.125	26.2 ± 0.4	b
	Gardian	V6	0.000	1.000	0.000	0.000	0.250	30.3 ± 0.7	c
Lolium perenne			**0.115**	**0.115**	**0.115**	**0.115**	**0.115**		
	Juras	V1	0.000	0.000	0.000	0.333	0.166	19.9 ± 0.7	a
	Gagny	V2	0.000	0.000	0.000	0.333	0.166	20.4 ± 0.7	a
	E6V5	V3	0.000	0.000	0.000	0.000	0.166	21.8 ± 0.8	b
	Aberstar	V4	1.000	0.000	0.000	0.333	0.166	21.9 ± 0.4	b
	Rgmaroc	V5	0.000	0.000	1.000	0.000	0.166	23.7 ± 0.4	c
	Tonnus	V6	0.000	1.000	0.000	0.000	0.166	24.9 ± 0.1	d
Trifolium repens			**0.115**	**0.115**	**0.115**	**0.115**	**0.115**		
	Abervantage	V1	1.000	0.000	0.000	0.333	0.333	15.7 ± 0.4	a
	Aran	V2	0.000	0.000	0.000	0.333	0.333	18.0 ± 0.5	b
	Giga	V3	0.000	1.000	1.000	0.333	0.333	19.7 ± 0.2	c
Trifolium pratense			**0.115**	**0.115**	**0.115**	**0.115**	**0.115**		
	Formica	V1	0.000	0.000	1.000	0.333	0.333	23.1 ± 1.1	-
	Diplo	V2	1.000	1.000	0.000	0.666	0.666	23.8 ± 0.8	-
Lotus corniculatus			**0.078**	**0.078**	**0.078**	**0.078**	**0.078**		
	Leo	V1	1.000	0.000	0.000	0.500	0.500	15.3 ± 0.5	a
	PX-ete	V2	0.000	0.000	1.000	0.000	0.250	20.2 ± 0.9	b
	Altus	V3	0.000	1.000	0.000	0.500	0.250	24.2 ± 1.3	c
Medicago sativa			**0.231**	**0.231**	**0.231**	**0.231**	**0.231**		
	Luzelle	V1	0.000	0.000	0.000	0.000	0.166	38.0 ± 0.7	a
	Rafia	V2	0.000	0.000	1.000	0.000	0.166	52.4 ± 1.7	b
	Galaxie	V3	0.000	1.000	0.000	0.333	0.166	53.3 ± 1.3	b
	Kali	V4	0.000	0.000	0.000	0.333	0.166	53.6 ± 1.2	b
	timbale	V5	1.000	0.000	0.000	0.333	0.166	54.0 ± 1.7	b
	Meldor	V6	0.000	0.000	0.000	0.000	0.166	55.0 ± 0.7	b

The proportion is given as a fraction of seed mass. Each cultivar used is described by a height value (cm) from measurements of plants in a nursery of isolated plants (mean ± s.e.). Different letters indicate significant differences between cultivars of a species ($P<0.05$). A cultivar code is assigned for cultivars not registered in an official catalogue. "Designation" refers to the cultivar code used in the paper. For all mixtures, seeds were from the same lots and provided by company Jouffray-Drillaud (JD).

Table S2. Number of individuals sampled for the cultivar abundance measure by fingerprinting

Species	Block	Mixture M-4					Mixture M-5				
		Maximum number sampled	Years 1	Year 3	Year 5	P_{NA}	Maximum number sampled	Years 1	Year 3	Year 5	P_{NA}
Dactylis glomerata											
	I	48	47	48	45	0.121	96	92	91	92	0.109
	II	48	46	44	41	0.069	96	94	91	92	0.058
Festuca arundinacea											
	I	48	37	48	48	0.053	96	39	32	92	0.104
	II	48	42	48	47	0.087	96	89	39	74	0.055
Lolium perenne											
	I	48	47	14	33	0.000	96	96	17	8	0.000
	II	48	46	14	21	0.000	96	93	13	17	0.000
Trifolium repens											
	I	48	46	37	25	0.120	48	46	36	32	0.097
	II	48	48	41	17	0.170	48	48	45	23	0.069
Trifolium pratense											
	I	32	32	14	-	0.044	32	32	20	-	0.000
	II	32	32	19	-	0.020	32	31	30	-	0.033
Lotus corniculatus											
	I	32	32	5	3	0.025	48	48	8	4	0.012
	II	32	32	3	4	0.026	48	48	8	7	0.064
Medicago sativa											
	I	48	46	12	11	0.130	96	89	34	32	0.065
	II	48	38	5	8	0.078	96	77	27	25	0.016

The number varies according to the species in function to the number of cultivars established, 16 individuals (maximum) per cultivar within species. *Trifolium pratense* had disappeared by the fifth year. P_{NA} is the proportion of individuals no-assigned among the three samplings (years one, three, and five).

Table S3. Primer combinations (SSR) from Polymerase Chain Reaction for *Lolium Perenne*

SSR markers	5' Primer sequence	3' primer sequence
LpSSR066	GCCAGTGCCCATTCCGATAA	CCCCACTCCAACCAAAGCAA
LpSSR058	CGATGAACTCAAGGGGGATT	GCACCGGTCTAGGGACAGAA
rv0641	TGCATAACTTCACTGCAGCATA	AGAAACTCGGTAGAAGGACCTC
M15185	GGTCTGGTAGACATGCCTAC	TACCAGCACAGGCAGGTTC
B4D7op	CTGGCSCTGTGCTCCGYG	TCGCCGTCCACCTACCAC
LmgSSR01-08H	ATGGAACTCTGGCACACCAG	GCATGGCTACATCCTTCCAG

Table S4. Primer combinations per species for Amplified Fragment Length Polymorphism

Species	AFLP selective primer	Number of markers
Dactylis glomerata		**162**
	E-ATT/M-CCG	60
	E-ACA/M-CGG	102
Festuca arundinacea		**96**
	E-ACA/M-CTCG	38
	E-AAG/M-CCTA	21
	E-ACC/M-CAG	37
Trifolium repens		**116**
	E-ACA/M-CCT	65
	E-AAC/M-CCA	51
Trifolium pratense		**78**
	E-ACA/M-CTT	56
	E-ATC/M-CCT	22
Lotus corniculatus		**161**
	E-AAG/M-CTG	42
	E-AAA/M-CAC	51
	E-ACT/M-CTA	39
	E-ATC/M-CAT	29
Medicago sativa		**381**
	E-ACA/M-CTG	82
	E-ACC /M-CAA	41
	E-ACA /M-CGT	53
	E-AAA /M-CGC	40
	E-ATG/M-CCG	37
	E-AGC /M-CCC	26
	E-ACA/M-CTGG	20
	E-AAA/M-CTCG	23
	E-ATG/M-CGC	24
	E-ATA/M-CGG	35

Table S5. Proportion of no-assignation (P_{NA}) in the reference source per cultivar and per species (mean of cultivars)

Species	Cultivar	designation	P_{NA}
Dactylis glomerata			**0.060**
	E1V5	V1	0.022
	Accord	V2	0.022
	Lucullus	V3	0.110
	Vaillant	V4	0.076
	Otop	V5	0.098
	President	V6	0.044
Festuca arundinacea			**0.037**
	E3V5	V1	0.000
	Soni	V2	0.021
	Mariellendo	V3	0.042
	Elodie	V4	0.087
	Noria	V5	0.042
	Gardian	V6	0.031
Lolium perenne			**0.000**
	Juras	V1	0.000
	Gagny	V2	0.000
	E6V5	V3	0.000
	Aberstar	V4	0.000
	Rgmaroc	V5	0.000
	Tonnus	V6	0.000
Trifolium repens			**0.046**
	Abervantage	V1	0.064
	Aran	V2	0.054
	Giga	V3	0.021
Trifolium pratense			**0.027**
	Formica	V1	0.021
	Diplo	V2	0.032
Lotus corniculatus			**0.041**
	Leo	V1	0.000
	PX-ete	V2	0.075
	Altus	V3	0.049
Medicago sativa			**0.065**
	Luzelle	V1	0.022
	Rafia	V2	0.032
	Galaxie	V3	0.133
	Kali	V4	0.101
	timbale	V5	0.066
	Meldor	V6	0.034

Table S6 | Models describing the dynamics of species-niches time in each block, trough and genetic selection.

Models	Specific mean (μ)	Specific dispersion (σ)		Genetic Diversity (div)	Cultivar proportion (prop)	WAIC (± s.e)	Δ_{WAIC} (± s.e.)
Model selection - niches							
Mod_0						27443 (±83)	3554 (±100)
Mod_t			X			26066 (±106)	2177 (±86)
$Mod_{t,\mu}$	X		X			24205 (±108)	316 (±50)
$Mod_{t,\mu,\sigma}$	X	X	X			23951 (±101)	62 (±38)
Mod_{div}	X	X	X	X		23889 (±103)	0
Model selection - genetic selection							
$Mod_{0'}$						42126 (±100)	6112 (±133)
Mod_{prop}					X	41938(±102)	5924(±122)
$Mod_{t,\mu,\sigma'}$	X	X	X			36129 (±133)	115 (±25)
$Mod_{t,prop}$	X	X	X		X	36014 (±133)	0

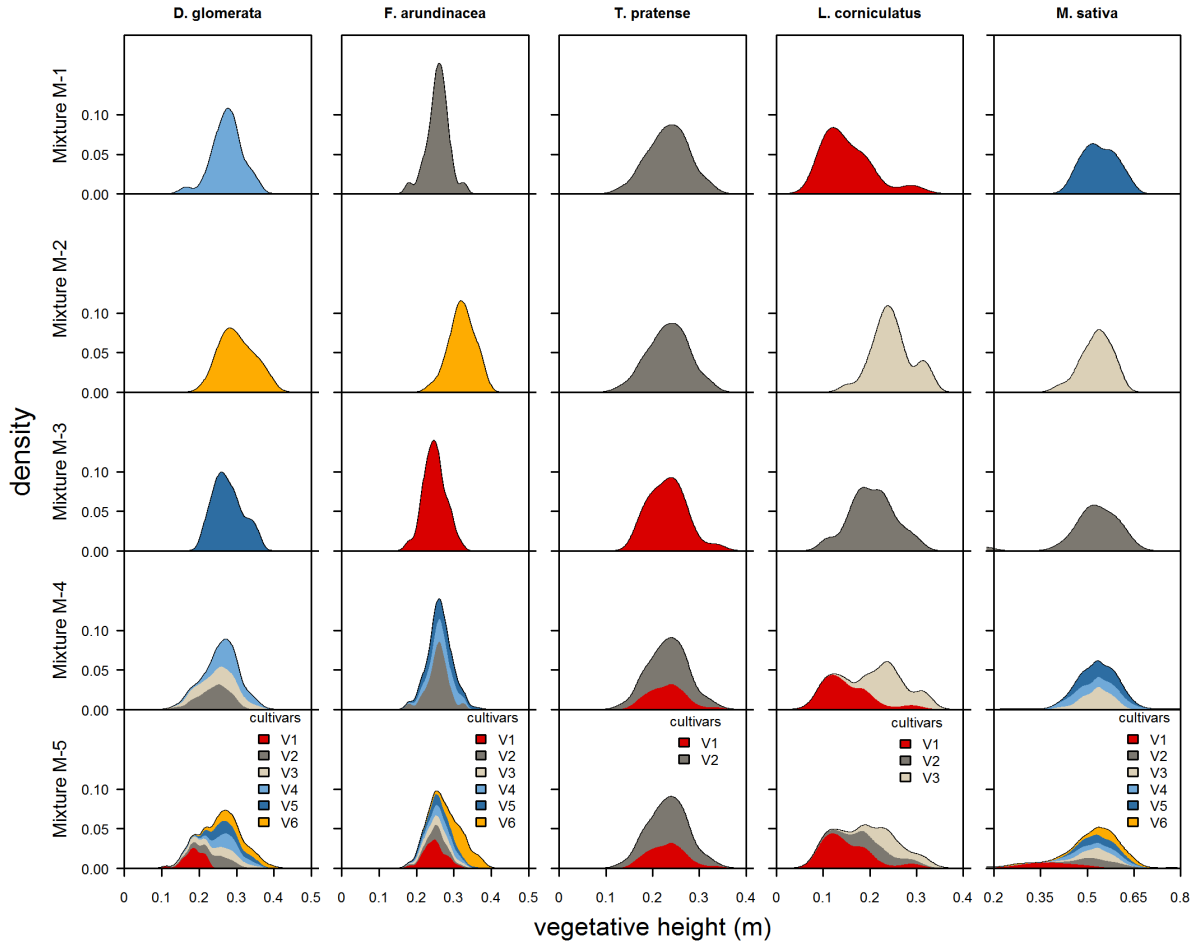


Figure S1 | Range of phenotypic variability of mixtures for species with an elongation strategy in relation to light competition. Each cultivar of each species used in the mixtures was established in a nursery as isolated plants with 30 individuals per cultivar. Vegetative height was measured on all individuals to estimate the phenotypic variability of each cultivar with the value distributions estimated from the nursery. For multi-cultivar mixtures, the distributions were estimated using a bootstrap method with respect to the different cultivar proportions within species and the distribution of values of each cultivar estimated from the nursery.

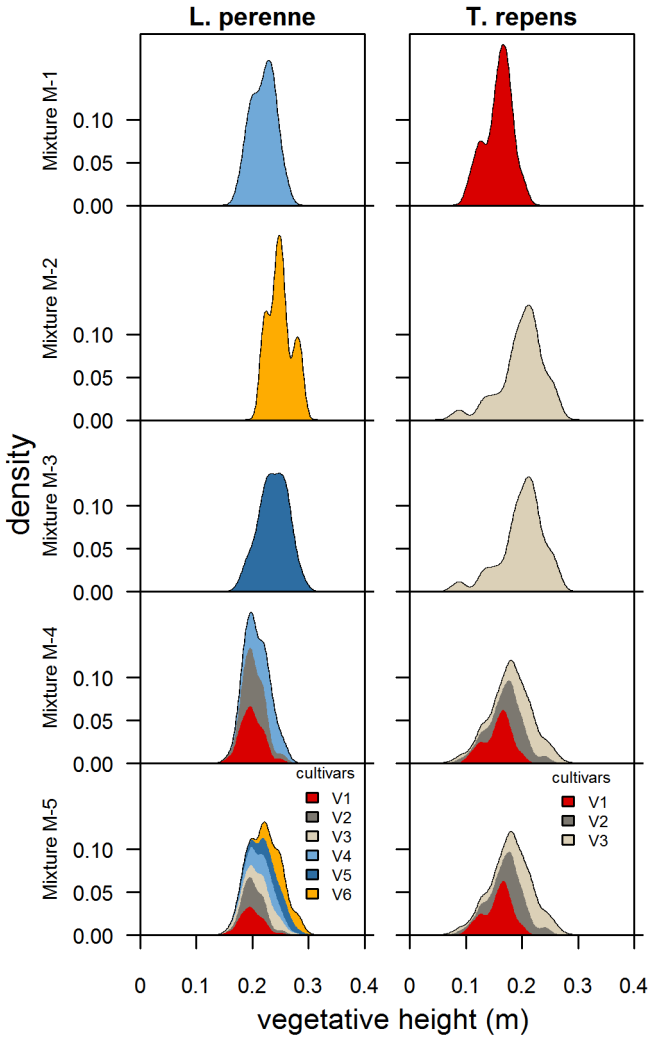


Figure S2 | Range of phenotypic variability of mixtures for species with horizontal spreading strategy in relation to light competition. Each cultivar of each species used in the mixtures was established in a nursery of isolated plant with 30 individuals per cultivar. Vegetative height was measured in all individuals to estimate phenotypic variability of cultivars with the distribution values estimated from the nursery. For multi-cultivar mixtures, distribution was estimated using a bootstrap method with respect to the different cultivar proportions within species and the distribution of values of each cultivar estimated from the nursery.

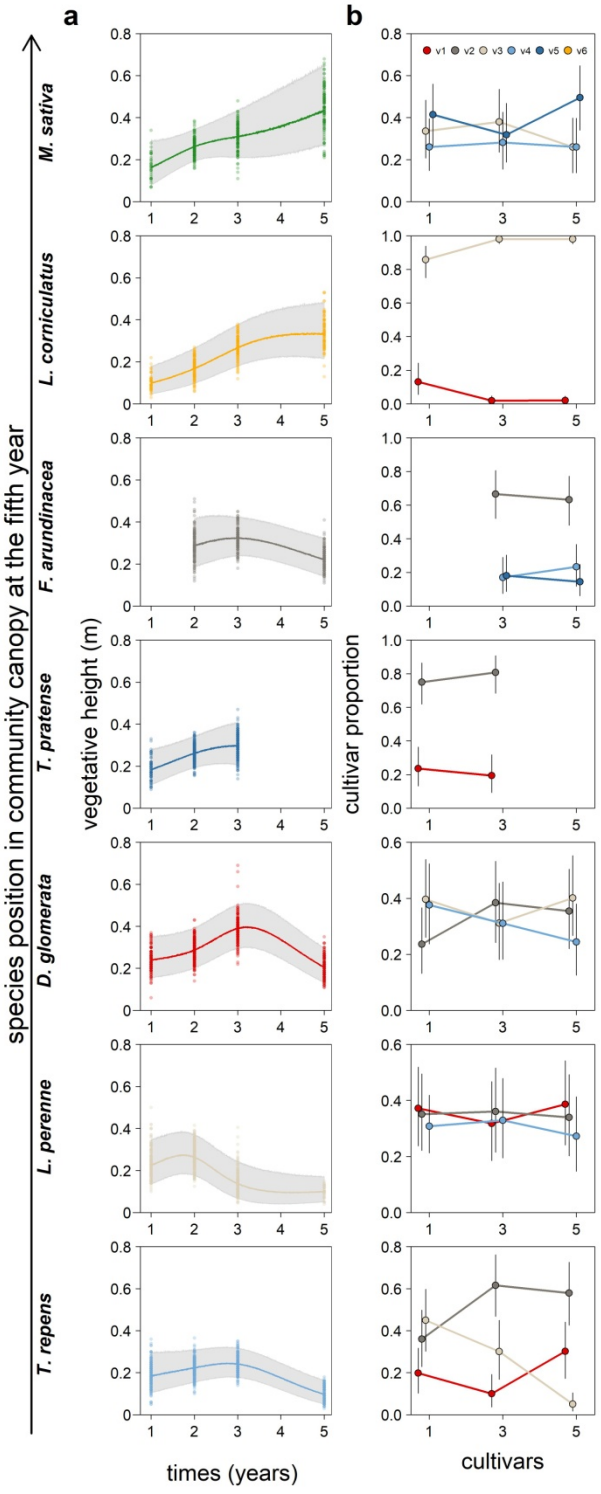


Figure S3 | Phenotypic (all mixtures) and genotypic dynamics (Mixture M-4) throughout the five years of the experiment. a, Evolution of species trait distribution over time. Solid lines represent the mean of each species trait distribution and the shaded areas represent the 95% predicted interval, after accounting for the variability between mixtures. **b**, Cultivar

proportions for each species at years one, three and five after establishment (respectively black, grey and white bars) for the mixture with intermediate complex structure (M-4). Species are sorted by their heights at the fifth year and cultivars are ranked along their height values in the nursery for isolated plants (from shortest to tallest). *Trifolium pratense* had disappeared the fifth year, but was still present in the third year. *Festuca arundinacea* was missing in the first year because its establishment took longer than other species. Cultivar proportions were computed from individual samples which were taken several weeks after phenotypic measurements (when *Festuca arundinacea* was present*)*. Values are means ± s.e.

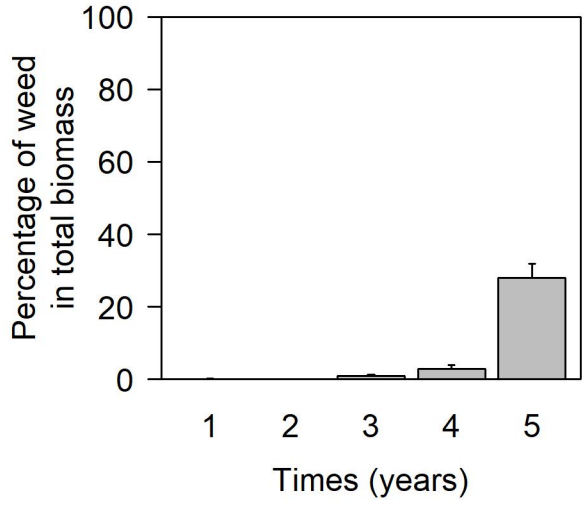


Figure S4 | Percentage of weed in total biomass production by year. The percentage corresponds to the mean of mixtures. mean + s.e.

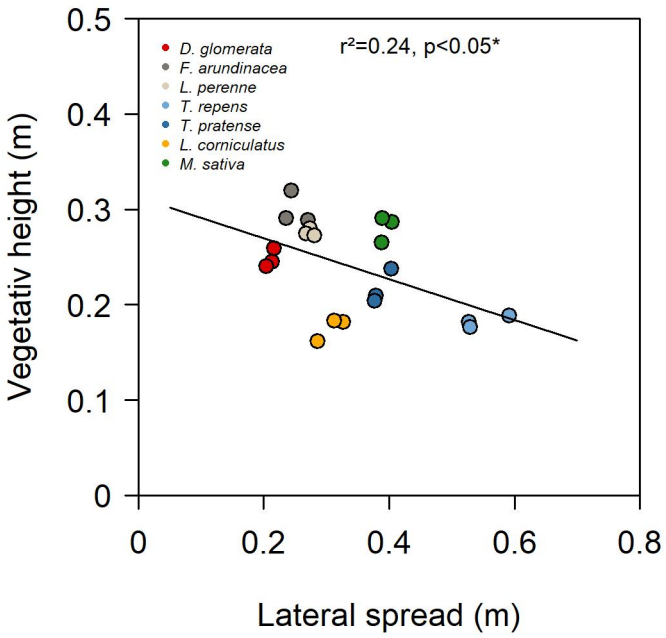


Figure S5: Correlation between vegetative height and lateral spread measured in common garden. Each point represents the mean of species for each block. So each species is represented by three points, the regression is significant ($p<0.05$) with $r^2=0.24$. Lateral spread corresponds to the diameter occupied by plant on the ground.

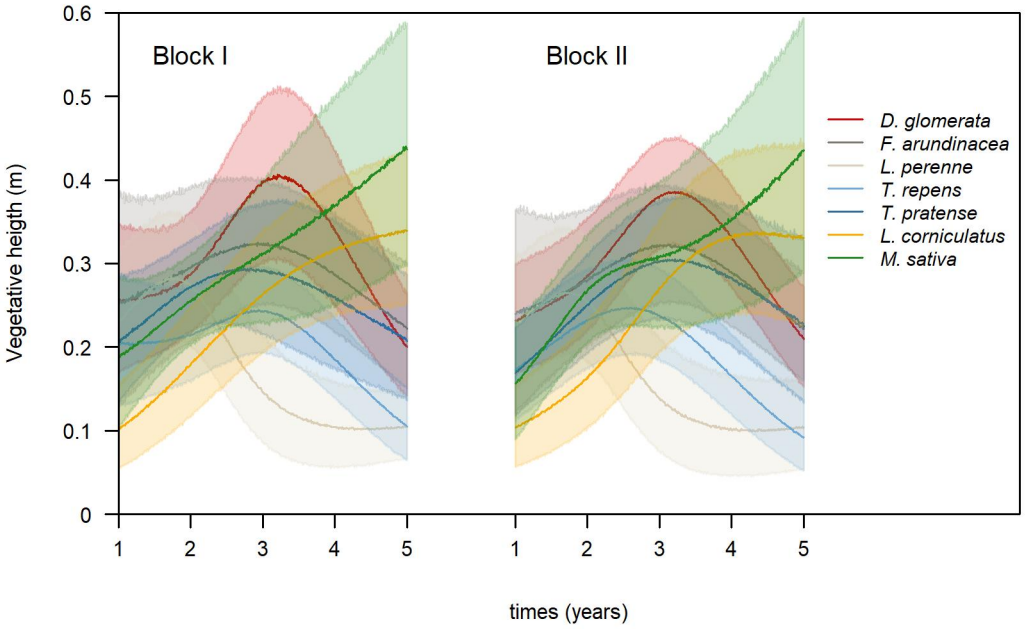


Figure S6 | Modelled vegetative height distribution of each species in each block through time, with shaded area representing 90% predictive interval around the mean after accounting for variability between mixtures.

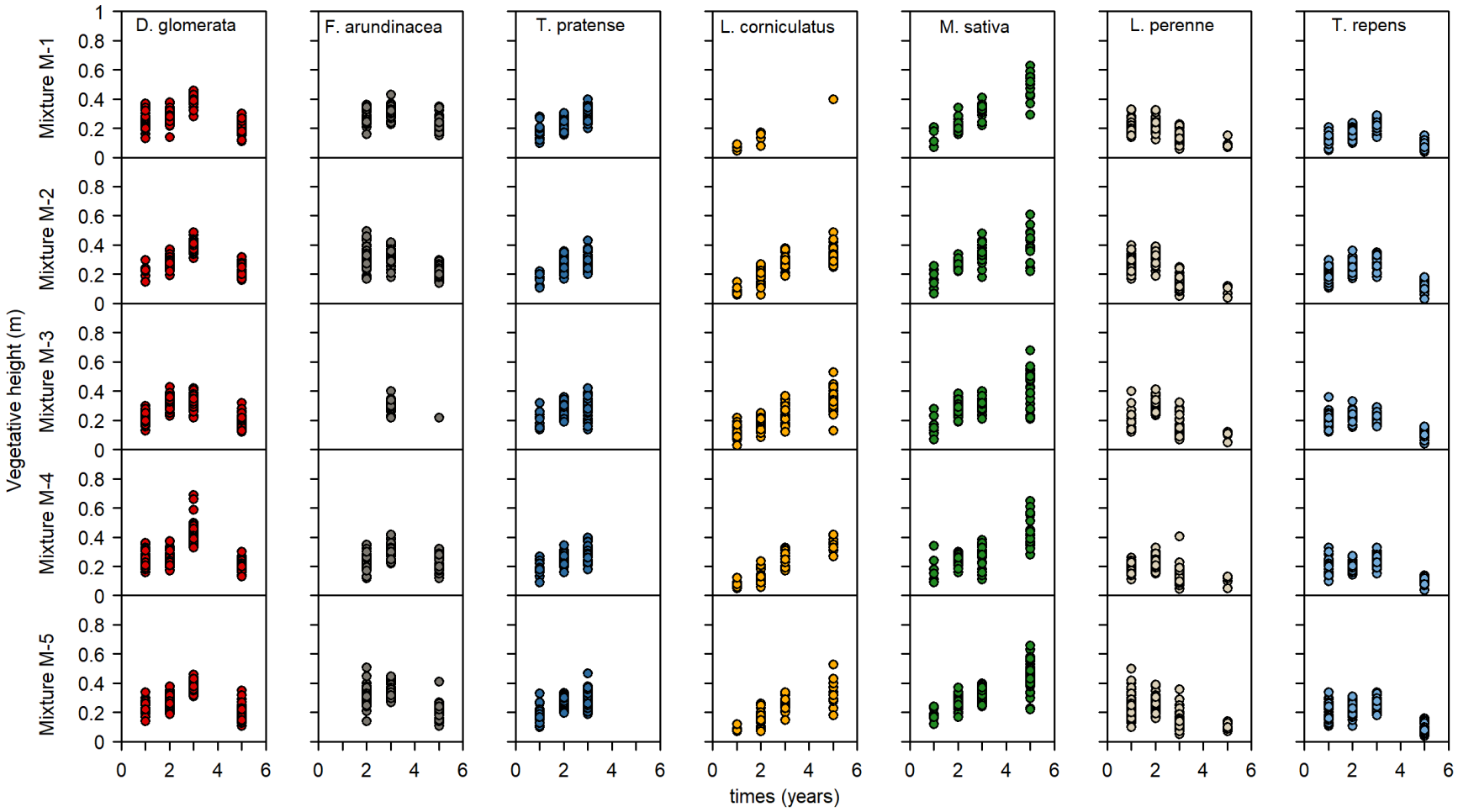


Figure S7 | Changes of vegetative heights over time for each species, *in situ*, in each mixture (M-1 to M-5).

Complete description of hierarchical distributional models and their validations

GLM framework

The classical framework of the generalized linear models (GLM) allows to model any random variable distributed following an exponential family distribution (such as normal, poisson, gamma…). Many of these distributions are constantly bounded, implying a nonlinear variation of their mean and variance close to these bounds. For example, the variation of the mean of a strictly positive value, such as a count, is not linear close to zero, just as the variation of a probability is nonlinear close to its bounds (zero and one). The GLM framework allows to model the effects of linear predictors on a *function* of the response variable, which makes the latter linear. The corresponding function is caller the *link function*, $g()$. The variance, which has also a non-linear behavior at the bound, is then linked to the mean by a *variance function* (Smyth 1989). Thus, the variation of the mean and variance of an exponential family distribution following a series of linear predictors $\mathbf{x_i^T \beta}$ is

$$\mu_i = g^{-1}(\mathbf{x_i^T \beta})$$
$$\sigma_i^2 = \phi_i w_i^{-1} v(\mu_i)$$

Where μ_i and σ_i^2 are the mean and the variance of the distribution for the observation i, respectively, and g^{-1} the inverse of the link function. $v()$ is a non-negative variance function specific to each probability distribution and w_i^{-1} are known weights. ϕ_i is an unknown dispersion parameter for the observation i.

Distributional modelling

Many distributions, even if they are not properly belonging to the exponential family, might be reparametrized to model conjointly mean and variance. Given the fact that we can define the link between mean and variance as below, one would be able to regress both μ_i and ϕ_i as the result of deterministic equations. In this framework, we can define a distributional model for location and scale as follow (Rigby & Stasinopoulos 2005):

$$\mathbf{Y} \sim f(\boldsymbol{\mu}, \boldsymbol{\phi})$$
$$g_1(\boldsymbol{\mu}) = \mathbf{X}\beta$$
$$g_2(\boldsymbol{\phi}) = \mathbf{X}\gamma$$

Where Y is a column vector of the response variable, μ and ϕ are vectors of linear parameters, X a matrix of predictors, and β and γ are vector coefficients linking predictors to the mean and the dispersion, respectively. $g_1()$ and $g_2()$ are the link function for the mean and dispersion, respectively.

Gamma distribution

We modeled strictly positive vegetative height with a gamma distribution, with α being the shape parameter and β being the rate parameter. The formulation with independent mean, μ and dispersion, ϕ, is as follow:

$$\mathbf{Y} \sim Gamma\,(\boldsymbol{\alpha} = \boldsymbol{\mu}^2\boldsymbol{\phi}, \boldsymbol{\beta} = \boldsymbol{\mu}\boldsymbol{\phi})$$
$$log(\boldsymbol{\mu}) = \mathbf{X}\beta$$
$$log(\boldsymbol{\phi}) = \mathbf{X}\gamma$$

This parametrization is straightforward to recover, given that the first two moments of the gamma distribution are defined as follow:

$$E(Y) = \frac{\alpha}{\beta}$$
$$Var(Y) = \frac{\alpha}{\beta^2}$$

Hierarchical additive model formulation

Because of the structured nature of our design (individuals within species within mixtures within blocks), we used hierarchical parameters for mixture and species categorical effects (Gelman *et al.* 2013). Thus, intercepts for each species and each mixtures were modelled as belonging to the same respective population with estimated variance. The complete formulation of the model Mod_3, which explores the adjustment of each species mean and dispersion through time is:

$$y_i \sim Gamma(\mu_i, \phi_i)$$
$$log(\mu_i) = \beta_0 + \beta_B + \beta_M + \beta_S + f_{1s}(t_i)$$
$$log(\phi_i) = \gamma_0 + \gamma_B + \gamma_M + \gamma_S + f_{2s}(t_i)$$
$$f_{1s} = \sum_{k=1}^{K} \alpha_{1Sk}\, b_k(t_i)$$
$$f_{2s} = \sum_{k=1}^{K} \alpha_{2Sk}\, b_k(t_i)$$
$$\beta_M \sim normal(0, \sigma_{\beta_M})$$
$$\beta_S \sim normal(0, \sigma_{\beta_S})$$
$$\gamma_M \sim normal(0, \sigma_{\gamma_M})$$
$$\gamma_S \sim normal(0, \sigma_{\gamma_S})$$
$$\alpha_{1Sk} \sim normal(0, \sigma_{\alpha_{1Sk}})$$
$$\alpha_{2Sk} \sim normal(0, \sigma_{\alpha_{2Sk}})$$

Where y_i is the height value of individual i. The β and γ parameters are the coefficients for mean and dispersion, respectively. Thus, β_0 and γ_0 are the intercepts, while β_B and γ_B are the coefficient for the individuals harvested in the second block. β_M and γ_M are hierarchical deviation parameters for each mixture m, and β_S and γ_S are hierarchical intercepts for each

species. f_{1s} and f_{2s} are smooth functions describing the evolution of mean and dispersion of each species through each value of time t_i, respectively. f_{1s} and f_{2s} are defined as the weighted sum of the $K = 4$ cubic basis b_k, with $\alpha_{1_{sk}}$ and $\alpha_{1_{sk}}$ being the weights of each basis for each species, distributed normally with estimated variances $\sigma_{\alpha_{1Sk}}$ and $\sigma_{\alpha_{1Sk}}$, respectively.

We used weakly informative priors to optimize posterior sampling with appropriate constraints and scaling, without eliciting actual knowledge. Hierarchical priors are centered on 0 with relatively low scale to potentially shrink coefficient and avoid overfitting, while avoiding algorithms to get lost in the posterior surface. Student-t priors with 3 degree-of-freedom have thick tails, allowing algorithms to explore correctly the posterior surface even if the scale is insufficiently large. Priors of Mod_3 are defined as follow, $student^+$ begin positive truncated student-t distributions:

$$\begin{aligned}
\beta_0 &\sim student(3,3,3) & \gamma_0 &\sim student(3,0,3)\\
\beta_B &\sim student(3,0,1) & \gamma_B &\sim student(3,0,1)\\
\sigma_{\beta_M} &\sim student^+(3,0,1) & \sigma_{\gamma_M} &\sim student^+(3,0,1)\\
\sigma_{\beta_S} &\sim student^+(3,0,1) & \sigma_{\gamma_S} &\sim student^+(3,0,1)\\
\sigma_{\alpha_{1sk}} &\sim student^+(3,0,1) & \sigma_{\alpha_{2sk}} &\sim student^+(3,0,1)
\end{aligned}$$

The model Mod_4, which described the adjustment of species height mean and dispersion for three level of population's genetic diversity is formulated as follow:

$$\begin{aligned}
y_i &\sim Gamma(\mu_i, \phi_i)\\
log(\mu_i) &= \beta_0 + \beta_B + \beta_M + \beta_{SD} + f_{1SD}(t_i)\\
log(\phi_i) &= \gamma_0 + \gamma_B + \gamma_M + \gamma_{SD} + f_{2SD}(t_i)\\
f_{1SD} &= \sum_{k=1}^{K} \alpha_{1SDk}\, b_k(t_i)\\
f_{2sd} &= \sum_{k=1}^{K} \alpha_{2SDk}\, b_k(t_i)\\
\beta_M &\sim normal(0, \sigma_{\beta_M})\\
\beta_S D &\sim normal(0, \sigma_{\beta_S D})\\
\gamma_M &\sim normal(0, \sigma_{\gamma_M})\\
\gamma_S D &\sim normal(0, \sigma_{\gamma_S D})\\
\alpha_{1SDk} &\sim normal(0, \sigma_{\alpha_{1SDk}})\\
\alpha_{2SDk} &\sim normal(0, \sigma_{\alpha_{2SDk}})
\end{aligned}$$

With β_{SD} and γ_{SD} being intercepts and f_{1SD} and f_{2SD} being smooth functions of time for each species s at three level of genetic diversity, d (low, medium and high), describing the mean and dispersion of height distribution, respectively.

The priors are as follow:

$$\begin{aligned} \beta_0 &\sim student(7,3,3) & \gamma_0 &\sim student(7,0,3) \\ \beta_B &\sim student(7,0,1) & \gamma_B &\sim student(7,0,1) \\ \sigma_{\beta_M} &\sim student^+(7,0,1) & \sigma_{\gamma_M} &\sim student^+(7,0,1) \\ \sigma_{\beta_{SD}} &\sim student^+(7,0,1) & \sigma_{\gamma_{SD}} &\sim student^+(7,0,1) \\ \sigma_{\alpha_{1SDk}} &\sim student^+(7,0,1) & \sigma_{\alpha_{2SDk}} &\sim student^+(7,0,1) \end{aligned}$$

The model Mod_2', which described the adjustment of species height mean and dispersion in function of time *and* in function of changes in cultivars abundances is defined as follow. Because cultivar proportions have only been measured for three years, the we reduced the number of cubic basis $K = 3$.

$$y_i \sim Gamma(\mu_i, \phi_i)$$
$$log(\mu_i) = \beta_0 + \beta_B + \beta_M + \beta_S + \beta_V + \beta_{1V} frequency + f_{1S}(t_i)$$
$$log(\phi_i) = \gamma_0 + \gamma_B + \gamma_M + \gamma_S + \gamma_V + \gamma_{1V} frequency + f_{2S}(t_i)$$
$$f_{1SD} = \sum_{k=1}^{K} \alpha_{1SDk}\, b_k(t_i)$$
$$f_{2sd} = \sum_{k=1}^{K} \alpha_{2SDk}\, b_k(t_i)$$
$$\beta_M \sim normal(0, \sigma_{\beta_M})$$
$$\beta_S \sim normal(0, \sigma_{\beta_S})$$
$$\gamma_M \sim normal(0, \sigma_{\gamma_M})$$
$$\gamma_S \sim normal(0, \sigma_{\gamma_S})$$
$$\begin{bmatrix} \beta_V \\ \beta_{1V} \\ \gamma_V \\ \gamma_{1VS} \end{bmatrix} = MVNormal(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, S)$$
$$\alpha_{1Sk} \sim normal(0, \sigma_{\alpha_{1Sk}})$$
$$\alpha_{2Sk} \sim normal(0, \sigma_{\alpha_{2Sk}})$$

With β_V and γ_V being hierarchical intercepts describing the effect of the presence of each cultivar on the height distribution of species s in the mixture m and the block b. β_{1V} and γ_{1V} are hierarchical slopes describing the effect on species s height distributions of the change in cultivar v abundance through time. All of these parameters where distributed multinormally with S being a covariance matrix computed as follow:

$$S = \begin{bmatrix} \sigma_{\beta_V} & 0 & 0 & 0 \\ 0 & \sigma_{\beta_{1V}} & 0 & 0 \\ 0 & 0 & \sigma_{\gamma_V} & 0 \\ 0 & 0 & 0 & \sigma_{\gamma_{1V}} \end{bmatrix} R \begin{bmatrix} \sigma_{\beta_V} & 0 & 0 & 0 \\ 0 & \sigma_{\beta_{1V}} & 0 & 0 \\ 0 & 0 & \sigma_{\gamma_V} & 0 \\ 0 & 0 & 0 & \sigma_{\gamma_{1V}} \end{bmatrix}$$

Where R is an estimated 4x4 correlation matrix. The priors for this model are as follow, the LKJcorr distribution being a modified beta distribution to become a prior of correlation matrices defined on [-1,1].

$$\begin{aligned}
\beta_0 &\sim student(3,3,3) & \gamma_0 &\sim student(3,0,3) \\
\beta_B &\sim student(3,0,1) & \gamma_B &\sim student(3,0,1) \\
\sigma_{\beta_M} &\sim student^+(3,0,1) & \sigma_{\gamma_M} &\sim student^+(3,0,1) \\
\sigma_{\beta_S} &\sim student^+(3,0,1) & \sigma_{\gamma_S} &\sim student^+(3,0,1) \quad R \sim LKJcorr(2) \\
\sigma_{\beta_V} &\sim student^+(3,0,1) & \sigma_{\gamma_V} &\sim student^+(3,0,1) \\
\sigma_{\beta_1 V} &\sim student^+(3,0,1) & \sigma_{\gamma_1 V} &\sim student^+(3,0,1) \\
\sigma_{\alpha_{1Sk}} &\sim student^+(3,0,1) & \sigma_{\alpha_{2Sk}} &\sim student^+(3,0,1)
\end{aligned}$$

Model diagnosis

Sampling behaviors

The *No-U-Turn sampler* implemented in *stan* returned several diagnostics. For every model, we ensured that there were no divergent transitions. Divergent transitions arise when the curvature of an area of the posterior surface is too high to be adequately explored by the sampler. This diagnostic is very important, and this situation shall be avoided, because it leads to biased estimates (Gelman *et al.* 2013). Well-chosen priors and the cleverly parametrized *stan* code produced by *brms* (Bürkner 2018) avoided such suboptimal sampling behavior and parametrizing the step size at 0.95 in the sampler was sufficient to fit every model without divergences. Once the sampler behaves correctly, the second important concern is the mixing of the different chains. The chains represent independent instances of posterior surface exploration, beginning at various starting points. A rule of thumb is that a Gelman-Rubin split $\hat{R}$ greater than 1.1 indicates bad mixing of chains (Gelman *et al.* 2013). Every $\hat{R}$ in the models were lower than 1.01. We also inspected each chain visually to ensure the absence of problematic behaviors. Each chain were constituted of 2000 iterations including 1000 warm-up iterations.

Posterior-Predictive checks

We checked that models recovered key features of data by confronting visually predicted and observed distributions.

Mod3

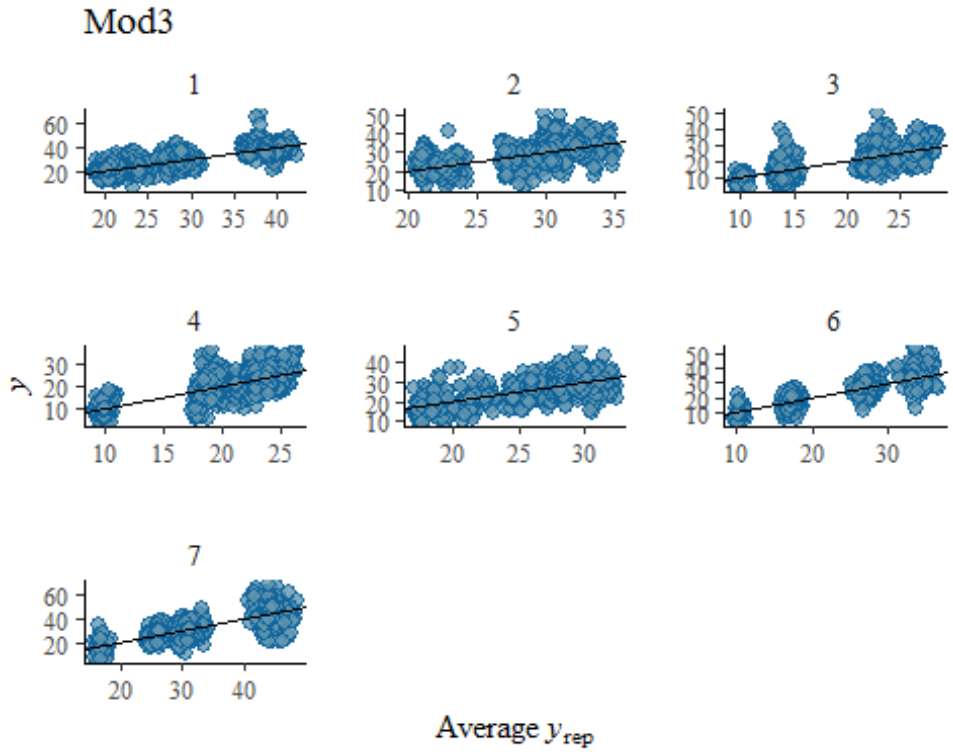


Figure A 1: Scatterplots of observed individual vegetative heights in function of predicted values for each species. Line represents the 1:1 relationship. Species 1: *Dactylis glomerata*, 2: *Festuca arundinacea*; 3: *Lolium perenne*; 4: Trifolium repens, 5: Trifolium pratense, 6: *Lotus corniculatus*; 7: *Medicago sativa*

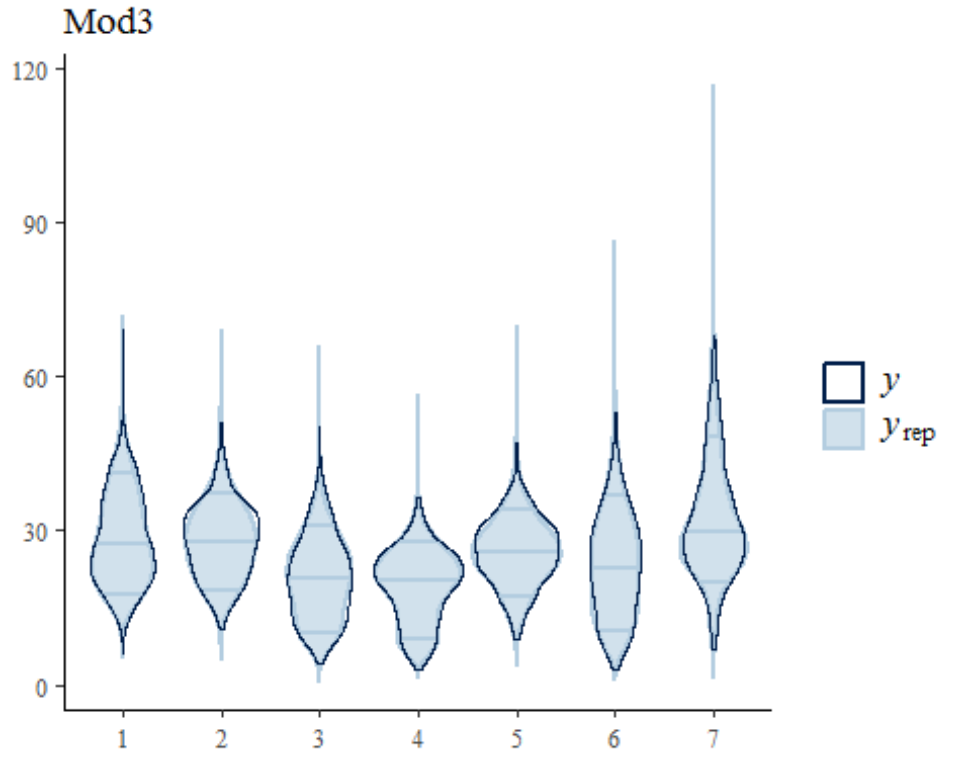


Figure A 2: Violin plots of each species observed and predicted vegetative heights distribution. Species 1: *Dactylis glomerata*, 2*: Festuca arundinacea*; 3: *Lolium perenne*; 4: *Trifolium repens*, 5: *Trifolium pratense*, 6: *Lotus corniculatus*; 7: *Medicago sativa*

Mod4

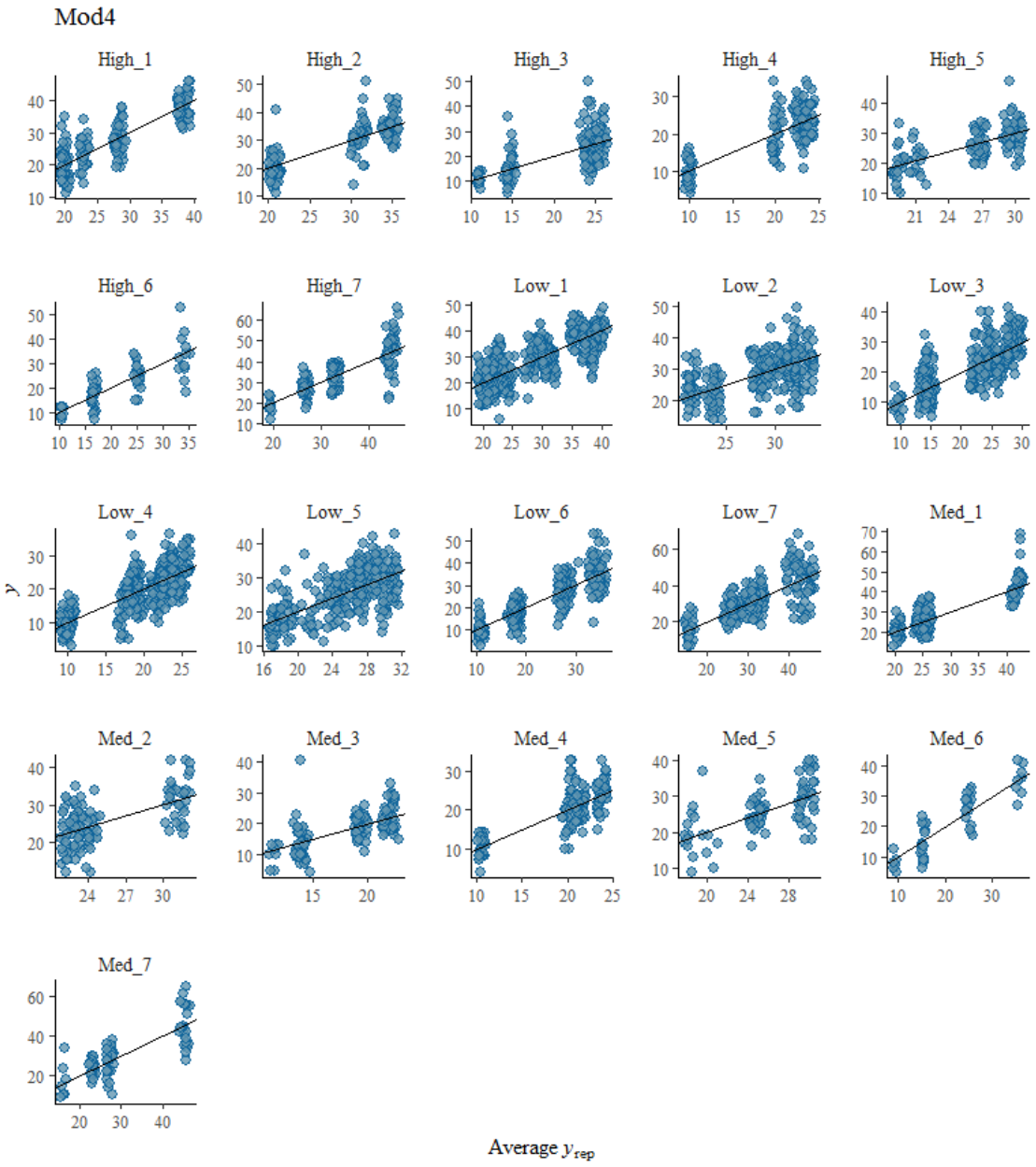


Figure A 3: Scatterplots of observed individual vegetative heights in function of predicted values, for each species at three level of genetic diversity. Line represents the 1:1 relationship. Species 1: *Dactylis glomerata*, 2: *Festuca arundinacea*; 3: *Lolium perenne*; 4: *Trifolium repens*, 5: *Trifolium pratense*, 6: *Lotus corniculatus*; 7: *Medicago sativa*

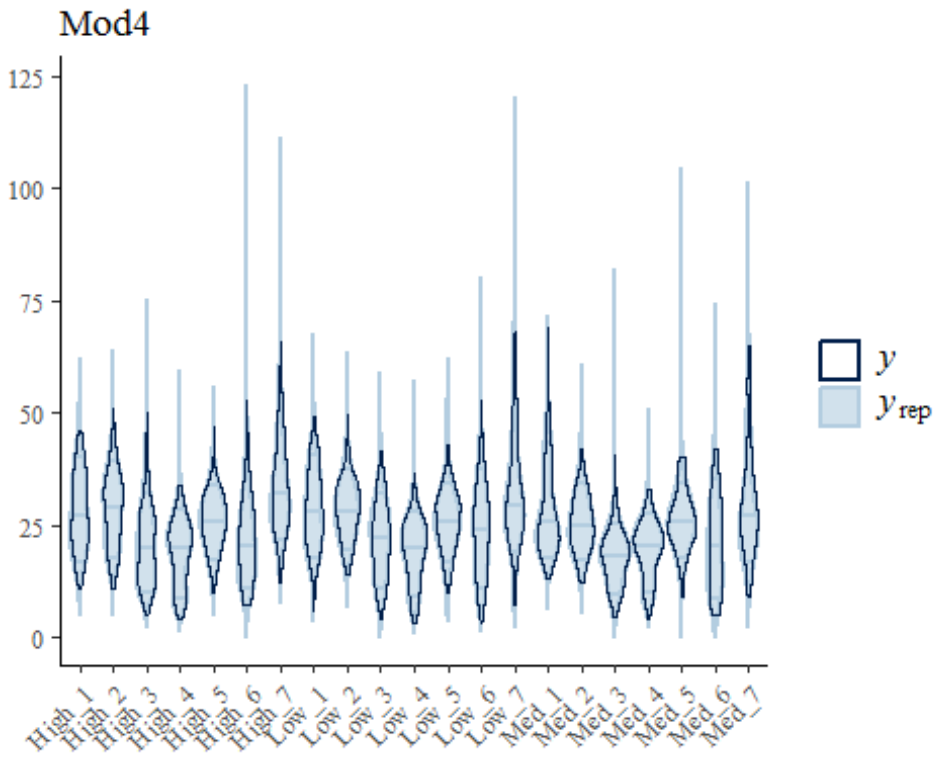


Figure A 4: Violin plots of each species observed and predicted vegetative height distribution, at three levels of genetic diversity. Species 1: *Dactylis glomerata*, 2: *Festuca arundinacea*; 3: *Lolium perenne*; 4: *Trifolium repens*, 5: *Trifolium pratense*, 6: *Lotus corniculatus*; 7: *Medicago sativa*

Mod2'

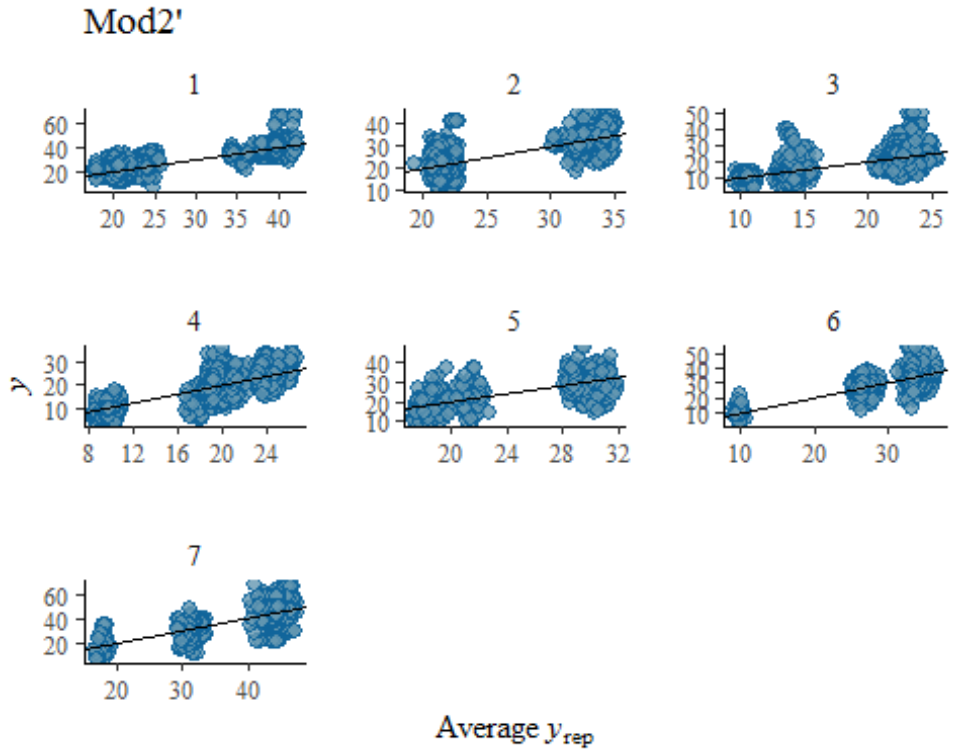


Figure A 5: Scatterplot of observed individual vegetative heights in function of predicted values. Line represents the 1:1 relationship. Species 1: Dactylis glomerata, 2: *Festuca arundinacea*; 3: *Lolium perenne*; 4: Trifolium repens, 5: Trifolium pratense, 6: *Lotus corniculatus*; 7: *Medicago sativa*

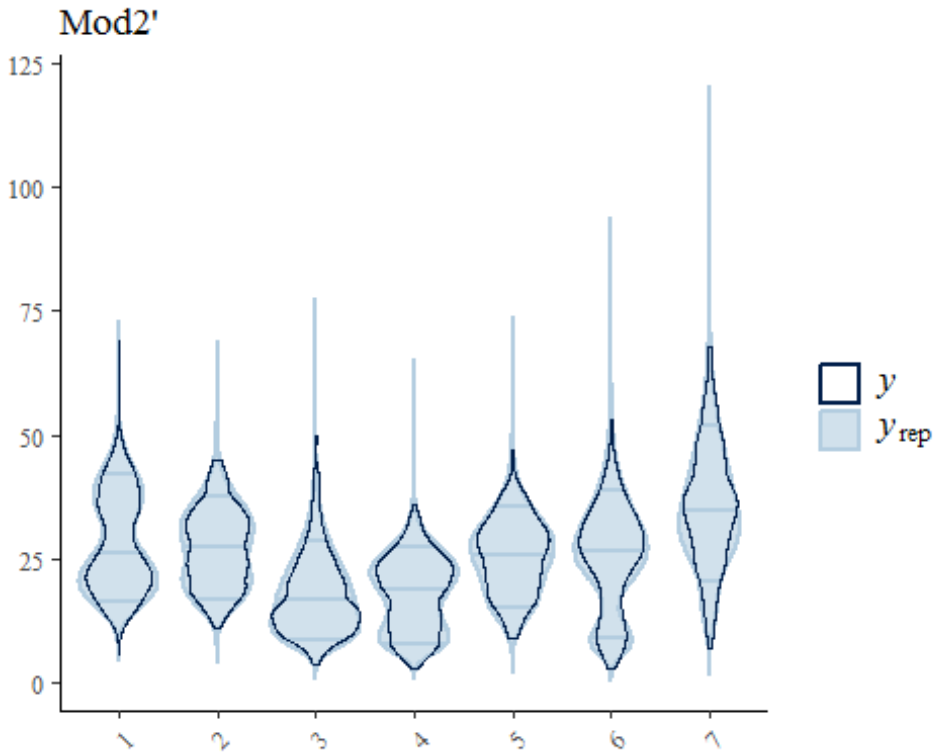


Figure A 6: Violin plot of each species observed and predicted vegetative height distribution. Species 1: *Dactylis glomerata*, 2: *Festuca arundinacea*; 3: Lolium perenne; 4: *Trifolium repens*, 5: *Trifolium pratense*, 6: *Lotus corniculatus*; 7: *Medicago sativa*.

Literature cited

Bürkner, P. C. Advanced Bayesian Multilevel Modeling with the *R* Package *brms*. R J. 10, 395–411 (2018).

Gelman, A. et al. Bayesian data analysis. (Chapman and Hall/CRC, 2013).

Rigby, R. A. & Stasinopoulos, D. M. Generalized additive models for location, scale and shape. Appl. Stat. 54, 507–554 (2005).

Smyth, G. K. Generalized Linear Models with Varying Dispersion. J. R. Stat. Soc. Ser. B 51, 47–60 (1989).