SUPPLEMENTARY INFORMATION FOR

Evolution of species complementarity in response to drought in a grassland biodiversity experiment

Authors: Yuxin Chen, Anja Vogel, Cameron Wagg, Tianyang Xu, Maitane Iturrate Garcia, Michael Scherer-Lorenzen, Alexandra Weigelt, Nico Eisenhauer, Bernhard Schmid

***Correspondence**: yuxin.chen@xmu.edu.cn and bernhard.schmid@uzh.ch

Supplementary Table 1 | Species information and number of replicates of species composition (identity of species pair) before the drought event in the glasshouse. Numbers in brackets indicate the numbers of replicates for ambient- vs. drought-selected plants, respectively. Data in the diagonal are for monocultures (in blue, planted in blocks 1–4); data in the lower diagonal are for 2-speices mixtures (in red, planted in blocks 1–4). The last two rows are for individuals from pots with one individual planted in blocks 1–4 and 5 (in orange), respectively. Upper-case letters indicate functional groups (G=grass, S=short herb, T=tall herb, L=legume).

Species full name	Short name	AP1 (G)	AP2 (G)	BP (S)	CB (T)	LC (L)	PL (S)	PM (S)	PT (G)	PV (S)	RA (T)	TO (S)	TD (L)
Alopecurus pratensis	AP1	(11, 11)	--	--	--	--	--	--	--	--	--	--	--
Avenula pubescens	AP2	(8, 7)	(9, 9)	--	--	--	--	--	--	--	--	--	--
Bellis perennis	BP	(10, 10)	--	(14, 14)	--	--	--	--	--	--	--	--	--
Crepis biennis	CB	(1, 1)	--	--	(2, 2)	--	--	--	--	--	--	--	--
Lotus corniculatus	LC	--	--	--	--	(1, 1)	--	--	--	--	--	--	--
Plantago lanceolata	PL	(8, 8)	--	--	(1, 1)	--	(8, 7)	--	--	--	--	--	--
Plantago media	PM	--	--	--	(1, 1)	--	--	(16, 16)	--	--	--	--	--
Poa trivialis	PT	(11, 10)	(7, 7)	--	--	--	--	(16, 15)	(15, 15)	--	--	--	--
Prunella vulgaris	PV	--	(4, 4)	(8, 8)	--	--	(7, 7)	--	--	(8, 8)	--	--	--
Rumex acetosa	RA	--	(5, 6)	(12, 12)	--	(1, 1)	--	--	--	(8, 8)	(14, 14)	--	--
Taraxacum officinale	TO	--	--	--	--	--	(4, 4)	(3, 3)	(4, 4)	--	--	(4, 4)	--
Trifolium dubium	TD	--	--	--	--	--	(5, 5)	--	(6, 5)	--	--	--	(7, 7)
Individual (blocks 1–4)		(11, 11)	(9, 9)	(14, 14)	(2, 2)	(2, 3)	(7, 8)	(16, 16)	(15, 15)	(8, 8)	(14, 14)	(6, 5)	(5, 5)
Individual (block 5)		(22, 16)	(12, 13)	(13, 15)	(5, 4)	(2, 2)	(11, 11)	(14, 13)	(17, 15)	(4, 4)	(8, 11)	(3,3)	(12, 11)

Supplementary Table 2 | Significance tests against zero for net biodiversity effect (NE), complementarity effect (CE) and sampling effect (SE) before, during and after the drought event in the glasshouse for the two selection treatments. Block and species composition were set as fixed- and random-effects terms, respectively.

	Ambient-selected plants				Drought-selected plants			
	df	ddf	*F*	*P*	df	ddf	*F*	*P*
Before drought								
NE (127, 128)	1	17.3	4.540	**0.048 (+)**	1	16.6	4.798	**0.043 (+)**
CE (126, 128)	1	13.8	1.675	0.217 (+)	1	14.4	1.653	0.219 (+)
SE (126, 128)	1	18.7	3.201	0.090 (+)	1	17.4	9.429	**0.007 (+)**
During drought								
NE (126, 128)	1	17.4	0.149	0.705 (-)	1	18.6	2.309	0.145 (-)
CE (124, 124)	1	20.0	0.070	0.788 (-)	1	16.9	5.985	**0.026 (-)**
SE (124, 124)	1	19.6	0.051	0.824 (-)	1	18.4	0.013	0.910 (-)
After drought								
NE (109, 110)	1	13.4	0.022	0.882 (-)	1	16.2	10.330	**0.005 (+)**
CE (97, 91)	1	9.9	<0.001	0.993 (-)	1	13.4	21.480	**<0.001 (+)**
SE (97, 91)	1	10.9	0.867	0.372 (-)	1	14.6	14.690	**0.002 (-)**

df, numerator degrees of freedom; ddf, denominator degrees of freedom (these reflect residual degrees of freedom among the 15–21 species pairs [= species compositions] for which biodiversity effects were calculated). *F* and *P* indicate *F* ratios and *P* values of the significance tests, respectively. ± besides the *P* values represents the direction of effects. Numbers within brackets indicate the numbers of pots for plants under ambient or drought conditions in the glasshouse, respectively.

Supplementary Table 3 | Significance tests for the effects of selection treatment on net biodiversity effect (NE), complementarity effect (CE) and sampling effect (SE) before, during and after the drought event in the glasshouse. Block and selection treatment were set as fixed-effects terms; species composition and its interaction with selection treatment were set as random-effects terms.

	df	ddf	*F*	*P*
Before drought				
NE (n=255)	1	11.6	0.035	0.855 (+)
CE (n=254)	1	12.5	0.029	0.867 (+)
SE (n=254)	1	15.7	0.524	0.480 (+)
During drought				
NE (n=254)	1	12.9	1.671	0.219 (-)
CE (n=248)	1	17.8	2.853	0.109 (-)
SE (n=248)	1	18.7	0.070	0.794 (+)
After drought				
NE (n=219)	1	9.1	14.490	**0.004 (+)**
CE (n=188)	1	14.0	22.110	**0.001 (+)**
SE (n=188)	1	11.4	9.988	**0.009 (-)**

df, numerator degrees of freedom; ddf, denominator degrees of freedom (these reflect residual degrees of freedom among the selection responses of 15–21 species pairs [= species compositions] for which biodiversity effects were calculated). *F* and *P* indicate *F* ratios and the *P* values of the significance tests, respectively. ± besides the *P* values represents the direction of difference between drought vs. ambient selection treatments. Numbers within brackets indicate the numbers of mixtures.

Supplementary Table 4 | Significance tests for the effects of selection treatment on stability calculated separately for mixtures and monocultures. Block and selection treatment were set as fixed-effects terms; species composition and its interaction with selection treatment were set as random-effects terms.

	df	ddf	*F*	*P*
Resistance				
Mixture (255)	1	15.4	1.690	0.212 (-)
Monoculture (209)	1	5.1	0.668	0.450 (+)
Recovery				
Mixture (255)	1	6.9	7.388	**0.030 (+)**
Monoculture (206)	1	9.8	0.044	0.839 (-)
Resilience				
Mixture (255)	1	15.8	1.020	0.328 (+)
Monoculture (209)	1	10.2	0.003	0.955 (+)

df, numerator degrees of freedom; ddf, denominator degrees of freedom (these reflect residual degrees of freedom across species compositions). *F* and *P* indicate *F* ratios and the *P* values of the significance tests, respectively. $\pm$ besides the *P* values represents the direction of difference between drought vs. ambient selection treatments. Numbers within brackets indicate the numbers of pots.

Supplementary Table 5 | Significance tests for the effects of selection treatment on the difference in stabilities between mixtures and monocultures. Block and selection treatment were set as fixed-effects terms; species composition and its interaction with selection treatment were set as random-effects terms.

	df	ddf	*F*	*P*
Resistance (254)	1	16.5	4.627	**0.047 (-)**
Recovery (248)	1	19.0	6.550	**0.020 (+)**
Resilience (254)	1	14.7	0.686	0.421 (+)

df, numerator degrees of freedom; ddf, denominator degrees of freedom (these reflect residual degrees of freedom across species compositions). *F* and *P* indicate *F* ratios and the *P* values of the significance tests, respectively. $\pm$ besides the *P* values represents the direction of difference between drought vs. ambient selection treatments. Numbers within brackets indicate the numbers of mixtures.

Supplementary Table 6 | Significance tests for the effects of selection treatment on neighbor interaction intensity in mixtures and monocultures before, during and after the drought event in the glasshouse. Block and selection treatment were set as fixed-effects terms; species composition and its interaction with selection treatment were set as random-effects terms.

	df	ddf	*F*	*P*
Before drought				
Mixture (257)	1	11.2	0.037	0.851 (-)
Monoculture (216)	1	10.4	0.639	0.442 (-)
During drought				
Mixture (253)	1	14.1	6.330	**0.025 (-)**
Monoculture (205)	1	9.7	1.710	0.222 (-)
After drought				
Mixture (245)	1	9.3	0.245	0.632 (+)
Monoculture (195)	1	5.2	0.140	0.720 (-)

df, numerator degrees of freedom; ddf, denominator degrees of freedom (these reflect residual degrees of freedom across species compositions). *F* and *P* indicate *F* ratios and the *P* values of the significance tests, respectively. ± besides the *P* values represents the direction of difference between drought vs. ambient selection treatments. Numbers within brackets indicate the numbers of pots.

Supplementary Table 7 | Significance tests for the difference between heterospecific and conspecific interactions before, during and after the drought event in the glasshouse for the two selection treatments. Block and species composition as set as fixed- and random-effects term, respectively

	df	ddf	*F*	*P*
Before drought				
Drought (127)	1	14.9	0.158	0.696 (+)
Ambient (130)	1	6.3	2.348	0.174 (-)
During drought				
Drought (125)	1	14.2	0.635	0.439 (-)
Ambient (128)	1	17.7	1.963	0.178 (+)
After drought				
Drought (120)	1	15.6	22.680	**<0.001 (+)**
Ambient (123)	1	13.8	4.495	0.053 (+)

df, numerator degrees of freedom; ddf, denominator degrees of freedom (these reflect residual degrees of freedom across species compositions). *F* and *P* indicate *F* ratios and the *P* values of the significance tests, respectively. ± besides the *P* values represents the direction of difference between heterospecific vs. conspecific interaction. Numbers within brackets indicate the numbers of mixtures.

Supplementary Table 8 | Significance tests for the effects of selection treatment on the difference between heterospecific and conspecific interaction before, during and after the drought event in the glasshouse. Block and selection treatment were set as fixed-effects terms; species composition and its interaction with selection treatment were set as random-effects terms.

	df	ddf	*F*	*P*
Before drought (n=257)	1	9.1	1.985	0.192 (+)
During drought (n=253)	1	15.4	2.677	0.122 (-)
After drought (n=243)	1	10.1	10.640	**0.008 (+)**

df, numerator degrees of freedom; ddf, denominator degrees of freedom (these reflect residual degrees of freedom across species compositions). *F* and *P* indicate *F* ratios and the *P* values of the significance test, respectively. ± besides the *P* values represents the direction of difference between drought vs. ambient selection treatments. Numbers within brackets indicate the numbers of mixtures.

Supplementary Table 9 | Significance tests for the effects of selection treatment on traits measured on individual plants without neighbors from general linear models. Numbers within brackets indicate the numbers of individual plants.

	df	*P*
Leaf relative chlorophyll content before drought (n=420)		
Species	11	<0.001
Selection treatment	1	0.847
Species x selection treatment	11	0.175
Leaf area before drought (n=422)		
Species	11	<0.001
Selection treatment	1	0.257
Species x selection treatment	11	0.400
Leaf mass per area before drought (n=422)		
Species	11	<0.001
Selection treatment	1	0.330
Species x selection treatment	11	**0.019**
Leaf osmometric potential before drought (n=359)		
Species	11	<0.001
Selection treatment	1	0.973
Species x selection treatment	10	0.506
Leaf stomatal conductance before drought (n=329)		
Species	11	<0.001
Selection treatment	1	0.500
Species x selection treatment	11	0.886

Leaf stomatal conductance during drought (n=152)		
Species	10	0.012
Selection treatment	1	0.454
Species x selection treatment	10	0.974
Root-shoot biomass ratio after drought (n=368)		
Species	11	<0.001
Selection treatment	1	0.526
Species x selection treatment	11	0.488

df and *P* indicate numerator degrees of freedom and the *P* values of the significance test, respectively.

Supplementary Table 10 | Significance tests for the effects of selection treatment on trait dissimilarity between interacting species in mixtures from mixed-effects models. Block and selection treatment were set as fixed-effects terms; species composition and its interaction with selection treatment were set as random-effects terms. Traits were measured on plants within mixtures.

	df	ddf	F	*P*
Leaf relative chlorophyll content (n=237)	1	14.7	0.912	0.355 (+)
Leaf area (n=232)	1	19.0	3.660	**0.071 (+)**
Leaf mass per area (n=232)	1	12.0	0.041	0.843 (-)
Three traits (n=230)	1	12.3	1.483	0.246 (+)

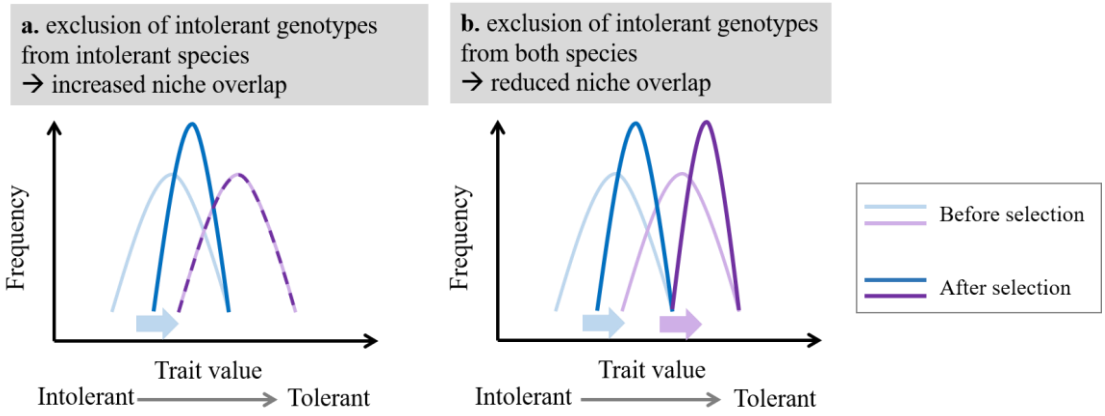
df, numerator degrees of freedom; ddf, denominator degrees of freedom (these reflect residual degrees of freedom across species composition). *F* and *P* indicate *F* ratios and the *P* values of the significance test, respectively. $\pm$ besides the *P* values represents the direction of difference between drought vs. ambient selection treatments. Numbers within brackets indicate the numbers of mixtures.

Supplementary Table 11 | Effects of drought selection x functional group richness history (FGR, as factor) in the field on net biodiversity effect (NE), complementarity effect (CE) and sampling effect (SE) before, during and after the drought event in the glasshouse. Block, drought selection, FGR and drought selection x FGR were set as fixed-effects terms; species composition, species composition x drought selection, species composition x FGR and species composition x drought selection x FGR were set as random-effects terms.

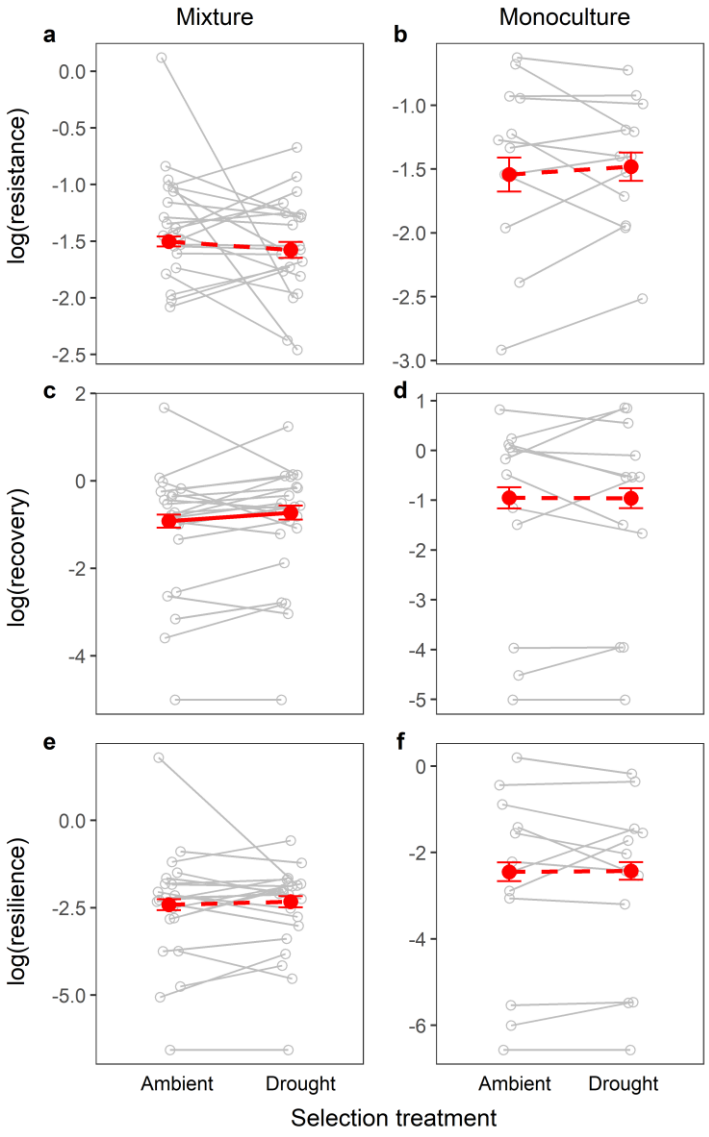
	df	ddf	*F*	*P*
Before drought				
NE (n=242)	3	6.9	0.769	0.547
CE (n=241)	3	8.6	1.736	0.232
SE (n=241)	3	15.1	0.907	0.461
During drought				
NE (n=241)	3	15.4	0.845	0.490
CE (n=235)	3	10.2	1.881	0.196
SE (n=235)	3	15.0	0.370	0.776
After drought				
NE (n=206)	3	16.0	0.334	0.801
CE (n=176)	3	14.0	0.230	0.873
SE (n=176)	3	14.0	1.410	0.282

df, numerator degrees of freedom; ddf, denominator degrees of freedom. *F* and *P* indicate *F* ratios and the *P* values of the significance tests, respectively. Numbers within brackets indicate the numbers of mixtures. We excluded 13 pots from this analysis because their composed two species were from different FGR field plots.

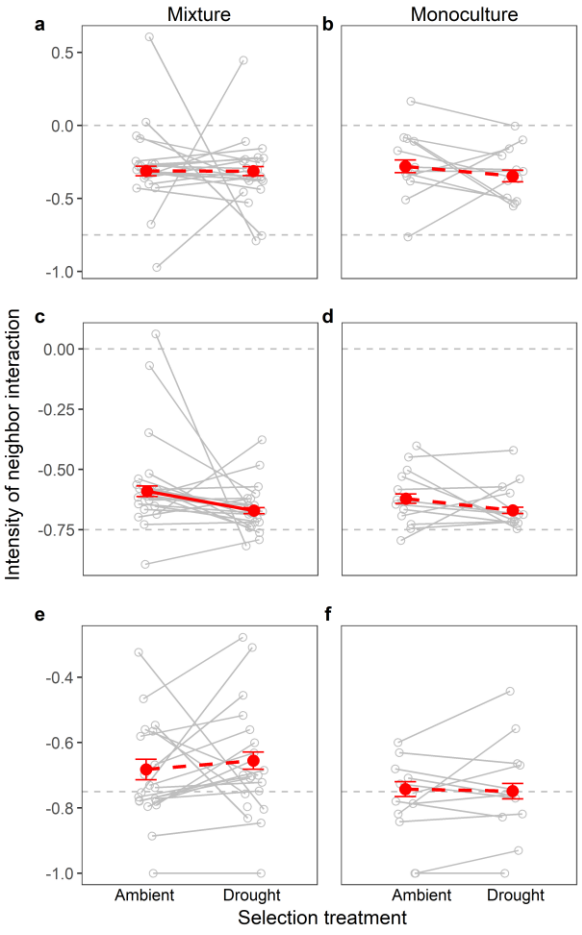
SUPPLEMENTARY FIGURES



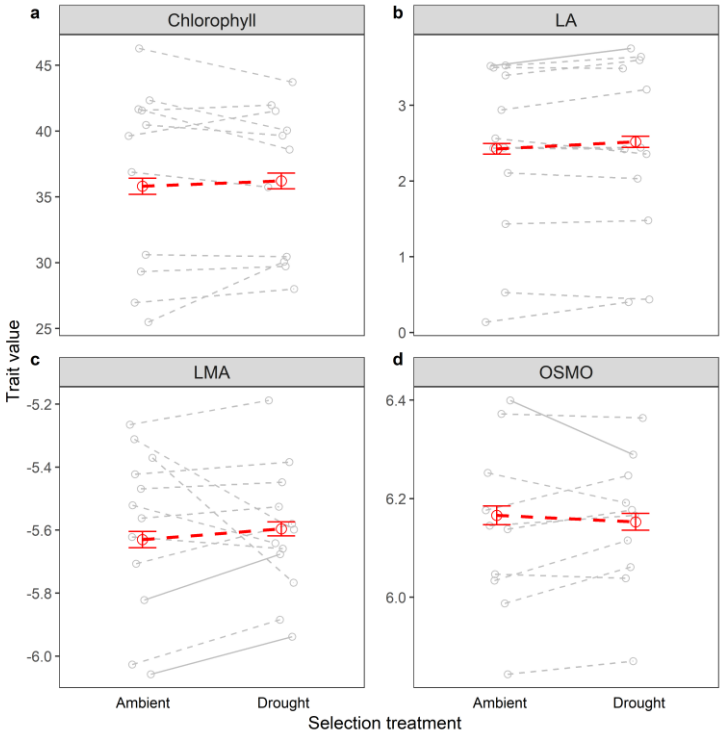
Supplementary Figure 1 | Illustration on how selection for traits associated with drought tolerance may change niche overlaps between two species within communities, under the assumption that selection arises from a filtering of pre-existing genotypes without recombination and without new mutations. Panel a shows how exclusion of intolerant genotypes from intolerant species (blue curves) but not from tolerant species (purple curves) would shift the overall trait distribution of intolerant species closer to that of tolerant species, thus, increasing the overlap of trait distribution or niche. Panel b shows how exclusion of intolerant genotypes could occur both in tolerant and intolerant species when there is strong intraspecific competition within tolerant species. In this case, the trait distribution of both species would shift to the side of higher tolerance and shrink in their ranges, thus reducing the niche overlap between species. Light and dark colors represent the scenarios before and after selection, respectively. Arrows represent the directions of selection.



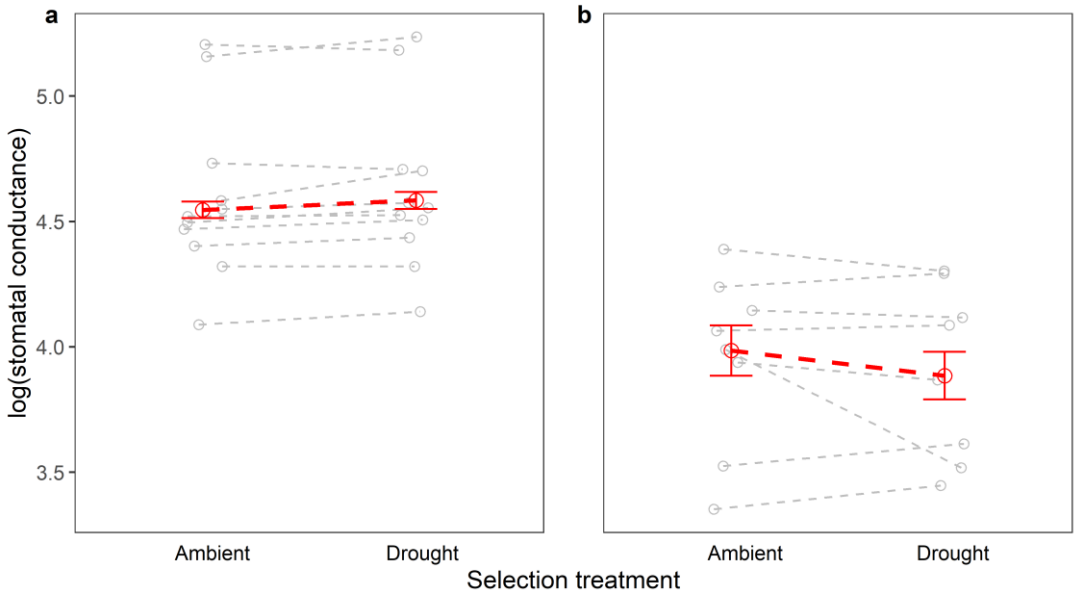
Supplementary Figure 2 | Difference in biomass stabilities (resistance [a–b], recovery [c–d] and resilience [e–f]) in response to the drought event in the glasshouse between selection treatments (plants selected under ambient vs. drought conditions), calculated separately for mixtures (first column) and monocultures (second column). The solid red line indicates a significant difference between the two selection treatments ($P < 0.05$ in mixed-model analysis of variance, see Supplementary Table 4). Red points and error bars show means ± standard error. Grey points represent means for species pairs (standard errors for species pairs were not shown). Grey lines connect the two selection treatments for each species pair.



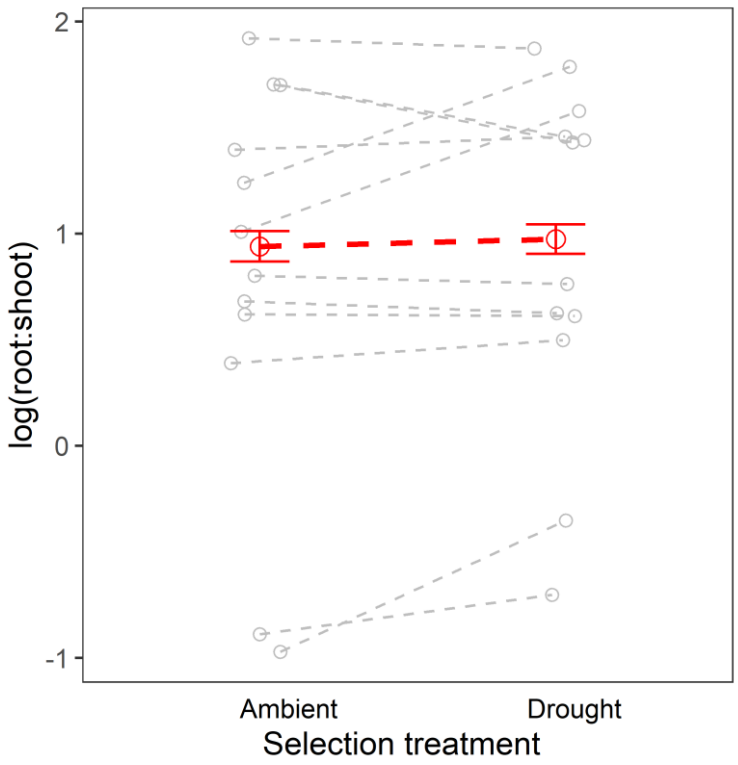
Supplementary Figure 3 | Effects of selection treatment (ambient- vs. drought-selected plants) on neighbor interaction intensity for 2-species mixtures (first column) and monocultures (second column) before (a–b), during (c–d) and after (e–f) the drought event in the glasshouse. The solid red line indicates a significant difference between the two selection treatments ($P < 0.05$ in mixed-model analysis of variance, see Supplementary Table 6). Red points and error bars show means ± standard error. Grey points represent means for species pairs (standard errors for species pairs were not shown). Grey solid lines connect the same species pair between the two selection treatments for each species pair. Grey dashed lines at zero indicate the case when individual plants with vs. without neighbors have the same biomass. Grey dashed lines at –0.75 indicate the case when an individual grown alone in a pot is four times bigger than an individual grown in a pot with four individuals.



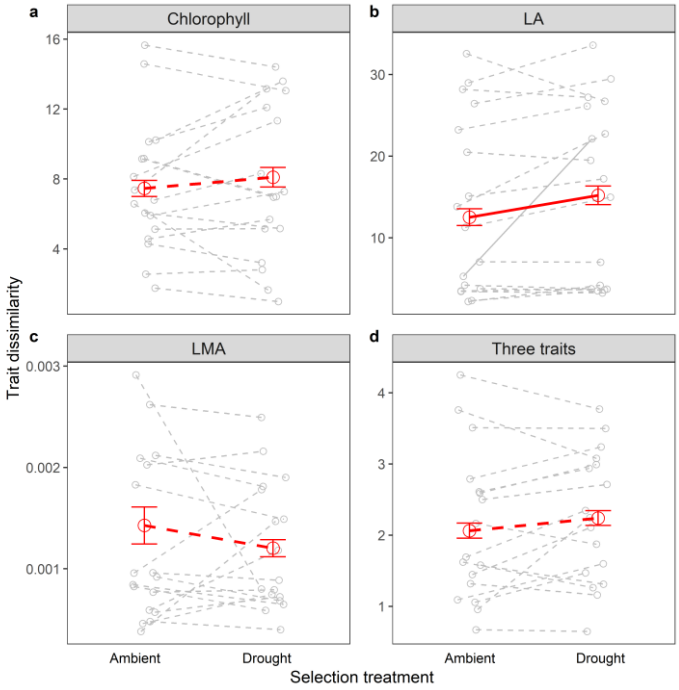
Supplementary Figure 4 | Effects of selection treatment (ambient- vs. drought-selected plants) on trait values (leaf relative chlorophyll content [a], leaf area [b, LA], leaf mass per area [c, LMA] and leaf osmometric potential [d, OSMO]) measured on individual plants before the drought event in the glasshouse. Red points and error bars show means $\pm$ standard error of all individual plants. Grey points represent means for species (standard errors for species were not shown). Grey lines connected the same species between the selection treatments. Solid lines represent statistically significant difference ($P < 0.05$) in trait values between the two selection treatments. The statistical tests for all species together were conducted by fitting block and selection treatment as fixed-effects terms, species composition and its interaction with selection treatment as random-effects terms in mixed-effects models ($P > 0.10$; results not shown). The statistical tests for each species were conducted by fitting block and selection treatment in general linear models. Note that the values of LA, LMA and OSMO were log-transformed.



Supplementary Figure 5 | Effects of selection treatment (ambient- vs. drought-selected plants) on leaf stomatal conductance measured on individual plants before (a) and during (b) the drought event in the glasshouse. Red points and error bars show means ± standard error of all individual plants. Grey points represent means for species (standard errors for species were not shown). Grey lines connected the same species between the selection treatments. Dashed lines represent statistically insignificant difference ($P > 0.10$) in trait values between the two selection treatments. The statistical tests for all species together were conducted by fitting selection treatment as fixed-effects terms, species composition and its interaction with selection treatment as random-effects terms in mixed-effects models (results not shown). The statistical tests for each species were conducted by fitting selection treatment in general linear models. Block was additionally included as fixed-effect term in panel a.



Supplementary Figure 6 | Effects of selection treatment (ambient- vs. drought-selected plants) on biomass ratio between root and shoot measured on individual plants after the drought event in the glasshouse. Red points and error bars show means $\pm$ standard error of all individual plants. Grey points represent means for species (standard errors for species were not shown). Grey lines connected the same species between the selection treatments. Dashed lines represent statistically insignificant difference ($P > 0.05$) in trait values between the two selection treatments. The statistical tests for all species together were conducted by fitting block and selection treatment as fixed-effects terms, species composition and its interaction with selection treatment as random-effects terms in mixed-effects models (results not shown). The statistical tests for each species were conducted by fitting block and selection treatment in general linear models.



Supplementary Figure 7 | Effects of selection treatment (ambient- vs. drought-selected plants) on trait dissimilarity between interacting species within 2-speices mixtures before the drought event in the glasshouse. Trait dissimilarities were measured for leaf relative chlorophyll content (a), leaf area (b, LA), leaf mass per area (c, LMA) and the three traits together (d). Red points and error bars show means ± standard error of all mixtures. Grey points represent means for species pair (standard errors for species pairs were not shown). Grey lines connected the same species pair between the selection treatments. The solid red line indicates a marginally significant difference (averaged across species pairs) between the two selection treatments ($P < 0.10$) from a mixed-effects model, in which block and selection treatment were set fixed-effects terms, species composition and its interaction with selection treatment were set as random-effects terms (Supplementary Table 10). The solid grey line indicates a significant difference for a specific species pair between the two selection treatments ($P < 0.05$) from a general linear model, in which block and selection treatment were set as fixed-effects terms.