PROCEEDINGS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Shifts and disruptions in resource-use trait syndromes during the evolution of herbaceous crops

Rubén Milla, Javier Morente-López, J. Miguel Alonso-Rodrigo, Nieves Martín-Robles and F. Stuart Chapin III

Proc. R. Soc. B 2014 **281**, 20141429, published 3 September 2014

PROCEEDINGS OF THE ROYAL SOCIETY B

rspb.royalsocietypublishing.org

Research

Cite this article: Milla R, Morente-López J, Alonso-Rodrigo JM, Martín-Robles N, Stuart Chapin III F. 2014 Shifts and disruptions in resource-use trait syndromes during the evolution of herbaceous crops. *Proc. R. Soc. B* **281**: 20141429.
http://dx.doi.org/10.1098/rspb.2014.1429

Received: 11 June 2014
Accepted: 6 August 2014

Subject Areas:
ecology, plant science

Keywords:
breeding, competition, domestication, phenotypic integration, plant ecological strategies, structural equation models

Author for correspondence:
Rubén Milla
e-mail: ruben.milla@gmail.com

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2014.1429 or via http://rspb.royalsocietypublishing.org.

Royal Society Publishing

Shifts and disruptions in resource-use trait syndromes during the evolution of herbaceous crops

Rubén Milla[1], Javier Morente-López[1], J. Miguel Alonso-Rodrigo[1], Nieves Martín-Robles[1] and F. Stuart Chapin III[2]

[1]Departamento de Biología y Geología, Área de Biodiversidad y Conservación, Escuela Superior de Ciencias Experimentales y Tecnología, Universidad Rey Juan Carlos, c/Tulipán s/n, Móstoles 28933, Spain
[2]Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775, USA

Trait-based ecology predicts that evolution in high-resource agricultural environments should select for suites of traits that enable fast resource acquisition and rapid canopy closure. However, crop breeding targets specific agronomic attributes rather than broad trait syndromes. Breeding for specific traits, together with evolution in high-resource environments, might lead to reduced phenotypic integration, according to predictions from the ecological literature. We provide the first comprehensive test of these hypotheses, based on a trait-screening programme of 30 herbaceous crops and their wild progenitors. During crop evolution plants became larger, which enabled them to compete more effectively for light, but they had poorly integrated phenotypes. In a subset of six herbaceous crop species investigated in greater depth, competitiveness for light increased during early plant domestication, whereas diminished phenotypic integration occurred later during crop improvement. Mass-specific leaf and root traits relevant to resource-use strategies (e.g. specific leaf area or tissue density of fine roots) changed during crop evolution, but in diverse and contrasting directions and magnitudes, depending on the crop species. Reductions in phenotypic integration and overinvestment in traits involved in competition for light may affect the chances of upgrading modern herbaceous crops to face current climatic and food security challenges.

1. Introduction

Plant domestication has had far-reaching evolutionary and ecological consequences for both plants and people [1,2]. On the plant side, conscious and unconscious human selection led to the evolution of new phenotypes. More palatable plant organs with fewer secondary compounds [3], larger and/or more numerous fruits or seeds [4], or modifications in vegetative aboveground morphology [5] are examples of allelic variants favourable to human purposes that were selected for in most domestication processes [5]. Other relevant traits also emerged during domestication, but in variable ways, depending on the crop species or functional group. For instance, most cereals lost spontaneous seed shattering, and many other herbaceous crops reduced seed dormancy mechanisms, while woody fruit crops increased allocation to pericarp tissues [6].

In a parallel literature, though largely independently, ecological science found that selective filters operating in the wild consistently caused plants from unproductive infertile ecosystems to differ from those that evolved in more productive sites in a number of plant traits functionally linked to the use of mineral, water and light resources [7,8]. Plant strategies are groups of species that share traits whose variation impacts fitness differently in ecosystems with contrasting availability of resources [9]. Plant strategies, thus considered as sets of functionally interlinked traits [9], are closely connected with the concept of phenotypic integration. Phenotypic integration is the phenotypic covariance structure of multiple characters that bear functional links [10]. The traits of organisms with high phenotypic integration respond in coordinated fashion to changes in the

environment, in contrast to poorly integrated organisms [11]. Plant strategies and phenotypic integration literatures provide specific predictions as to the direction of phenotypic evolution in habitats with contrasting levels of resource supply, which may correspond to the shift in environmental pressures that plants experienced when humans began to raise them in croplands.

Initial domestication and further crop improvement moved plant species from wild habitats, where resources were relatively scarce, to cultivated land, where water and nutrients were generally more predictably available [12,13]. Plant strategy theory would predict a convergence of evolutionary trajectories during domestication towards plant strategies that acquire water, nutrients and light rapidly to support fast growth at the expense of less efficient nutrient retention [7,8]. If this were the case, traits indicative of effective hoarding of light and carbon, and rapid below-ground resource acquisition and growth, would be more characteristic of current crops than of their wild progenitors. Additionally, ecologists have postulated that abiotically less stressful conditions should relax selection for tight coordination among functionally related traits, resulting in plants with less phenotypic integration [10,14]. That postulate, joined to the fact that selection and breeding have focused on a few beneficial traits, rather than on the complete phenotype, might have reduced phenotypic integration during crop evolution. The concept of phenotypic integration is broadly applied from the physiological to the ecological and evolutionary scales, and from coordination of narrowly related functions and structures to the integration of the complete individual [15]. For the purposes of this work, we consider phenotypic integration as the tendency of functionally linked traits of the whole phenotype to evolve in a coordinated fashion after an evolutionary divergence. The focal divergence here is the split of crop genotypes from their wild progenitors during plant domestication.

In this paper, we test (i) whether high-resource-use strategies have evolved during the transition from wild progenitors to domesticated forms of herbaceous crop species, and (ii) whether resource-use traits have evolved in a coordinated fashion during domestication and further crop improvement. We pose those questions to a uniquely large set of 30 diverse crop–progenitor comparisons (table 1; electronic supplementary material S1 and S2, and figures S1 and S2). We further assess, in a smaller subset of six herbaceous crop species, which traits and covariation patterns changed during the early evolution of landraces versus the more intensive later breeding and development of modern crop varieties.

2. Material and methods

(a) Study system

We studied the process of domestication in 30 herbaceous crop species important to human food supply (table 1). These include a diverse array of phylogenetically and functionally different crops with distinct domestication geographies and histories (electronic supplementary material, figures S1 and S2), the ample majority of them being annual species. In an extensive experiment, we compared two accessions for each of the 30 crops: a crop cultivar and a related wild species known to have made the greatest contribution to the gene pool of the crop (hereafter termed 'progenitor'; electronic supplementary material, table S1). In a second intensive experiment, we compared nine accessions for six of these 30 crops. Three of these nine accessions were geographically diverse provenances of the putative wild progenitor. Three others were landraces, representative of an initial stage of domestication. The final three were commercial varieties that have undergone modern breeding improvement programmes. The accessions of the intensive experiment were selected to include a broad range of geographical wild provenances (wild), of ethnographically and geographically diverse landraces (landrace) and of varietal diversity for modern crops (improved). The six crop species chosen for the intensive experiment, based on their taxonomic and functional diversity and agronomic relevance, were maize, barley, pea, pepper, sunflower and collard. See the electronic supplementary material, table S1 for accession identifiers, seed donors, domestication status and literature source for wild progenitor assignment of the extensive experiment, and electronic supplementary material, table S2 for the same information for the intensive experiment.

(b) Plant growth and trait measurements

We adopted a common garden approach to document phenotypes of all accessions involved in this project in identical environments. Ten to 20 plants of each accession in the project were grown and measured under two experimental regimes. The first regime was devised to obtain early growth, height and leaf trait data for our 30-species extensive dataset. The second was designed to provide root measurements and total plant dry mass for both datasets, and early growth, height and leaf trait data for the six-species intensive dataset. Plant measurements were carried out to obtain individual scores of each of the following nine traits (see Full materials and methods section in electronic supplementary material for trait selection criteria): seed size (mg), total plant dry mass (g), seedling absolute growth rate (cm d^{-1}), plant canopy height (cm), leaf size (cm^2), specific leaf area ($cm^2 g^{-1}$), leaf dry matter content ($g\,g^{-1}$), specific root length ($m\,g^{-1}$) and tissue density of fine roots ($g\,cm^{-3}$). A total of 1562–2618 (depending on traits) plant individuals were phenotyped. See Full materials and methods section in the electronic supplementary material for detailed protocols and procedures.

(c) Statistical analyses

Below we describe the several statistical procedures that we followed to test for the effects of domestication on individual and grouped trait scores, and to test for shifts in inter-trait relationships and phenotypic integration during the evolution of crops. Analytical details of each of the procedures, in-depth explanations for the rationale of the construction of structural equation models and motivation for the choice of Aster models to test for different levels of phenotypic integration are provided in the Full materials and methods section in the electronic supplementary material. Additional supporting analyses are also provided in the electronic supplementary material (items 4 and 5).

(i) Effects of domestication status and crop identity on traits and on groups of traits

Traits were considered separately for analyses, and also in groups of two or three, according to well-known strong physiological or developmental linkages among traits. Grouping of traits for further data analysis was performed through factor analyses. Four principal components analyses (PCAs) were run separately for each of the following four groups of log-scaled traits. First, seed size and total dry mass data were reduced to a first PCA axis intended to represent the organ and plant size variation among individuals (SIZE hereafter). Second, seedling absolute growth rate, plant canopy height and leaf size were reduced to a single PCA axis aimed to synthesize competitive ability for light capture (LIGHT COMP hereafter). Third, a PCA was run on leaf economic traits using specific leaf area and leaf dry matter content (LEAF ECON

Table 1. Common and botanical name of each of the 30 crop species of this project, together with that of its assigned wild progenitor and family affiliation. Species in bold were those investigated in greater depth in the intensive experiment. See the electronic supplementary material, table S1 for more detailed information, particularly on bibliographic references used for assigning wild progenitors for each crop.

crop	wild ancestor	family	common name
Avena sativa	*A. sterilis*	Poaceae	oats
Beta vulgaris	*B. vulgaris*	Amaranthaceae	chard
Brassica oleracea	***B. oleracea***	**Brassicaceae**	**collard**
Capsicum anuum	***C. anuum***	**Solanaceae**	**pepper**
Capsicum bacattum	*C. bacattum*	Solanaceae	chili pepper
Cicer arietinum	*C. reticulatum*	Fabaceae	chickpea
Cichorium endibia	*C. intybus*	Asteraceae	chicory
Cynara cardunculus	*C. cardunculus*	Asteraceae	cardoon
Eruca sativa	*E. sativa*	Brassicaceae	rocket
Glycine max	*G. soja*	Fabaceae	soyabean
Gossypium hirsutum	*G. hirsutum*	Malvaceae	cotton
Helianthus annuus	***H. annuus***	**Asteraceae**	**sunflower**
Hordeum vulgare	***H. spontaneum***	**Poaceae**	**barley**
Lathyrus sativus	*L. cicera*	Fabaceae	chickling vetch
Lens culinaris	*L. orientalis*	Fabaceae	lens
Lupinus luteus	*L. luteus*	Fabaceae	lupins
Medicago lupulina	*M. lupulina*	Fabaceae	black medic
Oryza sativa	*O. nivara*	Poaceae	rice
Pennisetum glaucum	*P. glaucum*	Poaceae	millet
Pisum sativum	***P. humile***	**Fabaceae**	**peas**
Secale cereale	*S. ancestrale*	Poaceae	rye
Sesamum indicum	*S. indicum*	Pedaliaceae	sesame
Solanum lycopersicon	*S. pimpinellifolium*	Solanaceae	tomato
Sorghum sudanense	*S. bicolor*	Poaceae	sorghum
Spinacea oleracea	*S. turkestanica*	Amaranthaceae	spinach
Trifolium repens	*T. repens*	Fabaceae	white clover
Triticum durum	*T. dicoccoides*	Poaceae	wheat
Vicia faba	*V. narbonensis*	Fabaceae	fava bean
Vigna unguiculata	*V. unguiculata*	Fabaceae	cowpea
Zea mays	***Z. mays***	**Poaceae**	**maize**

hereafter). The first axis of that PCA was negatively related to structural investment in leaf tissue, which commonly correlates positively with leaf longevity and negatively with carbon fixation rates and mass-based nutrient status of leaves. Lastly, a fourth PCA was run on root economic traits with specific root length and density of fine roots as component variables (ROOT ECON hereafter). First axes of the PCA analyses above explained 64–82% of variance in their component variables (see the electronic supplementary material, table S3) and were thus used as summary proxies of each of the four plant functions considered in this paper. Parameter-free PERMANOVA analyses were then used to assess the effects of domestication status, crop identity and their interaction on trait scores and on PCA axes (see Full materials and methods section in the electronic supplementary material for details).

(ii) Coordinated evolution of traits during domestication and further improvement

First, structural equation modelling (SEM) was used to investigate functional links among multiple traits, and their putative coordinated evolution during domestication [16]. Here, we briefly describe the datasets we used to build the several SEMs. In the electronic supplementary material (Full materials and methods section), we also explain the rationale for model construction and provide statistical details on the estimation of goodness of fit and on the statistical significance of model parameters. Finally, in the electronic supplementary material (item 4), we provide results of additional SEM analyses intended to provide further empirical support for the general validity of the *a priori* inter-trait relationship scheme depicted in figure 2*a*.

SEMs were implemented for two separate types of datasets. First, we put together all log-scaled arithmetic mean trait and PCA scores for each accession present in both the extensive and intensive databases ('complete' dataset hereafter; $n = 114$ accessions). Second, we calculated, independently for the extensive and intensive databases, the magnitude of the domestication status effect over log-trait and PCA scores as follows. For the extensive database, we subtracted the average score of each wild (W) accession from that of its crop (C) counterpart. This evolutionary change is denoted as $\Delta C - W_{\mathrm{Trait}}$ throughout the paper. For the

intensive database, we calculated an effect of domestication and an effect of subsequent improvement separately. The domestication effect was taken by subtracting the average score of each wild (WI) accession from that of its landrace (LR) counterpart ($\Delta LR - WI_{Trait}$ hereafter). Since the intensive database included three accessions for each domestication status per crop species, we subtracted every possible combination of wild accession from every landrace for each species separately. This is analogous to common practices in phenotypic plasticity literature making use of relative distance plasticity indices [17]. This procedure yielded nine $\Delta LR - WI_{Trait}$ scores per crop species included in the intensive database. We proceeded in the same way to compute the improvement effect: every score of a landrace (LR) accession was subtracted from each of the three improved (IM) accessions available for each crop species ($\Delta IM - LR_{Trait}$ hereafter).

The above calculations resulted in four separate datasets: (i) a 'complete' dataset ($n = 114$ log-scaled accession average log-trait and PCA scores); (ii) '$\Delta C - W_{Trait}$' dataset ($n = 30$ wild-to-crop evolutionary transitions for each trait and PCA score); (iii) '$\Delta LR - WI_{Trait}$' dataset ($n = 54$ wild-to-landrace evolutionary transitions for each trait and PCA score); (iv) '$\Delta IM - LR_{Trait}$' dataset ($n = 54$ landrace-to-improved evolutionary transitions for each trait and PCA score).

After that, based on previous knowledge and on the patterns of trait correlations that we observed in our study, we designed an overall causal conceptual structure that linked our four groups of variables in a model (see Full materials and methods section in the electronic supplementary material for details). Considering the conceptual model, we generated several tentative specific models with our empirical variables, and the specific model receiving the highest statistical support for our 'complete' dataset was selected as a baseline model capturing common covariation patterns among the several traits involved in this study. This model was further validated with an independent dataset from the literature (electronic supplementary material, item 4). Then, the '$\Delta C - W_{Trait}$', '$\Delta LR - WI_{Trait}$' and '$\Delta IM - LR_{Trait}$' datasets, representing trait shifts during crop evolution, were fitted to the baseline model (see Full materials and methods in electronic supplementary material for details).

Finally, statistical comparison of phenotypic integration levels among datasets in figure 4 was carried out using log-likelihood ratio tests of Aster models, robust to bimodality of dependent variables (see Full materials and methods and table S8 in the electronic supplementary material).

3. Results

Seed and plant sizes generally increased with crop evolution, as did proxies for light competitive ability such as leaf size, plant canopy height and early absolute growth rates of seedlings (figure 1*a*; electronic supplementary material, figure S3). Traits related to resource-acquisition rates of leaves and roots, such as dry mass content of leaves or specific root length, changed in diverse directions, depending on the crop species (figure 1*a*; electronic supplementary material, S3). For six crop species, we dissected the crop evolution process into an early domestication and a later improvement stage. Early domesticates, represented by landraces, were larger plants and more effective light competitors, but were similar to progenitors in their leaf and root resource-acquisition traits (figure 1*b*; electronic supplementary material, S4). Modern cultivars, however, did not differ consistently from comparable landraces in any measured trait (figure 1*c*; electronic supplementary material, S5). For all traits and evolutionary comparisons, we found significant domestication status × crop identity effects (electronic supplementary material, tables S5–S7), signalling crop specificity (see the electronic supplementary material, figures S3–S5).

To analyse patterns of coordinated evolution of traits we first devised a model of inter-trait relationships, based on previous knowledge (electronic supplementary material, Full materials and methods section; figure 2*a*). Goodness of fit of the model to the data collected on all accessions was high (figure 2*b*), validation with independent datasets yielded remarkably similar covariance structures (electronic supplementary material, Full materials and methods section), and inter-trait correlations were generally strong. This supports the validity of our *a priori* model. We then used this baseline trait-syndrome model to investigate interlinked shifts in plant traits in response to initial domestication and to crop improvement. Note that path or correlations coefficients $< |0.05|$ were constrained to be zero to achieve model identification in all SEM models (e.g. $SIZE_{PCA} \rightarrow ROOT\ ECON_{PCA}$ in figure 2*b*).

The general strength and pattern of inter-trait coordination weakened when the *differences in trait scores* between wild and crop accessions were the underlying measured variables, rather than the values of traits themselves. This means that a given change in a trait or group of traits (e.g. increased seed size) did not necessarily entail a parallel shift in otherwise inter-connected traits (e.g. increase in seedling absolute growth rate; figure 3*a*). Note that *differences in trait scores* were widely different among traits and crop species (see Supplementary phenotypic integration analyses section in the electronic supplementary material), so we reject homogeneity of $\Delta C - W_{Trait}$ scores as the reason for the weakening of coordinated coevolution of traits. This weakening of trait correlation was supported by formal analyses of phenotypic integration (figure 4). Coordinated evolution of traits (i.e. integration), measured jointly as the magnitude and significance of Spearman's correlation coefficients between plant traits, was low when transiting from wild progenitors to domesticated forms of the 30 crops studied here (figure 4*b*).

We then looked at coordinated evolution in more detail for our subset of six crop species. We found that early domestication retained several of the strong trait-to-trait relationships that we had previously identified from literature and our own data (figure 3*b*). This includes allometric relationships between size traits and ability to compete for light and coordinated evolution of leaf and root economic traits (figure 3*b*). However, later crop improvement reduced covariation among most of these traits (figure 3*c*). Direct comparison of trait covariance matrices between early domestication and later breeding confirmed that changes occurring during more recent crop improvement took place with little coordination among traits (figure 4*c*–*d*).

4. Discussion

We found that, despite multiple crop-specific peculiarities, during crop evolution plants changed towards larger and more effective competitors for light, with poorly integrated phenotypes, in our set of herbaceous crop species. Enhanced competitive ability for light occurred during early domestication, whereas decreased phenotypic integration took place later, during recent crop improvement. No consistent phenotypic reaction to domestication was detected as regards leaf and root resource economy traits. These patterns have several important implications. First, increased size and capacity to

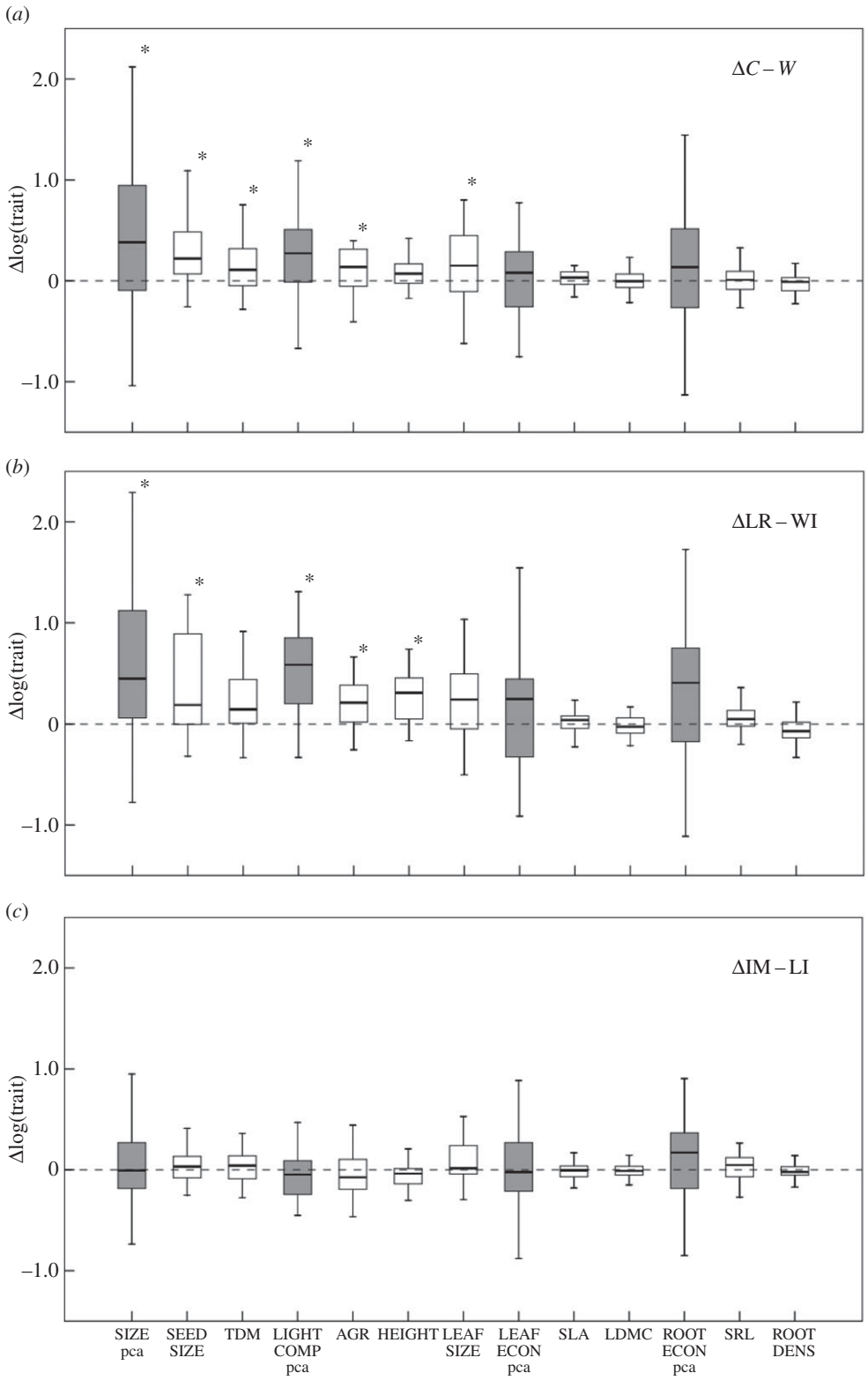


Figure 1. Shifts in traits during crop evolution. Direction and magnitude of (log)trait divergences (*a*) during the evolution of 30 crop species, (*b*) during early domestication of six species and (*c*) during further improvement of the same six species for the nine traits studied (white boxes), and for the four PCA axes summarizing trait covariation in figures 2 and 3 (grey boxes). Reference line: crop = wild (*a*); landrace = wild (*b*); or improved = landrace (*c*). Seed size (seed mass), TDM (total plant dry mass), AGR (absolute growth rate of seedlings), HEIGHT (canopy height), LEAF SIZE (projected leaf area), SLA (leaf area per leaf dry mass), LDMC (leaf dry matter content), SRL (root length per unit root mass), ROOT DENS (root tissue density). Crop identities were collapsed for clarity (crop-wise plots in the electronic supplementary material, figures S3–S5). Asterisks, domestication status significant at $p = 0.05$.

compete effectively with weeds for light is an evolutionary trajectory expected in high-resource environments [7,18]. Aggressive light competitors, however, divert resources to heterotrophic tissue (e.g. stems) for outcompeting neighbours [18]. This comes at the cost of diminishing community-level productivity and increasing risks of lodging, which are undesirable in agricultural stands [12,19,20]. Consequently, several of the major grain cereal crops were bred during the Green Revolution for shorter, more productive plants [12]. This could account for the lack of increased light competitiveness during the latest improvement stage of herbaceous crop evolution, at least for barley (figure 1*c*). In contrast to our

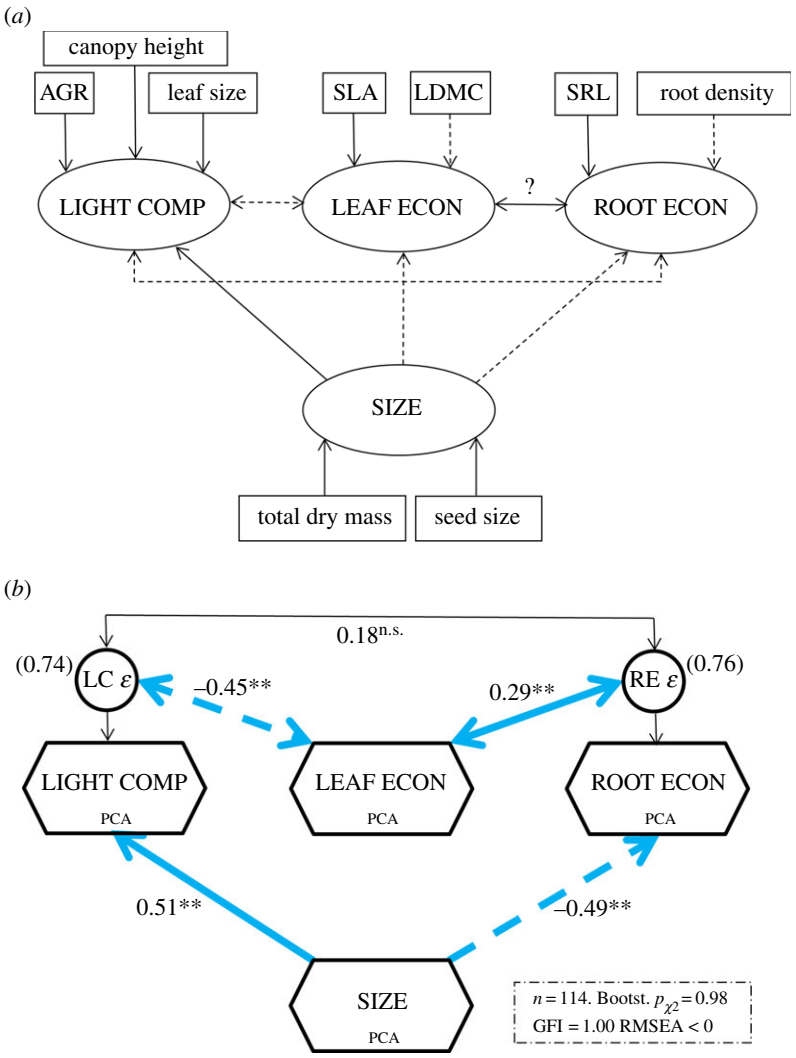


Figure 2. Baseline structural equation model for inter-trait relationships. (*a*) Conceptual *a priori* model. Solid and dashed lines are positive and negative relationships, respectively. Double-headed arrows denote uncertain direction of causality. Question marks denote uncertain sign (+ or −) of the relationship. SIZE, LIGHT COMP, LEAF ECON and ROOT ECON signify plant functions. SIZE increases with larger seeds and larger plants. LIGHT COMP increases with increasing leaf size, plant height and aboveground growth rates. LEAF ECON and ROOT ECON are larger for softer and photosynthetically faster leaves and roots with faster uptake rates, respectively. (*b*) Fit of the current paper's dataset to the *a priori* model. Statistically significant paths (→) or correlations (↔) in bold and denoted as $^{**}p < 0.01$ or $^{*}p < 0.05$. *Trait* ε = unexplained variance of dependent variables. Standardized path or correlations coefficients $< |0.05|$ (e.g. between SIZE and LEAF ECON) were constrained to be zero to achieve identification. Names of variables as in figure 1. (Online version in colour.)

initial hypothesis and the predictions of ecological strategies theory [7], leaf and root traits did not signal a generalized move towards a high-resource-use strategy during crop evolution. This result supports early literature noting absence of a systematic increase in area-based rates of photosynthesis during the evolution of cereals and other major herbaceous crops [21]. In addition, the few previous studies investigating shifts in specific leaf area accompanying crop evolution do not support any consistent pattern of change in this trait under domestication [22–24]. Wild progenitors of current herbaceous crops were probably already high-resource strategists in terms of biomass renewal and resource-acquisition rates. A look at specific leaf area scores of our species in the context of worldwide variation in that same trait supports this view (electronic supplementary material, figure S6). Fostering increases in an already fast strategy might provide diminishing physiological returns [25]. Further archaeo-ecophysiological work, however, is required to test the idea that wild progenitors were faster resource strategists than phylogenetically comparable co-occurring species (but see [26]).

Augmented size and competitive ability for light were not correlated with changes in proxies for leaf and root resource-use strategies (figures 1*a* and 3*a*). This is most interesting, and is at odds with previous theoretical and empirical literature highlighting metabolic costs of increases in body size or in heterotrophic tissue [27,28] (but see [29]). Note that decreased levels of coordinated coevolution among those traits occurred mainly during the improvement stage of herbaceous crop evolution, for the subset of six crops investigated more intensively (figure 4*d*). This suggests that historically recent breeding might explain the decrease in the degree of coordinated evolution of traits (i.e. phenotypic integration) observed for the larger 30-species dataset. During the improvement phase of herbaceous crops, a combination of enhanced understanding of genetics of specific traits and greater human control over the agro-environment may have allowed selection of new trait combinations, which neither succeeded in the wild nor during earlier plant domestication.

Phenotypic integration may arise through strong directional selection, through genetic linkages that promote joint

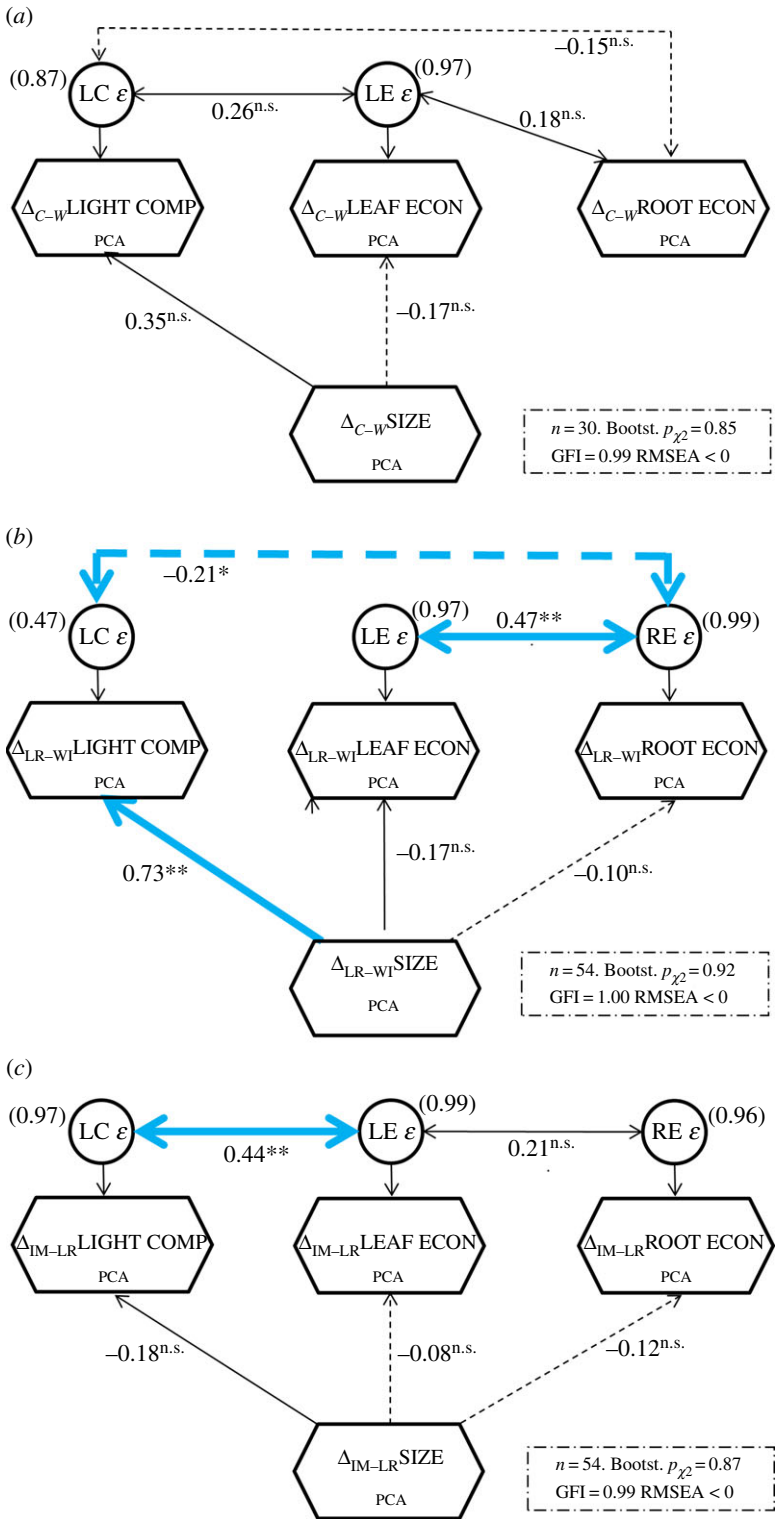


Figure 3. Shifts in plant trait relationships during crop evolution. Fit of the structural equation model in figure 2*b* to (*a*) the magnitude of trait shifts between crop and wild accessions of the 30-crop-species set ('Δ_{C-W}trait'), (*b*) the same shift in traits between wild and landrace accessions during early domestication of six crops ('Δ_{LR-WI}trait'), and (*c*) the same shift in traits during later improvement for the same six crops ('Δ_{IM-LR}trait'). Symbols and variables names as in figures 1 and 2. (Online version in colour.)

inheritance of traits, through selective sweeps and through genes with pleiotropic effects exerting biochemical control (e.g. response of many traits to abscisic acid) [28,29]. Disruption of genetic and some biochemical linkages among traits offers opportunities to develop varieties with new combinations of desired traits through breeding and genetic engineering. Indeed, that has been the case for the evolution and breeding of many crops, with obvious beneficial effects for humans. What disadvantages, if any, might then result from loss of coordinated evolution of functional traits during recent crop improvement? Phenotypic integration is promoted by environmental stresses (e.g. drought) that directly select for groups of traits with functional linkage (e.g. low minor vein densities and low maximum rates of photosynthesis) [30]. Previous work has documented increases in phenotypic integration in response to increasing environmental stress [10,14,31]. That literature, however, primarily focused on the plastic response of phenotypes to changes in the environment, rather than the

(*a*) traits

	SEED SIZE	TDM	AGR	HEIGHT	LEAF SIZE	SLA	LDMC	SRL
SEED SIZE								
TDM	**0.58**							
AGR	**0.61**	**0.39**						
HEIGHT	**0.26**	**0.38**	**0.60**					
LEAF SIZE	0.00	**0.37**	**0.32**	**0.68**				
SLA	**0.26**	0.00	0.00	**0.48**	**0.52**			
LDMC	0.00	0.00	0.00	**0.38**	**0.27**	**0.68**		
SRL	**0.61**	**0.54**	**0.30**	0.00	0.00	0.00	**0.21**	
ROOT DENS	**0.21**	**0.22**	0.00	0.00	0.00	**0.20**	**0.37**	**0.65**

(*b*) Δ_{C-W}

	SEED SIZE	TDM	AGR	HEIGHT	LEAF SIZE	SLA	LDMC	SRL
SEED SIZE								
TDM	**0.61**							
AGR	0.00	0.00						
HEIGHT	**0.39**	**0.49**	**0.38**					
LEAF SIZE	**0.47**	0.00	0.00	0.00				
SLA	0.00	0.00	0.00	0.00	0.00			
LDMC	0.00	0.00	0.00	0.00	0.00	**0.53**		
SRL	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
ROOT DENS	0.00	0.00	0.00	0.00	0.00	0.00	0.00	**0.37**

(*c*) Δ_{LR-WI}

	SEED SIZE	TDM	AGR	HEIGHT	LEAF SIZE	SLA	LDMC	SRL
SEED SIZE								
TDM	**0.67**							
AGR	**0.30**	0.00						
HEIGHT	**0.68**	**0.56**	**0.62**					
LEAF SIZE	**0.56**	**0.68**	0.00	**0.32**				
SLA	0.00	0.00	**0.28**	0.00	**0.36**			
LDMC	**0.47**	**0.31**	0.00	0.00	**0.61**	**0.46**		
SRL	**0.41**	**0.30**	**0.42**	**0.57**	0.00	0.00	0.00	
ROOT DENS	**0.37**	0.00	**0.33**	0.00	**0.34**	0.00	**0.58**	**0.33**

(*d*) Δ_{IM-LR}

	SEED SIZE	TDM	AGR	HEIGHT	LEAF SIZE	SLA	LDMC	SRL
SEED SIZE								
TDM	**0.35**							
AGR	0.00	0.00						
HEIGHT	0.00	0.00	**0.76**					
LEAF SIZE	**0.36**	**0.76**	0.00	0.00				
SLA	0.00	0.00	0.00	0.00	**0.29**			
LDMC	0.00	0.00	**0.28**	0.00	0.00	**0.56**		
SRL	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
ROOT DENS	0.00	0.00	0.00	0.00	0.00	**0.32**	**0.48**	**0.62**

Figure 4. Decrease in phenotypic integration with crop evolution. Matrices of Spearman's correlation coefficients among the nine log-transformed traits measured in this study (figure 1). When a correlation was non-significant, the coefficient is statistically non-different from zero and was thus denoted as 0.00. Significant correlations are highlighted in bold and in grey background. (*a*) Correlation coefficients for the current paper's dataset of all accessions. (*b–d*) Correlation coefficients between shifts in traits (*b*) during crop evolution of the whole set of 30 crops, (*c*) due to early domestication of six crops and (*d*) during later improvement for the same six crops. Different lowercase letters (a, b) among matrices in the same row indicate significant differences in average correlation coefficient scores.

macroevolutionary changes reported here [11]. Although studies of coordinated plant trait macroevolution are abundant [32,33], phenotypic integration levels are rarely explicitly addressed in that research. The fact that annual herbs not only acclimate, but also evolve varying levels of inter-trait linkages as a consequence of crop domestication and breeding, is remarkable. Evolutionary disruption of those linkages might result in varieties that perform effectively in highly controlled environments such as high-input agriculture, but may be more vulnerable to stressful conditions that are expected with climatic change. For instance, ongoing increased aridity is perceived as a relevant threat for the marketability of *Brassica oleracea* L. var acephala (cabbage), a major European crop [34]. Cabbage has shown outstanding evolutionary plasticity in food quality characteristics. However, the fact that traits that promote an efficient use of water and mineral resources did not improve during the evolution of this crop (electronic supplementary material, figures S3–S5) may pose limitations on its widespread use in the near future, particularly in southern European countries. More detailed and crop-specific research is needed to understand the balance between benefits and disadvantages, in terms of yield or plant fitness, of diverse levels of phenotypic integration and of plasticity in crop species grown in contrasting environments.

In short, our results show that evolution of herbs under agricultural selective pressures resulted in phenotypes that are larger and more aggressive competitors for light than their wild progenitors, but similar in terms of mass-specific capabilities for root and leaf resource uptake and conservation. The capacity to compete for light was selected during initial domestication, but plants did not grow larger during recent crop improvement. More interestingly, we report that phenotypes did not change in a coordinated manner during crop evolution for traits that are otherwise tightly interrelated in wild habitats. Such decrease in phenotypic integration occurred in recent times, during breeding of modern commercial varieties.

We suggest that the set of commonly observed traits that accompany domestication should be extended to include traits of direct relevance to plant fitness in the wild. This may contribute to our knowledge of how herbaceous crop evolution proceeded in the past and to our opportunities to re-direct or emphasize those trajectories. Most of the common traits associated with domestication of annual herbs are chiefly made of traits that were selected for their agronomic benefit to humans [4]. Here, we propose that evolution of more effective light competitors should be considered a generalized consequence of herbaceous crop domestication, jointly with other traits previously identified [5]. Additionally, decreases in the strength of inter-trait linkages during crop evolution fit expectations from phenotypic integration literature for evolution under benign environments [10,14,31]. Potential implications of this remarkable result for augmenting crop yields, and for promoting yield sustainability and adaptability to stress factors in a changing climate, should prompt further research in this area.

Acknowledgements. We thank Jose Margalet, Mónica Álvaro, Tania de la Fuente and Carlos Díaz for assistance in data gathering, Martha Raynolds and Samuel Doshevsky for logistic support, Luis Balaguer for ideas on phenotypic integration, and two anonymous referees for useful comments on an earlier version of the manuscript.

Data accessibility. All data used in this manuscript are deposited in Dryad repository (doi:10.5061/dryad.dg85v).

Funding statement. We thank Universidad Rey Juan Carlos for funding a research stay of R.M. in the University of Alaska (Ayudas a la Movilidad grants, 2013), and Madrid Regional Government (grant REMEDINAL-2) and MINECO-Spain (grants nos. AGL2010-10935-E and CGL2011-28778) for project funding.

References

1. Hancock J. 2004 *Plant evolution and the origin of crop species*. Cambridge, MA: CABI Publishing.
2. Kingsbury N. 2009 *Hybrid: the history and science of plant breeding*. Chicago, IL: University of Chicago Press.
3. Magnussen LS, Hauser TP. 2007 Hybrids between cultivated and wild carrots in natural populations in Denmark. *Heredity* (*Edinb*). **99**, 185–192. (doi:10.1038/sj.hdy.6800982)
4. Doebley J. 2006 Unfallen grains: how ancient farmers turned weeds into crops exploring other worlds to learn. *Science* **312**, 1318–1319. (doi:10.1126/science.1128836)
5. Meyer RS, DuVal AE, Jensen HR. 2012 Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. *New Phytol.* **196**, 29–48. (doi:10.1111/j.1469-8137.2012.04253.x)
6. Meyer RS, Purugganan M. 2013 Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* **14**, 840–852. (doi:10.1038/nrg3605)
7. Craine JM. 2009 *Resource strategies of wild plants*. Princeton, NJ: Princeton University Press.
8. Lambers H, Poorter H. 1992 Inherent variation in growth rate between higher plants: a search for physiological causes and ecological consequences. *Adv. Ecol. Res.* **22**, 187–261. (doi:10.1016/S0065-2504(08)60148-8)
9. Grime JP, Thompson K, Hunt R, Hodgson JG, Cornelissen JHC, Rorison IH. 1997 Integrated screening validates primary axes of specialization in plants. *Oikos* **79**, 259–281. (doi:10.10.2307/3546011)
10. Schlichting CD. 1989 Phenotypic integration and environmental change. *Bioscience* **39**, 460–464. (doi:10.2307/1311138)
11. Pigliucci M. 2003 Phenotypic integration: studying the ecology and evolution of complex phenotypes. *Ecol. Lett.* **6**, 265–272. (doi:10.1046/j.1461-0248.2003.00428.x)
12. Denison R. 2012 *Darwinian agriculture how understanding evolution can improve agriculture*. Princeton, NJ: Princeton University Press.
13. Mckey D, Elia M, Pujol B, Duputié A. 2012 Ecological approaches to crop domestication. In *Biodiversity in agriculture: domestication, evolution and sustainability* (eds P Gepts, R Bettinger, S Brush, T Famula, P McGuire, C Qualset), pp. 377–406. Cambridge, UK: Cambridge University Press.
14. Gianoli E. 2004 Plasticity of traits and correlations in two populations of *Convolvulus arvensis* (Convolvulaceae) differing in environmental heterogeneity. *Int. J. Plant Sci.* **165**, 825–832. (doi:10.1086/422050)
15. Murren CJ. 2002 Phenotypic integration in plants. *Plant Species Biol.* **17**, 89–99. (doi:10.1046/j.1442-1984.2002.00079.x)
16. Shipley B. 2002 *Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference*. Cambridge, UK: Cambridge University Press.
17. Valladares F, Sanchez-Gomez D, Zavala MA. 2006 Quantitative estimation of phenotypic plasticity: bridging the gap between the evolutionary concept and its ecological applications. *J. Ecol.* **94**, 1103–1116. (doi:10.1111/j.1365-2745.2006.01176.x)
18. Falster DS, Westoby M. 2003 Plant height and evolutionary games. *Trends Ecol. Evol.* **18**, 337–343. (doi:10.1016/S0169-5347(03)00061-2)
19. Donald C. 1968 Breeding of crop ideotypes. *Euphytica* **17**, 385–403. (doi:10.1007/BF00056241)
20. Weiner J, Andersen S, Wille WK, Griepentrog H, Olsen J. 2010 Evolutionary agroecology: the potential for cooperative, high density, weed-suppressing cereals. *Evol. Appl.* **3**, 473–479. (doi:10.1111/j.1752-4571.2010.00144.x)
21. Evans LT. 1993 *Crop evolution, adaptation and yield*. Cambridge, UK: Cambridge University Press.
22. Pujol B, Salager J-L, Beltran M, Bousquet S, McKey D. 2008 Photosynthesis and leaf structure in domesticated cassava (Euphorbiaceae) and a close wild relative: have leaf photosynthetic parameters evolved under domestication? *Biotropica* **40**, 305–312. (doi:10.1111/j.1744-7429.2007.00373.x)
23. Wacker L, Jacomet S, Körner C. 2002 Trends in biomass fractionation in wheat and barley from wild ancestors to modern cultivars. *Plant Biol.* **4**, 258–265. (doi:10.1055/s-2002-25735)
24. González A, Lynch J, Tohme J, Beebe S, Macchiavelli R. 1995 Characters related to leaf photosynthesis in wild populations and landraces of common bean. *Crop Sci.* **35**, 1468–1476. (doi:10.2135/cropsci1995.0011183X003500050034x)
25. Cornwell W, Cornelissen JHC. 2013 Commentary: a broader perspective on plant domestication and nutrient and carbon cycling. *New Phytol.* **198**, 331–333. (doi:10.1111/nph.12219)
26. Cunniff J, Wilkinson S, Charles M, Jones G, Rees M, Osborne CP. 2014 Functional traits differ between cereal crop progenitors and other wild grasses gathered in the Neolithic fertile crescent. *PLoS ONE* **9**, e87586. (doi:10.1371/journal.pone.0087586)
27. West GB, Brown JH, Enquist BJ. 1997 A general model for the origin of allometric scaling laws in biology. *Science* **276**, 122–126. (doi:10.1126/science.276.5309.122)
28. Niklas KJ. 1994 *Plant allometry: the scaling of form and process*. Chicago, IL: University of Chicago Press.
29. Price CA, Wright IJ, Ackerly DD, Niinemets Ü, Reich PB, Veneklaas EJ. In press. Are leaf functional traits 'invariant' with plant size and what is 'invariance' anyway? *Funct. Ecol.* (doi:10.1111/1365-2435.12298)
30. Brodribb TJ, Jordan GJ, Carpenter RJ. 2013 Unified changes in cell size permit coordinated leaf evolution. *New Phytol.* **199**, 559–570. (doi:10.1111/nph.12300)
31. García-Verdugo C, Granado-Yela C, Manrique E, Rubio de Casas R, Balaguer L. 2009 Phenotypic plasticity and integration across the canopy of *Olea europaea* subsp. *guanchica* (Oleaceae) in populations with different wind exposures. *Am. J. Bot.* **96**, 1454–1461. (doi:10.3732/ajb.0800420)
32. Ackerly DD, Donoghue MJ. 1998 Leaf size, sapling allometry, and Corner's rules: phylogeny and correlated evolution in maples (*Acer*). *Am. Nat.* **152**, 767–791. (doi:10.1086/286208)
33. Moles AT, Ackerly DD, Webb CO, Tweddle JC, Dickie JB, Pitman AJ, Westoby M. 2005 Factors that shape seed mass evolution. *Proc. Natl Acad. Sci. USA* **102**, 10 540–10 544. (doi:10.1073/pnas.0501473102)
34. Salisbury P, Barbetti M. 2001 Breeding oilseed *Brassica* for climate change. In *Crop adaptation to climate change* (eds S Yadav, R Redden, J Hatfield, H Lotze-Campen, A Hall), pp. 448–463. Chichester, UK: Wiley & Sons.

Electronic Supplementary Material for

Shifts and disruptions in resource-use trait syndromes during the evolution of herbaceous crops

by Milla R*, Morente-López J, Alonso-Rodrigo JM, Martín-Robles N, Chapin III FS

*To whom correspondence should be addressed. E-mail: ruben.milla@gmail.com

This file includes:

1) Full Materials and Methods

2) Supplementary Tables S1 to S8:

Table S1: Botanical name, domestication status, and seed origin information of each accession of the extensive set of 30 crop-wild ancestor pairs used in this study.

Table S2: Botanical name, domestication status, and seed origin information of each accession of the intensive set of six crop species that were studied in more detail.

Table S3: PCA loadings of each log-scaled individual trait on the first PCA axes named as SIZE, LIGHT COMP, LEAF ECON, and ROOT ECON.

Table S4: Spearman correlation coefficients matrices.

Table S5. PERMANOVA results for treatment effects and interactions (Domestication Status, Crop identity, and Dom.Status * Crop id.) on traits and PCA eigenvalues of the extensive 30 species dataset.

Table S6. PERMANOVA results for treatment effects and interactions (Domestication Status, Crop identity, and Dom.Status * Crop id.) on traits and PCA eigenvalues of part of the intensive dataset where six species where investigated in more detail.

Table S7. PERMANOVA results for treatment effects and interactions (Domestication Status, Crop identity, and Dom.Status * Crop id.) on traits and PCA eigenvalues of part of the intensive dataset where six species where investigated in more detail.

Table S8: Log-likelihood ratio tests for comparing correlation matrices of Traits vs those of Trait variation during crop evolution (“Traits” vs “$\Delta_{C\text{-}W}$Trait”); and for comparing Trait evolution during early

domestication vs Trait evolution during later improvement ("$\Delta_{LR\text{-}WI}$Trait" vs "$\Delta_{IM\text{-}LR}$Trait").

3) Supplementary Figures S1 to S6:

Figure S1. Map showing the location of the seed origin of the 48 wild ancestor accessions of this project

Figure S2: Phylogenetic diversity of the crop species of this project.

Figure S3: Bisector plots of the score of each crop on the 4 PCA axes associated with trait variation in Size (A), Competitive ability for light (B), Leaf economics (C) and Root economics (D).

Figure S4: Bisector plots of PCA scores of three wild (x-axis) and three landrace (y-axis) accessions of each of the six crop species investigated in more detail.

Figure S5: Bisector plots of PCA scores of three landrace (x-axis) and three improved (y-axis) accessions of each of the six crop species investigated in more detail.

Figure S6: Frequency distribution of Specific Leaf Area in the Glopnet database *vs* that of wild accessions of the current paper.

4) Supplementary SEM analyses

5) Supplementary phenotypic integration analyses

6) References cited in Supplementary Material

Other Supplementary Material for this manuscript includes the following:

1) Supplementary Data (milla_etal_Supplementary_Data_S1.xls):

S.D. 1a: Extensive database: Trait data for a crop and a wild ancestor accession of each of 30 crop species.

S.D. 1b: Intensive database: Trait data for three wild, three landrace, and three improved cultivar accessions of each of six crop species.

1) FULL MATERIALS AND METHODS

Study system, and collection and selection of seed material

We studied the process of domestication in 30 herbaceous crop species important to human food supply (see Table 1 in the main body of the paper for a species list). These include a diverse array of phylogenetically and functionally different crops with distinct domestication geographies and histories (Figs S1 and S2, [1]). In an extensive experiment we compared two accessions for each of the 30 crop species: a modern domesticated crop cultivar and a related wild species known to be its most likely wild ancestor. In a second intensive experiment we compared nine accessions for six of those thirty crops. Three of those nine accessions were geographically diverse provenances of the putative wild ancestor. Three others were landraces, as representatives of an initial stage of domestication. The final three were commercial varieties that have undergone modern breeding improvement programs. The accessions of the intensive experiment were selected to include a broad range of geographical wild provenances (wilds), of ethnographically diverse landraces (landraces), and of varietal diversity for modern crops (improved). The six crop species selected for the intensive experiment were maize, barley, pea, pepper, sunflower, and collard. These species were chosen for their taxonomic and functional diversity and agronomic relevance. See Table S1 for accession identifiers, seed donors, domestication status, and literature source for wild ancestor assignment of the extensive experiment, and Table S2 for the same information for the intensive experiment.

Criteria for selecting traits, experimental approach and plant measurements

We selected a set of nine plant traits to encompass four independent plant functions: competitive ability for light, leaf resource-use strategy, root resource-use strategy, and size-allometry. We adopted a soft-traits approach to select specifically measured traits [2]. Seed and whole plant dry weight were selected as proxies for whole organism and organ sizes [3]. Specific Leaf Area and Leaf Dry Matter Content were used to signify leaf resource-use strategy [4]. Specific Root Length and Fine Root Tissue Density were analogously used to signify root resource-use strategy [5]. And Absolute Growth Rate in Height of seedlings, Maximum Canopy Height, and Leaf Area were employed as proxies to competitive ability for light [6].

During 2011 and 2012 we conducted several common garden experiments to build the trait databases for the extensive and the intensive experiments described in the previous subsection. In 2011, approximately 20 seeds for each of 60 crop-wild paired accessions aimed for the extensive database were weighed to the nearest microgram with a microbalance (MT XP6, Mettler-Toledo Inc., Westerville, Ohio, USA), set to germinate on moist filter paper in dark-cold growth chambers and, when radicle emergence was observable, transplanted to individual 5x5x10 cm containers filled with commercial potting soil, and set in a greenhouse (Universidad Rey Juan Carlos, Móstoles-Madrid, Spain, 40°18´48´´N-3°52´57´´W, mean annual temperature: 14ºC, mean annual precipitation: 481 mm, long-term data from http://opengis.uab.es/wms/iberia/mms/index.htm). Seedlings were kept in the greenhouse for three to six weeks, depending on developmental speed of the crop species. Then, 15 individuals per accession were transplanted outdoors to free-rooting planting beds in an experimental field beside the greenhouse. Watering in the greenhouse and in the experimental fields was supplied at dawn and/or sunset through regular automatic water sprinkling and drip irrigation, respectively, and as needed to maintain plants under optimal growing conditions. All the above plant growth procedures were carried out sequentially through the season, matching the most appropriate time of the year for the performance of each crop species. The two accessions of each crop species were always sown concurrently and at the same spatial location within the greenhouse and planting bed, and were transplanted to the planting bed at the same time.

The plants grown in 2011 were used to obtain 5-20 (median = 15) replicate scores, for each accession of the extensive experiment, of the following traits: 1) Seed Size (mg), as described above; 2) Leaf Size, measured as one-sided projected surface area (LA, cm^2); 3) Specific Leaf Area (cm^2 g), measured as leaf surface area per oven-dried mass of leaf laminas; 4) Leaf Dry Matter Content (g g^{-1}), as the oven-dried leaf mass divided per mass of the leaf taken to full turgidity; 5) Seedling Absolute Height Growth Rate (cm d^{-1}), as the early rate of increase in height at the seedling stage; and 6) Canopy Height (cm), measured, before transplanting to planting beds, as the distance from soil surface to the highest node of the main shoot. Protocols for trait measurement described below follow [3]. Both accessions of a given crop were always measured concurrently at the same spatial location, even if timing of measurements was variable among crop species (*e.g.* number of days between initial and final height measurement below), depending on species phenology and developmental rates. After full extension of the first pair of true leaves, plant height was measured to the nearest mm. Plant height was measured again before transplanting. One representative fully mature but non senescent leaf was harvested from each of 10-15 individuals per accession. Leaves were then scanned at 400 d.p.i. using a large A3 flatbed scanner for large leaves (*e.g. Beta vulgaris*, WinRHIZO ProLA2400 unit, Regent Instruments Inc., Quebec, Canada). After scanning, a 1-5 cm^2 piece of the lamina was cut avoiding major veins, and placed on top of a soaked piece of germination paper overnight at 4 ºC to reach full hydration. The remainder of the leaf lamina was oven-dried at 70 ºC. The following day, the 1-5 cm^2 piece was weighed to the nearest microgram (MT XP6, Mettler-Toledo Inc., Westerville, Ohio, USA), oven-dried, and re-weighed. After three days at 70 ºC, all leaf material was weighed to obtain dry mass. Scanned leaves were processed with ImageJ software (http://rsb.info.nih.gov/ij) to obtain leaf area.

In 2012 we conducted an additional experiment, aimed to contribute all data for the intensive database, plus root and plant dry mass data for the extensive database. Seed weighing and germination procedures were carried out as in 2011, but germinating seeds were set into special containers appropriately built to characterize the root system. To this end, we used Root Trainers Jumbo containers that provide extra large container depth (Spencer Lemaire Ltd., Canada). Root Trainers are curly in the inside. We thus stuffed plastic cylinders inside Root Trainers to avoid root guiding through the curls of the containers, and also gain extra container depth. Containers built this way were 42 cm in depth and 8 cm in diameter. Containers were filled with commercial 100% sand substrate, aimed to facilitate the complete and integral recovery of the root system at harvest time. Plant were grown in those containers until inspection of the lower end of the containers spotted root tips for at least 3-4 plants per accession. At that time (ranging from 3 to 6 weeks since sowing date), all accessions of a given crop were harvested. Watering regime and greenhouse growth conditions were as described above for 2011, but containers were fertilized twice a week with complete nutrient solution to allow regular development in the sandy substrate.

The plants grown in 2012 were used to obtain 5-15 (median = 11) replicate scores, per each accession belonging to either the extensive or the intensive database, of the following traits: 1) Total plant Dry Mass (TDM, g) of plants kept growing for the same amount of time for all accessions of a given crop species; 2) Specific Root Length (SRL, m g^{-1}) of the whole root system; and 3) Root Tissue Density (RTD, g cm^{-3}). To obtain measures of those traits we proceeded as follows. At harvest time, individual plants were carefully dug and roots washed of soil particles to recover the whole root system intact. Root systems were subsequently set in a water-filled glass tray (A3-size) and scanned (gray scale, 400 dpi; [7,8]) using a flat-bed scanner (WinRHIZO ProLA2400 unit, Regent Instruments Inc., Quebec, Canada) equipped with a light transparency unit. Scanned root images were converted to binary black and white with ImageJ (http://rsb.info.nih.gov/ij). Total root length and root volume were further measured using the morphological analysis of WinRhizo (WinRHIZO Pro, Regent Instruments Inc., Quebec, Canada). Finally, the above- and the below-ground biomass of each plant was oven-dried at 70ºC for three days and weighed separately. Specific Root Length (SRL, m g^{-1}) was calculated as the ratio between the length of the root system and its dry mass. Root Tissue Density (RTD, g cm^{-3}) was calculated by dividing root dry mass by its fresh volume, as provided by Winrhizo software. Total plant Dry Mass (TDM, g) was the sum of above- and below-ground dry mass of each plant. Additionally, but only for the accessions belonging to the intensive database, all seed size and aboveground plant traits were measured as described above for the 2011 procedures. Canopy Height was, in this case, measured just before harvesting plants.

Data analysis

Our dataset had 3.02% missing data (31 out of 1026 accession*trait scores, see Database S1 in Supplementary Material for specific trait scores missing for specific

accessions). Following recommended procedures we adopted a multiple imputation approach to deal with missing data [9]. We generated ten complete datasets using Bayesian imputation, as implemented in Amos 18.0 [10]. For each of the PCA, PERMANOVA, SEM or Aster procedures described below, we used each of the ten complete databases separately. Reported parameter estimates, and measures for magnitude of goodness of fit in the main body of the paper, are average scores of fitting models to those ten complete databases. Given the low amount of missing data, in no case did statistical significance or directionality of effects of any model parameter change as a function of the dataset employed.

Effects of domestication status and crop identity on traits and on groups of traits

Traits were considered separately for analyses, and also in groups of two or three according to well-known strong physiological or developmental linkages among traits. Grouping of traits for further data analysis was performed through reduction of dimensionality. Four Principal Components Analyses were run separately for each of the following four groups of log-scaled traits. First, Seed Size and Total Dry Mass data were reduced to a first PCA axis aimed to represent organ and plant size variation among individuals (SIZE hereafter). Second, Seedling Absolute Growth Rate, Plant Canopy Height, and Leaf Size were reduced to a single PCA axis aimed to synthesize competitive ability for light capture (LIGHT COMP hereafter). Third, a PCA was run on leaf economic traits using Specific Leaf Area and Leaf Dry Matter Content (LEAF ECON hereafter). The first axis of that PCA was negatively related to structural investment in leaf tissue, which commonly correlates positively with leaf longevity and negatively with carbon fixation rates and mass-based nutrient status of leaves [11]. Lastly, a fourth PCA was run on root economic traits with Specific Root Length and Density of Fine Roots as component variables (ROOT ECON hereafter). First axes of the PCA analyses above explained 64 to 82% of variance in their composing variables (see Table S3) and were thus used as summarizing proxies of each of the four plant functions considered in this paper.

To evaluate the effects of domestication status and of crop identity on log-transformed trait and PCA axes scores, we used parameter-free permutational ANOVA analyses [12](PERMANOVA hereon). A PERMANOVA approach was preferred, instead of General Linear Mixed Model procedures, because residuals of GLMs would not accommodate normality assumptions for a majority of models, even after transforming data. In short, statistical testing in PERMANOVA approaches is based on the use of permutation tests over distance matrices derived from dataset vectors and thus do not rely on GLM assumptions. PERMANOVA analyses were carried out over Euclidean distance matrices. Statistical significance testing of model parameters was done using 4999 permutations of the raw data. Several PERMANOVA analyses were run. The first set of analyses used the extensive database. Each analysis on that database included trait values or PCA scores as the dependent variable, Domestication Status (either crop or wild) as a fixed-effect predictor, and Crop Identity (crop botanical genus) and Dom. Stat. * Crop Id. interaction as random-effect predictors. Two additional rounds of analyses were run using the intensive database. In the first round for this database we

assessed trait changes during initial domestication. In these analyses, we removed all data from improved (IM) accessions and used only data for wilds (WI) and landraces (LR). In the second round for the intensive database we assessed trait changes during the crop improvement phase of domestication. We thus removed WI accessions from the database and kept those with either LR or IM domestication statuses. Model specifications for PERMANOVA analyses with the intensive database were identical to those described above for the extensive database. Analyses were run with PERMANOVA+ for PRIMER statistical package (PRIMER-E Ltd., Plymounth Marine Laboratory, UK).

Coordinated evolution of traits during domestication and further improvement

A) Structural Equation Modelling:

We used Structural Equation Modelling (SEM) to investigate functional links among multiple traits, and their putative coordinated evolution during domestication [13]. First, we describe the datasets we used to build the several SEMs, and the rationale for model construction. After that, we provide statistical details on the estimation of goodness of fit, and on the statistical significance of path coefficients. Finally, in Supporting Information item 5 we provide results of additional SEM analyses aimed to provide additional empirical support for the general validity of the a priori inter-trait relationship scheme depicted in Fig. 2A in the main body of the paper.

Datasets and model construction

SEMs were implemented for two separate types of data sets. First, we put together all log-scaled arithmetic mean trait and PCA scores for each accession present in both the extensive and intensive databases ("complete" dataset hereafter, n = 114 accessions). Second, we calculated, independently for the extensive and intensive databases, the magnitude of the domestication and improvement effects over log-trait and PCA scores as follows. For the extensive database, we substracted the average score of each wild (W) accession from that of its crop (C) counterpart. This evolutionary change is denoted as $\Delta_{C\text{-}W}$Trait throughout the paper. For the intensive database we calculated an effect of domestication and an effect of subsequent improvement, separately. The domestication effect was taken by subtracting the average score of each wild (WI) accession from that of its landrace (LR) counterpart ($\Delta_{LR\text{-}WI}$Trait, hereafter). Since the intensive database included three accessions for each domestication status per each crop species, we subtracted every possible combination of wild accession from every landrace, for each species separately. This yielded nine $\Delta_{LR\text{-}WI}$Trait scores per crop species included in the intensive database. We proceeded in the same way to compute the improvement effect: every score of a landrace (LR) accession was subtracted from each of the three improved (IM) accessions available for each crop species ($\Delta_{IM\text{-}LR}$Trait, hereafter).

The above calculations resulted in four separate datasets: 1) a "complete" dataset, n= 114 log-scaled accession average log-trait and PCA scores; 2) "$\Delta_{C\text{-}W}$Trait" dataset, n = 30 wild-to-crop evolutionary transitions for each trait and PCA score; 3) "$\Delta_{LR\text{-}WI}$Trait" dataset, n = 54 wild-to-landrace evolutionary transitions for each trait and PCA score; 4) "$\Delta_{IM\text{-}LR}$Trait" dataset, n = 54 landrace-to-improved evolutionary transitions for each trait and PCA score.

Based on previous knowledge and on the patterns of trait correlation that we observed in our study, we first designed an overall causal conceptual structure that identified four groups of variables in a model. Simple arrows were employed when sufficient knowledge of the direction of causality was available, whereas double-headed arrows were used if direction of causality was either unclear or could work in both ways. This structure included the following specific expectations:

1) Groups of traits affecting specific functions (*e.g.* capacity to compete for light, or root foraging ability), tend to co-vary tightly. For instances, Seedling Absolute Growth Rate and Canopy Height both promote above-ground competition ability and tend to be phenotypically linked [14]. There is abundant literature supporting this assumption for the other three groups of functions modelled here: leaf economics [11,15], root economics [16,17], and size [18,19].

2) Increases in size of individuals and organs tend to result in diminishing returns in terms of physiological revenue from resource capturing organs [20,21]. Therefore, we expect increases in size-related traits to co-evolve with slower leaf and root economic traits, but to promote faster vertical growth, taller canopies, larger leaves, and therefore greater competitive ability for light.

3) Duration, structural investment, and physiological performance of fine roots and leaves may, or may not, evolve in a coordinated fashion. Theoretical proposals suggest that high photosynthetic capacity and biomass renewal rates above-ground should require rapid nutrient uptake rates and short lived fine roots belowground [22,23]. However, empirical evidence is diverse [16,17,24,25].

4) Capacity to compete for light, that is, ability to shade competitors, may interact with leaf and root economics. Above ground, this might occur as a direct effect, via trade-offs between displaying large light-capturing surfaces *vs* smaller but physiologically more active ones [26]. But this correlation may also arise as an indirect allometric effect of increased size (see 2 above). Increased competitive ability requires increased investment in mostly heterotrophic stem and petiole tissue [14], which might incur diminishing returns from investment in leaf photosynthetic tissue and fine root water and nutrient uptake capacity. Therefore, we predict a negative relationship between

capacity to compete for light and fast leaf and root economics (i.e. fast photosynthesis and water and minerals uptake rates, and fast biomass renewal rates of productive tissues).

The model structure finally selected by goodness of fit estimates should 1) address the extent to which variation in the four separate suites of traits among all accessions in our database ("complete" dataset") is consistent with previous physiological and ecological knowledge (as outlined above); and 2) account for the extent of coordinated co-evolution of traits during domestication and further improvement, as reflected in datasets "$\Delta_{C\text{-}W}$Trait", "$\Delta_{LR\text{-}WI}$Trait", and "$\Delta_{IM\text{-}LR}$Trait".

Considering the above a priori constraints, we generated several tentative specific models, and the model that received the highest statistical support for our "complete" dataset is shown in Figure 2B. In this model, we used the first PCA axis of each functional grouping of traits, instead of raw trait scores. This accommodates our expectation number 1 above (*i.e.* intense within-group coupling of trait variation) and is analogous to the usage of latent variables in complete SEM structures [27]. Also, a simpler model excluding measured traits best complied with SEM rules of thumb for sample size to number of variables ratio for our dataset [27]. Computation of PCA axes is described above (subsection "Effects of domestication status and crop identity on traits and on groups of traits"). The model in Figure 2B was further fitted to the "$\Delta_{C\text{-}W}$Trait", "$\Delta_{LR\text{-}WI}$Trait", and "$\Delta_{IM\text{-}LR}$Trait" datasets to investigate co-evolution of traits during domestication and further improvement.

Goodness of fit measures and statistical significance of path coefficients

We used parameter-free approaches for assessing goodness of fit of models to data, and to evaluate whether single path and correlation coefficients were statistically different from zero. The degree of fit between the observed and expected covariance structures was first assessed by a χ^2 goodness-of-fit test. A significant goodness-of-fit test indicates that the model does not fit the data globally. Since sample size was small for one of the four SEM models ("$\Delta_{C\text{-}W}$Trait"), we estimated the statistical significance of χ^2 goodness-of-fit statistic using two independent methods, both robust to low sample sizes. First, the probability value of the obtained χ^2 was obtained with MCX2 (http://pages.usherbrooke.ca/jshipley/recherche/book.htm), which yields probability estimates for the maximum likelihood χ^2-statistic based on small sample sizes [13]. Second, χ^2 statistical significance was computed using the Bollen and Stine bootstrap test [28], which is also robust to small samples. However, significant χ^2 can result from violation of certain assumptions, whereas failure to reject a model (a non-significant χ^2) may result from inadequate statistical power [29]. Therefore, we also evaluated model fit to the data by means of the Goodness of Fit Index (GFI) and the Root Mean Square Error of Approximation (RMSEA), which are often used in SEM and are insensitive to sample size [30]. Values of GFI range between 0 and 1, and values >0.9 indicate an acceptable fit of the model to the data [30]. RMSEA values <0.1 also indicate

acceptable fit between model and data [30]. Bootstrapped and MCX2 probability values yielded practically equal statistical significances, thus only bootstrapped p-values are shown in the main body of the paper for simplicity.

Standardized path and correlation coefficients were estimated using maximum likelihood. Standardized partial-regression coefficients help interpret expected changes in the dependent variable in response to a unit of change in the predictor, while controlling for shared variance of the predictor and response with other predictors [31]. We used standardized coefficients to interpret inter-trait scaling relationships. Frequency distribution of the several dependent variables in each model did not always fit to a Gaussian distribution. Therefore, statistical significance was evaluated for each single standardized path and correlation coefficient in the model through bootstrapping [10]. Standardized path or correlations coefficients < |0.05| were constrained to be zero to achieve model identification. All SEM analyses were performed using AMOS 18.0 (SPSS Inc, Chicago, USA).

B) Analysis of phenotypic integration:

Phenotypic integration levels are often assessed based on trait-to-trait correlation matrices. Indices of phenotypic integration are computed based on magnitude of the absolute value coefficients in correlation matrices, and compared among experimental subjects or populations (*e.g.* eigenvalue-based INT index[32,33]). However, by definition, a non-significant correlation coefficient is indistinguishable from zero, and should be considered zero in any type of further analysis of collections of correlation coefficients. Thus, the most common approach is to quantify the number of significant correlations, and the magnitude of the absolute value of correlation coefficients, separately [34–36]. This procedure precludes the emission of a single unified answer to the hypothesis being tested, because two different components, significance and magnitude, are evaluated through separate statistical tests. However, if collections of correlation coefficients are taken as a unified dataset, such collections would most frequently accommodate a bimodal distribution, with an initial peak at zero (i.e. non-significant coefficient parameters), and a later one at the median of truly significant coefficients. Bimodality hinders most common procedures of general or generalized hypothesis testing and model evaluation [37]. To address this problem, and be able to provide a unified single test of whether phenotypic integration differed between datasets, we made use of *Aster* models [37,38]. *Aster* models were developed to address the typically bimodal distribution of lifetime fitness data [37]. Lifetime fitness is composed of the survival of a given percentage of individuals, coupled to the fecundity of survivors. In this context, Aster models account for the dependence of fitness components expressed later in ontogeny (*e.g.* fecundity) on processes expressed earlier (*e.g.* survival) [37]. This is achieved through the use of forest graph exponential family of canonical models that combine generalized linear modelling with survivorship analysis in a single statistical test [38]. We made use of Aster models here to analyze the bimodal distributions of our correlation-coefficient datasets. We dissected our correlation dataset in two “life history” components: (1) whether a correlation was statistically different from zero (analog to survivorship) and (2), if non-zero, absolute

magnitude of the correlation coefficient (analogous to fecundity). Statistical significance was assumed to accommodate a Bernoulli distribution, and magnitude of coefficients to follow a zero-truncated Poisson distribution.

We generated an Aster model where magnitude of correlation coefficients, conditional on prior statistical significance of coefficients, was used as the dependent variable. "Dataset" was the independent fixed effect factor, which had four levels: "complete", "$\Delta_{C\text{-}W}$Trait", "$\Delta_{LR\text{-}WI}$Trait", and "$\Delta_{IM\text{-}LR}$Trait"(see SEM procedures above for datasets nomenclature). Significance of the "Dataset" factor was tested through comparison of our model to a reduced model where "Dataset" was not included as a predictor. This comparison was carried out through log-likelihood ratio tests [37]. If "Dataset" was found to exert a significant effect over the magnitude of correlation coefficients, then multiple paired comparisons among the four levels of "Dataset" were carried out. Those multiple comparisons were also carried out through log-likelihood ratio tests. *Aster* models and log-likelihood ratio tests were run using the aster and anova.aster functions, respectively, of the *Aster* package [37,38] available for the R platform [39]. Detailed information and resources on the mathematical basis and usage of *Aster* models is available at http://www.stat.umn.edu/geyer/aster/.

2) SUPPLEMENTARY TABLES

Table S1: Botanical name, domestication status, and seed origin information of each accession of the extensive set of 30 crop-wild ancestor pairs used in this study. Domestication status (C: cultivated; W: wild ancestor). Seed donor (IPK: Germplasm bank of the Leibniz Institute of Plant Genetics and Crop Plant Research, Germany; NPGS: National Plant Germplasm System-USDA, U.S.A.; CRF: Centro Nacional de Recursos Fitogenéticos-INIA, Spain; ICARDA: International Center for Agricultural Research in Dry Areas-FAO, Syria; CGN: Center for Genetic Resources, The Netherlands; UPV: Seedbank of the Polythecnic University of Valencia, Spain; CIRAD: Centre de Coopération Internationale en Recherche Agronomique pour le Devélopemment, France; IRRI: International Rice Research Institute; JIC: John Innes Center; * commercial company). Accession identifier refers to the code assigned by each seed donor excepting the commercial companies (N.A., not applicable). Accession country refers to the country where the seeds were collected, if applicable. Ref: reference source for wild ancestor assignment (Ref list is provided as a footnote below the Table).

Botanical name	Dom. status	Accesion identifier	Seed donor	Accesion country	Ref.
Avena sativa	C	BGE024681	CRF	Spain	1
Avena sterilis	W	IG 100379 IFMI 3096	ICARDA	Turkey	1
Beta vulgaris	C	N.A.	Clause*	commercial	1
Beta vulgaris	W	1582	IPK	Italy	1
Brassica oleracea	C	N.A.	Rocalba*	commercial	2
Brassica oleracea	W	CGN18947	CGN	Germany	2
Capsicum anuum	C	N.A.	Mascarell*	commercial	2
Capsicum anuum	W	PI631137	NPGS	Guatemala	2
Capsicum bacattum	C	CGN23297	CGN	Peru	2
Capsicum bacattum	W	CGN23278	CGN	Argentina	2
Cicer arietinum	C	BGE024684	CRF	commercial	2
Cicer reticulatum	W	IG72945 ILWC116	ICARDA	Turkey	2
Cichorium endibia	C	N.A.	Rocalba*	commercial	3
Cichorium intybus	W	BGE032596	CRF	Spain	3
Cynara cardunculus	C	N.A.	Rocalba*	Spain	4
Cynara cardunculus	W	ES-01-14-0256	Semillas Silvestr	Spain	4
Eruca sativa	C	N.A.	Rocalba*	commercial	5
Eruca sativa	W	ERU 115	IPK	Pakistan	5
Glycine max	C	N.A.	Biográ*	commercial	6
Glycine soja	W	1039	IPK	Russia	6
Gossypium hirsutum	C	BGE006434	CRF	USA	2
Gossypium hirsutum	W	BG 6050	CIRAD	France	2
Helianthus annuus	C	HEL 226	IPK	USA	2
Helianthus annuus	W	PI413093	NPGS	USA	2
Hordeum vulgare	C	BGE000214	CRF	commercial	2
Hordeum spontaneum	W	BGE025385	CRF	Morocco	2
Lathyrus sativus	C	BGE014724	CRF	Spain	7
Lathyrus cicera	W	BGE019570	CRF	Spain	7
Lens culinaris	C	BGE024692	CRF	commercial	2
Lens orientalis	W	IG 72642 IFWL 119	ICARDA	Syria	2
Lupinus luteus	C	LO4500	CRF	commercial	8
Lupinus luteus	W	LO4579	CRF	Portugal	8
Solanum lycopersicum	C	N.A.	Clause*	commercial	2
Solanum pimpinellifolium	W	LA1383	NPGS	Peru	2
Medicago lupulina	C	N.A.	Intersemillas*	commercial	9
Medicago lupulina	W	IG 58734 IFMA 6092	ICARDA	Turkey	9
Oryza sativa	C	N.A.	Calasparra*	commercial	10
Oryza nivara	W	IRGC 104969	IRRI	China	10
Pennisetum glaucum	C	PI586660	NPGS	Burkina Fasso	11
Pennisetum glaucum	W	PI537068	NPGS	Niger	11
Pisum sativum	C	2600	JIC	commercial	2
Pisum humile	W	1794	JIC	Israel	2
Secale cereale	C	BGE010915	CRF	commercial	2
Secale ancestrale	W	PI618666	NPGS	Turkey	2
Sesamum indicum	C	N.A.	Rocalba*	commercial	12
Sesamum indicum	W	17	IPK	Yemem	12
Sorghum sudanense	C	N.A.	Rocalba*	commercial	2
Sorghum bicolor	W	PI524718	NPGS	Sudan	2
Spinacea oleracea	C	N.A.	Rocalba*	commercial	13
Spinacea turkestanica	W	CGN9546	CGN	Uzbekistan	13
Trifolium repens	C	N.A.	Intersemillas*	commercial	14
Trifolium repens	W	CGN22513	CGN	Kyrgyzstan	14
Triticum durum	C	BGE020911	CRF	commercial	2
Triticum diccocoides	W	352322	NPGS	Lebanon	2
Vicia faba	C	N.A.	Rocalba*	commercial	1
Vicia narbonensis	W	IG 111590 IFVI 5266	ICARDA	Tunisia	1
Vigna unguiculata	C	PI599213	NPGS	commercial	15
Vigna unguiculata	W	PI447516	NPGS	Nigeria	15
Zea mays	C	Ames26252	NPGS	Brazil	16
Zea mays	W	PI566674	NPGS	Mexico	16

Footnote: **1.** Hancock, JF. 2004. Plant Evolution and the origin of crop species. CABI Publishing, NY, USA. **2.** Sauer, JD. 1993. Historical geography of crop plants. A select roster. CRC Press. Boca Raton, USA. **3.** Kiær LP, et al. 2009. Genetic Resources and Crop Evolution, 56, 405-419. **4.** Sonnante G, Pignone D, and Hammer K. 2007. Annals of Botany 100: 1095–1100. **5.** Pignone D, and Gómez-Campo C. 2011. Eruca. In Wild Crop Relatives: Genomic and Breeding Resources, Oilseeds (Kole C, ed). Pp. 149-160. Springer-Verlag, Berlin. **6.** Hymowitz T, and Newell CE. 1981. Economic Botany 35: 272-288. **7.** Sarker A, El Moncim A, and Maxted N. 2001. Grasspea and chicklings. In Plant Genetic Resources of Legumes in the Mediterranean. Maxted and Bennett eds. Pp. 159-180.Kluwer Acad. Publishers, Dordrech, The Netherlands. **8.** Wolko B et al. 2011. Lupins. In Wild Crop Relatives: Genomic and Breeding Resources, Legume Crops and Forages (Kole C, ed). Pp. 153-206. Springer-Verlag, Berlin. **9.** A. Chandra, S. Verma, K.C. Pandey. 2011. Biochemical Systematics and Ecology 39: 711-717. **10.** Yamanaka S et al. 2003. Genetic Resources and Crop Evolution 50: 529-538. **11.** Lewis LR. 2010. The Professional Geographer 62: 377–395. **12.** Fuller, DQ. 2003. Asian Agri-History 7(2), 127–137. **13.** Andersen SB and Torp AM. 2011. Spinacea. in Wild Crop Relatives: Genomic and Breeding Resources, Vegetables. (Kole C, ed). Pp. 273-276. Springer-Verlag, Berlin. **14.** Frame J, Newbould P. 1986. Advances in Agronomy 40: 1-88. **15.** Norihiko Tomooka, Akito Kaga, Takehisa Isemura, and Duncan Vaughan. 2011. Vigna. In Wild Crop Relatives: Genomic and Breeding Resources, Legume Crops and Forages (Kole C, ed). Pp. 291-311. Springer-Verlag, Berlin. **16.** Wilkes G. 2007. Maydica 52:49-60

Table S2: Botanical name, domestication status, and seed origin information of each accession of the intensive set of six crop species that were studied in more detail. Domestication status (IM: improved; LR: landrace; WI: wild). Accessions identifiers and countries as in Table S1. Seeds donors names as in Table S1, but CIMMYT (Centro Internacional de Mejoramiento de Maiz y Trigo, Mexico), and CITA (Centro de Investigación Agraria de Aragón, Spain).

Botanical name	Dom. status	Accesion identifier	Seed donor	Accesion country
Hordeum spontaneum	WI	BGE025385	CRF	Morocco
Hordeum spontaneum	WI	PI 282671	NPGS	Asia
Hordeum spontaneum	WI	PI 662181	NPGS	Turkey
Hordeum vulgare	LR	clho1246	NPGS	USA
Hordeum vulgare	LR	PI 51209	NPGS	Israel
Hordeum vulgare	LR	PI 467371	NPGS	France
Hordeum vulgare	IM	N.A.	Batllé*	commercial
Hordeum vulgare	IM	N.A.	Batllé*	commercial
Hordeum vulgare	IM	N.A.	Batllé*	commercial
Capsicum anuum	WI	CGN22774	CGN	Guatemala
Capsicum anuum	WI	CGN22865	CGN	USA
Capsicum anuum	WI	CGN23200	CGN	Costa Rica
Capsicum anuum	LR	CGN21530	CGN	Nigeria
Capsicum anuum	LR	CGN24354	CGN	Mexico
Capsicum anuum	LR	CGN16901	CGN	China
Capsicum anuum	IM	N.A.	Fitó*	commercial
Capsicum anuum	IM	N.A.	Fitó*	commercial
Capsicum anuum	IM	N.A.	Fitó*	commercial
Zea mays	WI	27460	CIMMYT	Mexico
Zea mays	WI	27478	CIMMYT	Mexico
Zea mays	WI	27545	CIMMYT	Mexico
Zea mays	LR	2216	CIMMYT	Mexico
Zea mays	LR	2536	CIMMYT	Mexico
Zea mays	LR	10435	CIMMYT	Mexico
Zea mays	IM	N.A.	Fitó*	commercial
Zea mays	IM	N.A.	Fitó*	commercial
Zea mays	IM	N.A.	Fitó*	commercial
Helianthus annuus	WI	PI 468435	NPGS	USA
Helianthus annuus	WI	PI 435500	NPGS	USA
Helianthus annuus	WI	PI 435608	NPGS	USA
Helianthus annuus	LR	PI 496266	NPGS	China
Helianthus annuus	LR	PI 526256	NPGS	Zimbabwe
Helianthus annuus	LR	PI 600719	NPGS	USA
Helianthus annuus	IM	N.A.	Fitó*	commercial
Helianthus annuus	IM	N.A.	Fitó*	commercial
Helianthus annuus	IM	N.A.	Rocalba*	commercial
Brassica oleracea	WI	CGN18947	CGN	Germany
Brassica oleracea	WI	CGN06903	CGN	France
Brassica oleracea	WI	9804	UPM	Spain
Brassica oleracea	LR	CGN14079	CGN	Belgium
Brassica oleracea	LR	CGN18467	CGN	Turkey
Brassica oleracea	LR	BGHZ.1148	CITA	Spain
Brassica oleracea	IM	N.A.	Fitó*	commercial
Brassica oleracea	IM	N.A.	Batllé*	commercial
Brassica oleracea	IM	CGN18467	CGN	commercial
Pisum humile	WI	1794	JIC	Israel
Pisum humile	WI	JI 3239	JIC	Syria
Pisum humile	WI	W6 2044	NPGS	Turkey
Pisum sativum	LR	960	JIC	Turkey
Pisum sativum	LR	1033	JIC	India
Pisum sativum	LR	1281	JIC	Ethiopia
Pisum sativum	IM	N.A.	Fitó*	commercial
Pisum sativum	IM	N.A.	Fitó*	commercial
Pisum sativum	IM	N.A.	Fitó*	commercial

Table S3: PCA loadings of each log-scaled individual trait on the first PCA axes named as SIZE, LIGHT COMP, LEAF ECON, and ROOT ECON. % Variation accounted for by each first axis: SIZE = 77%; LIGHT COMP = 64%; LEAF ECON = 82%; and ROOT ECON = 80%.

PCA LEAFECON	Factor Loading on axis 1
log (SLA)	0,908
log (LDMC)	-0,908

PCA ROOTECON	Factor Loading on axis 1
log (SRL)	0,896
log (ROOTDENS)	-0,896

PCA SIZE	Factor Loading on axis 1
log (SEEDSIZE)	0,877
log (TDM)	0,877

PCA LIGHT COMP	Factor Loading on axis 1
log (HEIGHT)	0,910
log (AGR)	0,733
log (LEAFSIZE)	0,740

Table S4: **Spearman correlation coefficients matrices** among traits and PCA eigenvalues (current paper´s dataset, n= 114 accessions). Also, matrices are shown for differences in traits or PCA eigenvalues between wilds and crops of the extensive database ($\Delta_{C\text{-}W}$Trait, 30 crops, , n = 30), between Wilds and Landraces ($\Delta_{LR\text{-}WI}$Trait, multiple accessions of six crops, n = 54) and between Landraces and Improved accessions ($\Delta_{IM\text{-}LR}$Trait, multiple accessions of six crops, n = 54). * and ** correlation is significant at p = 0.05 or p = 0.01, respectively.

CURRENT PAPER´S DATASET

	log (SEEDSIZE)	log (HEIGHT)	log (AGR)	log (LEAFSIZE)	log (SLA)	log (LDMC)	log (SRL)	log (ROOTDENS)	log (TDM)
log (SEEDSIZE)	1								
log (HEIGHT)	,264**	1							
log (AGR)	,614**	,598**	1						
log (LEAFSIZE)	0,08	,683**	,319**	1					
log (SLA)	,260**	-,483**	0,089	-,518**	1				
log (LDMC)	-0,007	,384**	0,066	,265**	-,684**	1			
log (SRL)	-,613**	-0,088	-,303**	-0,034	-0,027	-,214*	1		
log (ROOTDENS)	,213*	0,036	0,037	0,011	-,195*	,371**	-,653**	1	
log (TDM)	,575**	,383**	,387**	,369**	-0,099	0,171	-,541**	,221*	1

	PCA LIGHT COMP	PCA LEAF ECON	PCA ROOT ECON	PCA SIZE
PCA LIGHT COMP	1			
PCA LEAF ECON	-,369**	1		
PCA ROOT ECON	-0,105	,222*	1	
PCA SIZE	,469**	0,013	-,495**	1

$\Delta_{C\text{-}W}$

	log (SEEDSIZE)	log (HEIGHT)	log (AGR)	log (LEAFSIZE)	log (SLA)	log (LDMC)	log (SRL)	log (ROOTDENS)	log (TDM)
log (SEEDSIZE)	1								
log (HEIGHT)	,386*	1							
log (AGR)	-0,164	,380*	1						
log (LEAFSIZE)	,471**	0,158	-0,098	1					
log (SLA)	-0,163	0,27	0,345	-0,092	1				
log (LDMC)	-0,059	-0,06	-0,184	-0,228	-,526**	1			
log (SRL)	-0,108	0,075	0,028	0,121	0,348	-0,201	1		
log (ROOTDENS)	-0,241	-0,036	0,23	-0,191	-0,003	-0,004	-,374*	1	
log (TDM)	,609**	,489**	0,079	0,22	-0,232	0,087	-0,022	-0,186	1

	PCA LIGHT COMP	PCA LEAF ECON	PCA ROOT ECON	PCA SIZE
PCA LIGHT COMP	1			
PCA LEAF ECON	0,234	1		
PCA ROOT ECON	0,031	0,179	1	
PCA SIZE	0,353	-0,233	0,031	1

Δ_{LR-WI}

	log (SEEDSIZE)	log (HEIGHT)	log (AGR)	log (LEAFSIZE)	log (SLA)	log (LDMC)	log (SRL)	log (ROOTDENS)	log (TDM)
log (SEEDSIZE)	1								
log (HEIGHT)	,678**	1							
log (AGR)	,299*	,620**	1						
log (LEAFSIZE)	,556**	,323*	-0,061	1					
log (SLA)	0,027	-0,228	-,280*	,364**	1				
log (LDMC)	-,470**	-0,187	-0,017	-,618**	-,455**	1			
log (SRL)	-,408**	-,569**	-,421**	-0,084	0,119	-0,074	1		
log (ROOTDENS)	-,373**	-0,005	,334*	-,344*	-0,251	,577**	-,332*	1	-0,029
log (TDM)	,668**	,563**	0,239	,683**	-0,132	-,311*	-,298*	-0,029	1

	PCA LIGHT COMP	PCA LEAF ECON	PCA ROOT ECON	PCA SIZE
PCA LIGHT COMP	1			
PCA LEAF ECON	0,145	1		
PCA ROOT ECON	-0,183	,487**	1	
PCA SIZE	,757**	0,222	0,019	1

Δ_{IM-LR}

	log (SEEDSIZE)	log (HEIGHT)	log (AGR)	log (LEAFSIZE)	log (SLA)	log (LDMC)	log (SRL)	log (ROOTDENS)	log (TDM)
log (SEEDSIZE)	1								
log (HEIGHT)	-0,142	1							
log (AGR)	-0,163	,758**	1						
log (LEAFSIZE)	,358**	-0,077	-0,116	1					
log (SLA)	0,037	0,018	0,146	,294*	1				
log (LDMC)	0,229	-0,2	-,281*	-0,246	-,561**	1			
log (SRL)	-0,236	-0,12	-0,206	0,071	-0,043	-0,22	1		
log (ROOTDENS)	-0,023	0,121	-0,016	-0,245	-,321*	,478**	-,624**	1	
log (TDM)	,352**	0,024	-0,058	,757**	-0,132	-0,005	-0,103	0,035	1

	PCA LIGHT COMP	PCA LEAF ECON	PCA ROOT ECON	PCA SIZE
PCA LIGHT COMP	1			
PCA LEAF ECON	,302*	1		
PCA ROOT ECON	-0,028	,344*	1	
PCA SIZE	0,188	-0,16	-0,134	1

Table S5. PERMANOVA results for treatment effects and interactions (Domestication Status, Crop identity, and Dom. Status * Crop id.) on traits and PCA eigenvalues of the extensive 30 species dataset. Domestication Status has two levels (Wild and Crop). Results in Fig. 1A in the main body of the paper refer to statistics in this table. Values of P below 0.05 are shown in boldface. See Supplementary Materials and Methods for details on analyses.

HEIGHT

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,3644	2,8153	0,1014
CROP ID	27	2,3889	231,97	**0,0002**
DOMEST × CROP ID	28	0,1548	15,034	**0,0002**

AGR

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	2,0846	5,7352	**0,0240**
CROP ID	27	3,0919	39,278	**0,0002**
DOMEST × CROP ID	28	0,4242	5,3886	**0,0002**

LEAF SIZE

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	2,5553	5,5261	**0,0288**
CROP ID	27	9,8790	600,38	**0,0002**
DOMEST × CROP ID	28	0,5575	33,879	**0,0002**

LDMC

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0039	0,096956	0,7598
CROP ID	27	0,3963	123,1	**0,0002**
DOMEST × CROP ID	28	0,0464	14,415	**0,0002**

SLA

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0171	0,37106	0,5384
CROP ID	27	0,3699	78,798	**0,0002**
DOMEST × CROP ID	28	0,0535	11,387	**0,0002**

SRL

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0117	0,14695	0,7016
CROP ID	27	1,9072	137,83	**0,0002**
DOMEST × CROP ID	28	0,0880	6,3633	**0,0002**

ROOT DENSITY

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0414	0,80399	0,3734
CROP ID	27	0,3496	35,793	**0,0002**
DOMEST × CROP ID	28	0,0570	5,8328	**0,0002**

SEED SIZE

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	8,3591	22,571	**0,0004**
CROP ID	27	8,1306	1242,9	**0,0002**
DOMEST × CROP ID	28	0,4579	69,992	**0,0002**

TDM

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	3,0802	9,6538	**0,0036**
CROP ID	27	0,8413	36,105	**0,0002**
DOMEST × CROP ID	28	0,3902	16,747	**0,0002**

PCA LIGHT COMP

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	9,1152	9,4993	**0,006**
CROP ID	27	17,8020	170,18	**0,0002**
DOMEST × CROP ID	28	1,1418	10,915	**0,0002**

PCA LEAF ECON

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,6245	0,28055	0,6196
CROP ID	27	16,5670	90,622	**0,0002**
DOMEST × CROP ID	28	2,5841	14,136	**0,0002**

PCA ROOT ECON

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,2072	0,19762	0,6468
CROP ID	27	12,7920	62,591	**0,0002**
DOMEST × CROP ID	28	1,1582	5,6669	**0,0002**

PCA SIZE

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	28,8900	14,276	**0,001**
CROP ID	27	11,5000	139,66	**0,0002**
DOMEST × CROP ID	28	2,4905	30,245	**0,0002**

Table S6. PERMANOVA results for treatment effects and interactions (Domestication Status, Crop identity, and Dom.Status * Crop id.) on traits and PCA eigenvalues of part of the intensive dataset where six species where investigated in more detail. Here, Domestication Status has two levels, selected to characterize the early domestication effect (Wild and Landrace). Results in Fig. 1B in the main body of the paper refer to statistics in this table. Values of P below 0.05 are shown in boldface. See Supplementary Materials and Methods for details on analyses.

HEIGHT

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	2,1810	14,111	**0,0136**
CROP ID	5	5,9821	321,25	**0,0002**
DOMEST × CROP ID	5	0,1922	10,324	**0,0002**

AGR

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	2,1030	6,5246	**0,0496**
CROP ID	5	17,1120	275,71	**0,0002**
DOMEST × CROP ID	5	0,3944	6,3552	**0,0002**

LEAF SIZE

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	1,7647	1,977	0,2216
CROP ID	5	5,3012	191,69	**0,0002**
DOMEST × CROP ID	5	1,1323	40,943	**0,0002**

LDMC

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0531	0,53819	0,5044
CROP ID	5	0,3379	80,017	**0,0002**
DOMEST × CROP ID	5	0,0995	23,553	**0,0002**

SLA

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0894	0,45997	0,5168
CROP ID	5	0,3911	68,34	**0,0002**
DOMEST × CROP ID	5	0,1960	34,249	**0,0002**

SRL

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,1044	0,65104	0,4598
CROP ID	5	1,8057	90,269	**0,0002**
DOMEST × CROP ID	5	0,1905	9,5216	**0,0002**

ROOT DENSITY

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,2009	2,5595	0,1778
CROP ID	5	0,5061	50,346	**0,0002**
DOMEST × CROP ID	5	0,0932	9,2712	**0,0002**

SEED SIZE

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	18,4970	8,5736	**0,034**
CROP ID	5	28,3460	233,5	**0,0002**
DOMEST × CROP ID	5	2,1811	17,967	**0,0002**

TDM

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	3,9804	3,0164	**0,1446**
CROP ID	5	6,7642	273,75	**0,0002**
DOMEST × CROP ID	5	1,3346	54,012	**0,0002**

PCA LIGHT COMP

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	12,5810	8,4608	**0,0288**
CROP ID	5	43,1520	419,91	**0,0002**
DOMEST × CROP ID	5	1,8706	18,203	**0,0002**

PCA LEAF ECON

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	6,7287	0,61314	0,4776
CROP ID	5	29,1060	87,969	**0,0002**
DOMEST × CROP ID	5	11,0640	33,44	**0,0002**

PCA ROOT ECON

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	5,6262	2,564	0,1714
CROP ID	5	28,4730	69,761	**0,0002**
DOMEST × CROP ID	5	2,5778	6,3157	**0,0004**

PCA SIZE

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	33,2370	5,7734	**0,0438**
CROP ID	5	46,9540	398,17	**0,0002**
DOMEST × CROP ID	5	5,8224	49,373	**0,0002**

Table S7. PERMANOVA results for treatment effects and interactions (Domestication Status, Crop identity, and Dom.Status * Crop id.) on traits and PCA eigenvalues of part of the intensive dataset where six species where investigated in more detail. Here, Domestication Status has two levels, selected to characterize the later improvement domestication effect (Landrace and Improved). Results in Fig. 1C in the main body of the paper refer to statistics in this table. Values of P below 0.05 are shown in boldface. See Supplementary Materials and Methods for details on analyses.

HEIGHT

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,1307	1,0046	0,358
CROP ID	5	5,3626	427,17	**0,0002**
DOMEST × CROP ID	5	0,1465	11,673	**0,0002**

AGR

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0597	0,29897	0,6078
CROP ID	5	16,9050	332,73	**0,0002**
DOMEST × CROP ID	5	0,2205	4,3397	**0,0012**

LEAF SIZE

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,3716	3,1386	0,1288
CROP ID	5	6,8134	249,47	**0,0002**
DOMEST × CROP ID	5	0,1311	4,8015	**0,0002**

LDMC

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0033	0,13612	0,7416
CROP ID	5	0,5757	145,84	**0,0002**
DOMEST × CROP ID	5	0,0248	6,2948	**0,0002**

SLA

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0143	2,0178	0,2212
CROP ID	5	0,6770	136,59	**0,0002**
DOMEST × CROP ID	5	0,0072	1,4505	**0,2086**

SRL

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0364	0,57607	0,4732
CROP ID	5	1,6657	110,85	**0,0002**
DOMEST × CROP ID	5	0,0804	5,3507	**0,0004**

ROOT DENSITY

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0128	1,196	0,3128
CROP ID	5	0,7979	115,75	**0,0002**
DOMEST × CROP ID	5	0,0121	1,7498	**0,1206**

SEED SIZE

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0743	0,49295	0,5122
CROP ID	5	36,2290	311,68	**0,0002**
DOMEST × CROP ID	5	0,1518	1,3057	**0,2758**

TDM

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,1122	0,98774	0,3648
CROP ID	5	7,9555	305,47	**0,0002**
DOMEST × CROP ID	5	0,1162	4,462	**0,0008**

PCA LIGHT COMP

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0012	0,0036021	0,9592
CROP ID	5	48,1810	521,41	**0,0002**
DOMEST × CROP ID	5	0,3576	3,8696	**0,0024**

PCA LEAF ECON

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0694	0,069926	0,8
CROP ID	5	55,9010	179,83	**0,0002**
DOMEST × CROP ID	5	1,0202	3,2817	**0,0082**

PCA ROOT ECON

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,7191	1,0898	0,3462
CROP ID	5	34,7950	119,28	**0,0002**
DOMEST × CROP ID	5	0,7922	2,7159	**0,0198**

PCA SIZE

Source of variation	df	MS	Pseudo-F	*P*
DOMEST	1	0,0706	0,23028	0,6466
CROP ID	5	59,8940	536,43	**0,0002**
DOMEST × CROP ID	5	0,3125	2,7988	**0,0176**

Table S8: Log-likelihood ratio tests for comparing correlation matrices of Traits vs those of Trait variation during crop evolution ("Traits" vs "$\Delta_{C\text{-}W}$Trait"); and for comparing Trait evolution during early domestication vs Trait evolution during later improvement ("$\Delta_{LR\text{-}WI}$Trait" vs "$\Delta_{IM\text{-}LR}$Trait"). Models under comparison are Aster models (see Materials and Methods in Supplementary Material). Change in deviance is twice the log likelihood ratio. A significant change in deviance indicates improvement of the previous model following the addition of Dataset as a predictor. Datasets codes for correlation coefficients belonging to "Traits", "$\Delta_{C\text{-}W}$Trait", "$\Delta_{IM\text{-}LR}$Trait", or "$\Delta_{LR\text{-}WI}$Trait" matrices. Null is a model including no predictor to provide a baseline against which to compare the effect of adding Dataset in the Full model.

Partial model name	Model formula				
Null	$y \sim 1 + \varepsilon$				
Full	$y \sim Dataset + \varepsilon$				
Datasets under comparison	Model comparison	d.f.	Model deviance	Change in deviance	P-value of log-likelihood ratio test
Trait vs $\Delta_{C\text{-}W}$Traits	Null	2	-2539.2		
	Null vs Full	3	-2553.9	14.689	**0.00013**
$\Delta_{LR\text{-}WI}$Traits vs $\Delta_{IM\text{-}LR}$Traits	Null	2	-3306.3		
	Null vs Full	3	-3313.1	68.084	**0.00907**

3) SUPPLEMENTARY FIGURES

Figure S1. Map showing the location of the seed origin of the 48 wild ancestor accessions of this project (30 for the extensive multi-crop experiment, and 18 for the 3 accessions for each of the 6 crops in the intensive experiment). Each accession is represented by a red point.

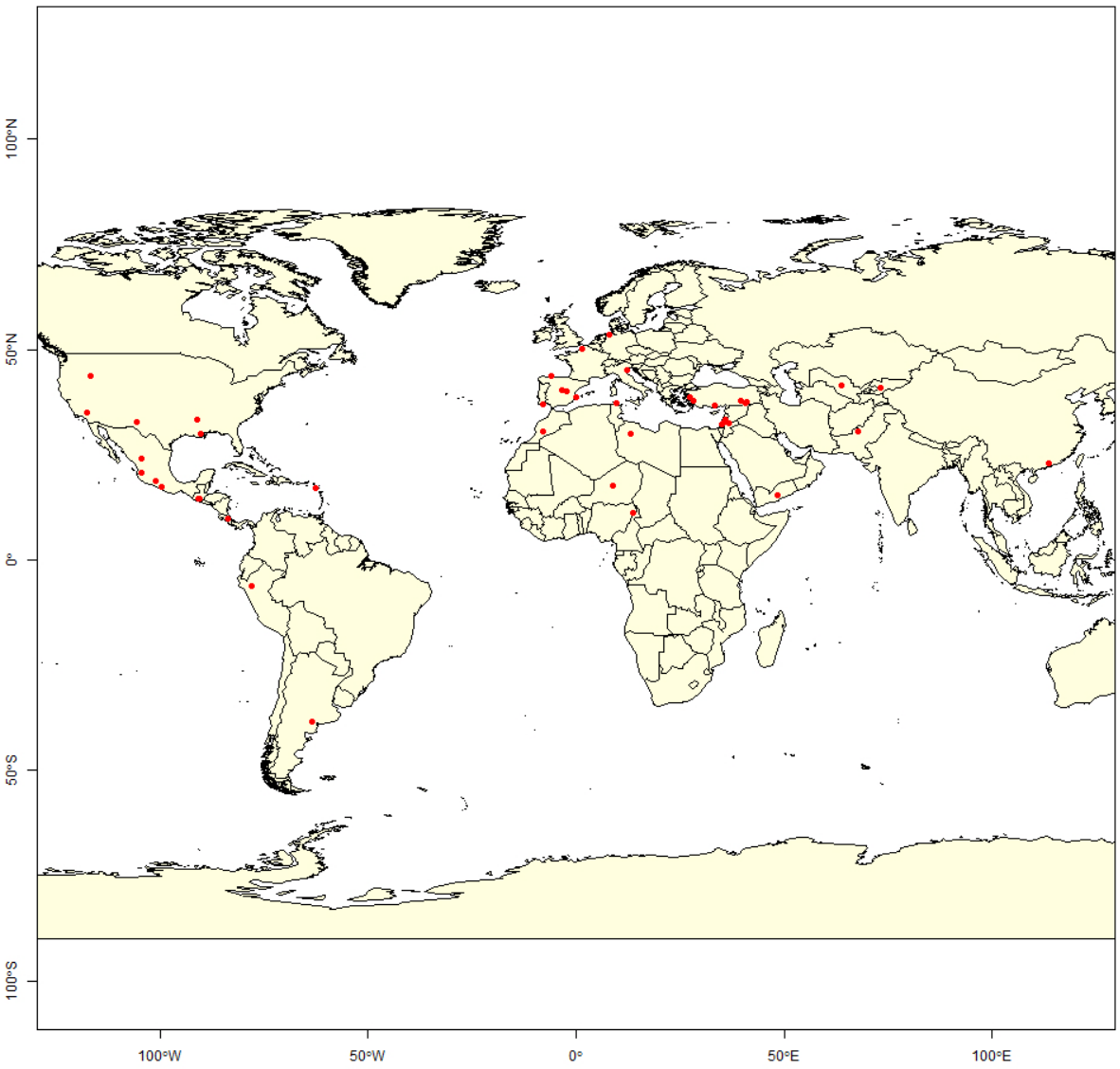


Figure S2: Phylogenetic diversity of the crop species of this project. The topology displayed was obtained from the maximally resolved seed plant tree available in Phylomatic (www.phylodiversity.net/phylomatic). For simplicity, branch lengths are not proportional to time since evolutionary divergence.

Figure S3: Bisector plots of the score of each crop on the 4 PCA axes associated with trait variation in Size (A), Competitive ability for light (B), Leaf economics (C) and Root economics (D). Each dot in the scatters is the wild (x-axis) and domesticated (y-axis) score for each crop belonging to the extensive experiment (30 crop species). Crops above y=x line showed higher PCA score in the domesticated than in the wild ancestor accession, and vice versa for crops below y=x.

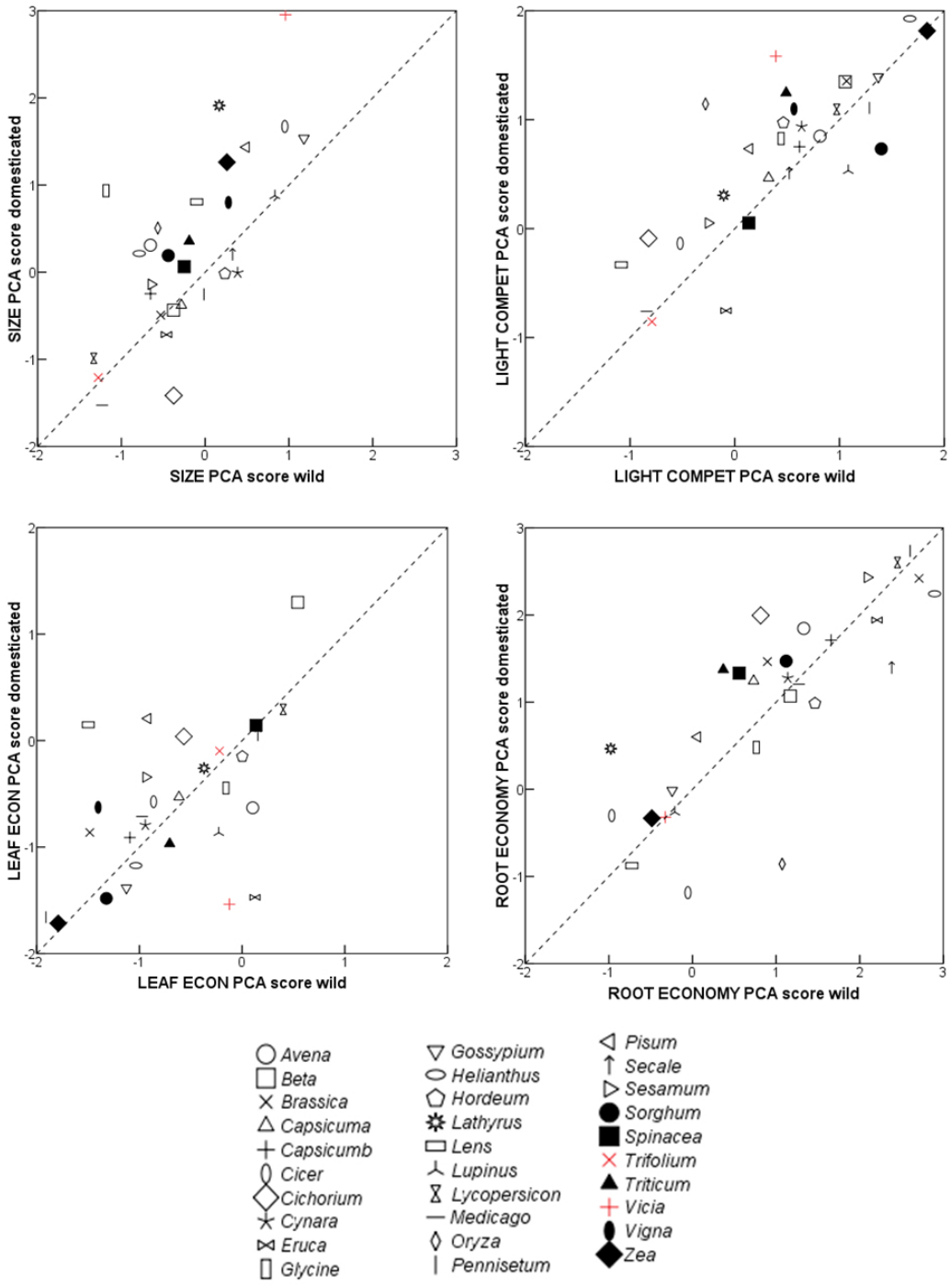


Figure S4: Bisector plots of PCA scores of three wild (x-axis) and three landrace (y-axis) accessions of each of the six crop species investigated in more detail.

Variables and reference lines as in Fig. S3.

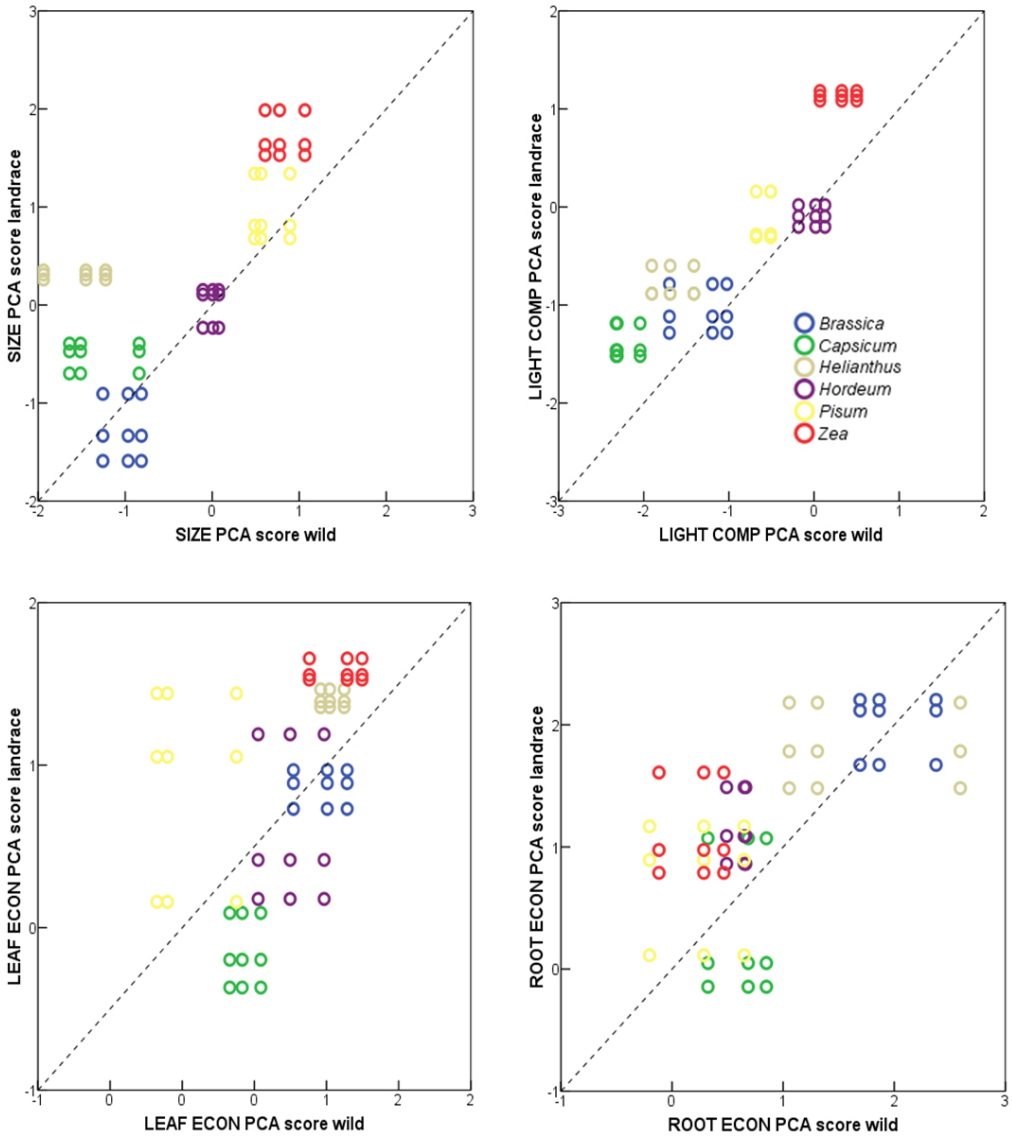


Figure S5: Bisector plots of PCA scores of three landrace (x-axis) and three improved (y-axis) accessions of each of the six crop species investigated in more detail. Variables and reference lines as in Fig. S3.

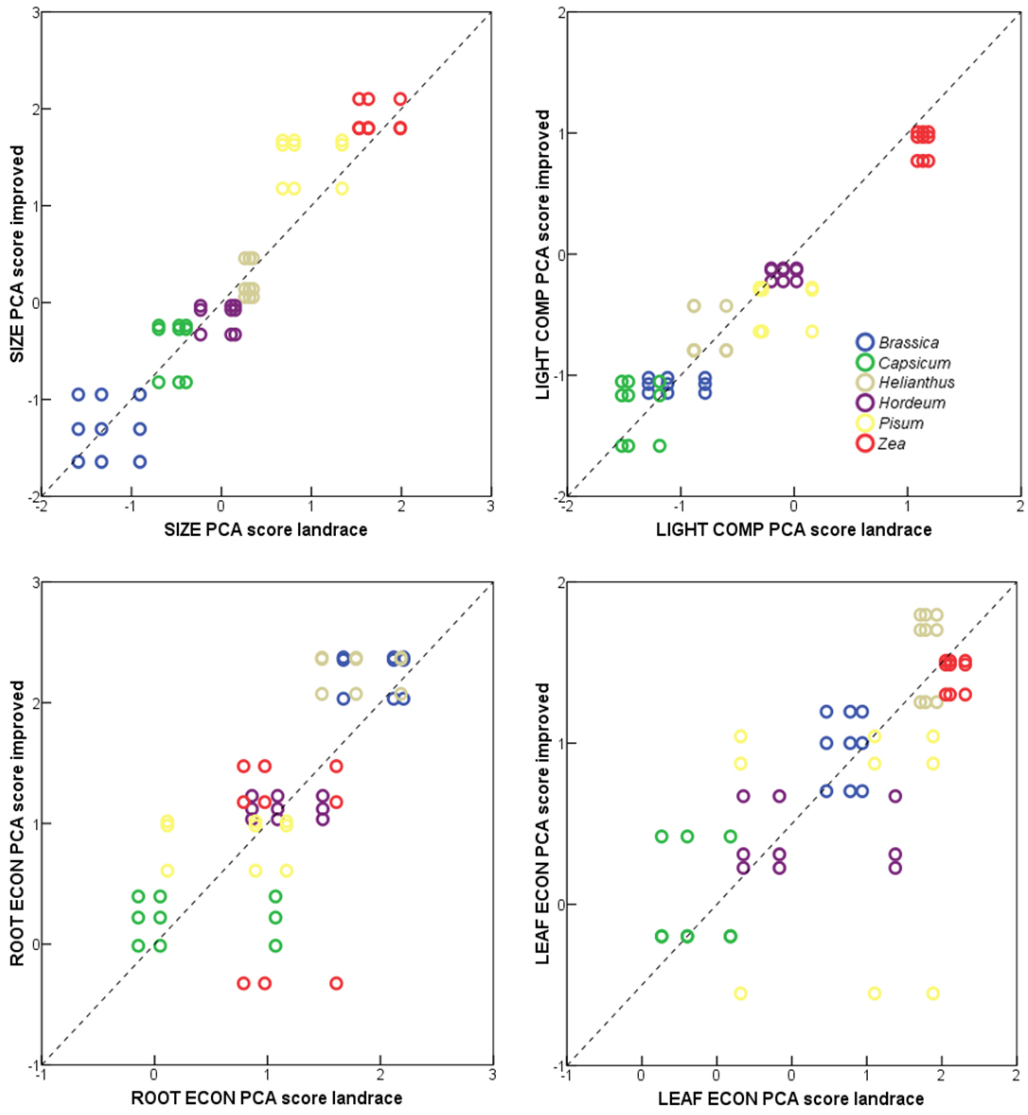
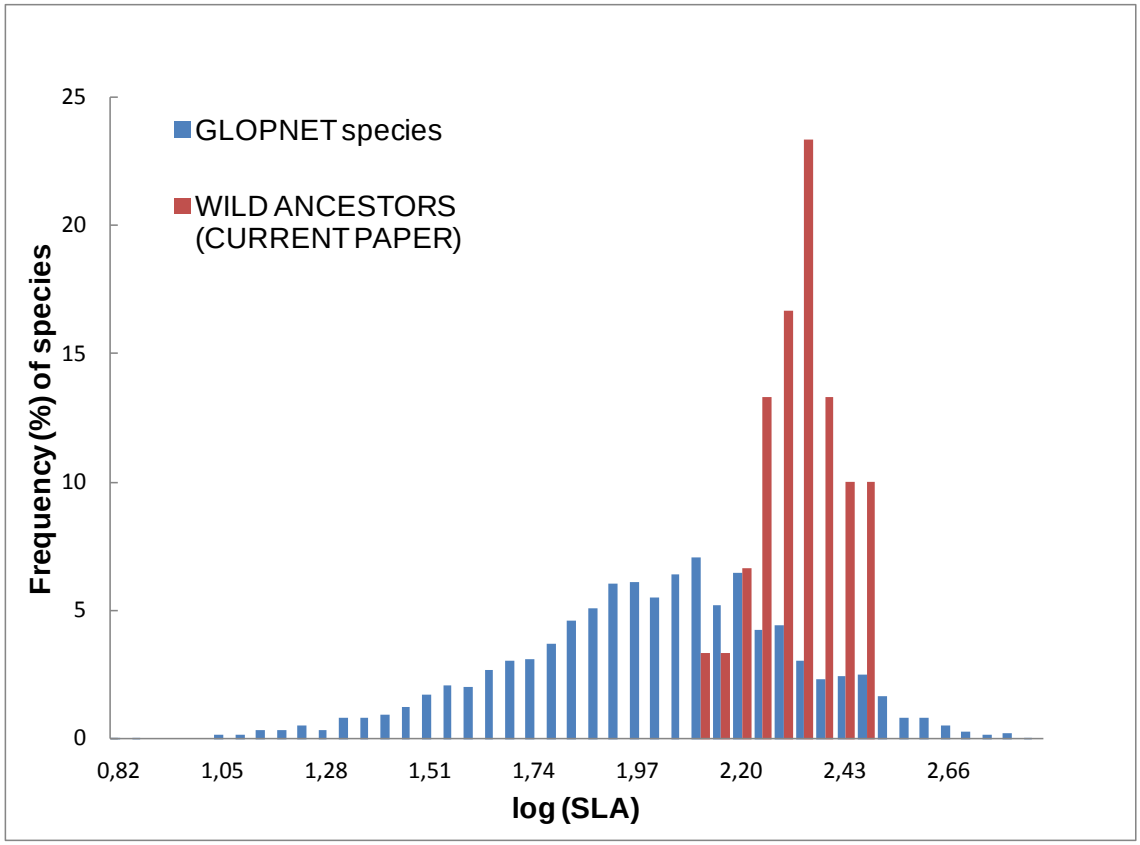


Figure S6: Frequency distribution of Specific Leaf Area in the Glopnet database *vs* that of wild accessions of the current paper. Y-axis is percentage of total species in each dataset. Glopnet data represents global variation in SLA, obtained from Supplementary Material of Wright et al. Nature 428, 821-827 at http://www.nature.com/nature/journal/v428/n6985/suppinfo/nature02403.html. Wild accessions of current paper dataset: n = 48 data, from 30 species). Glopnet dataset: n = 1958 data, from 1556 species.

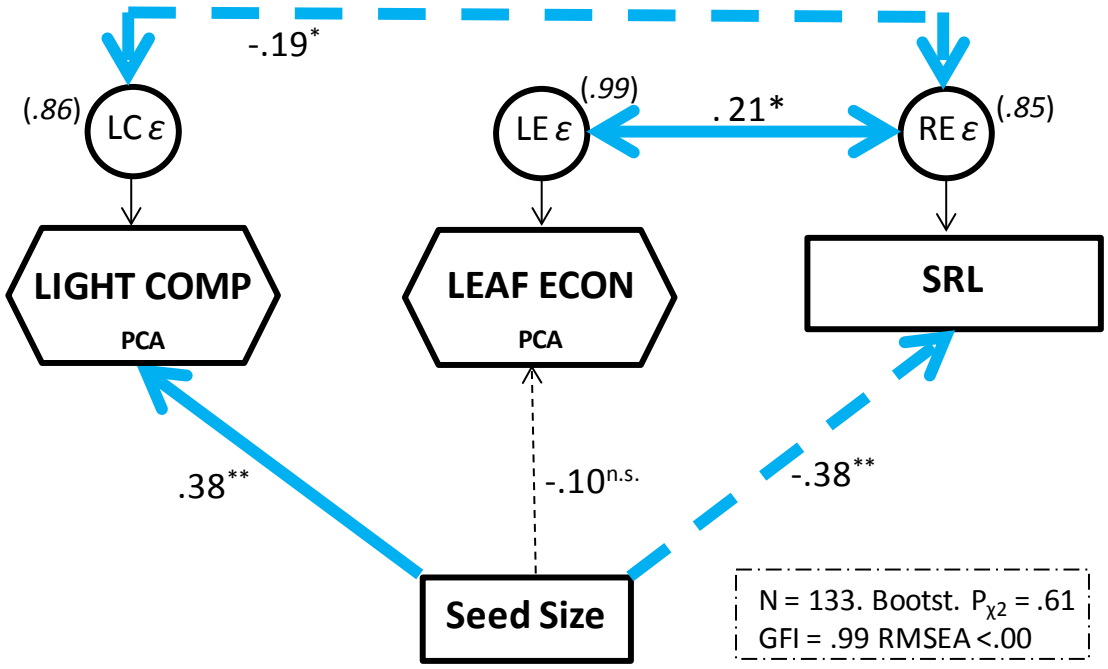


4) SUPPLEMENTARY STRUCTURAL EQUATION MODELLING ANALYSES

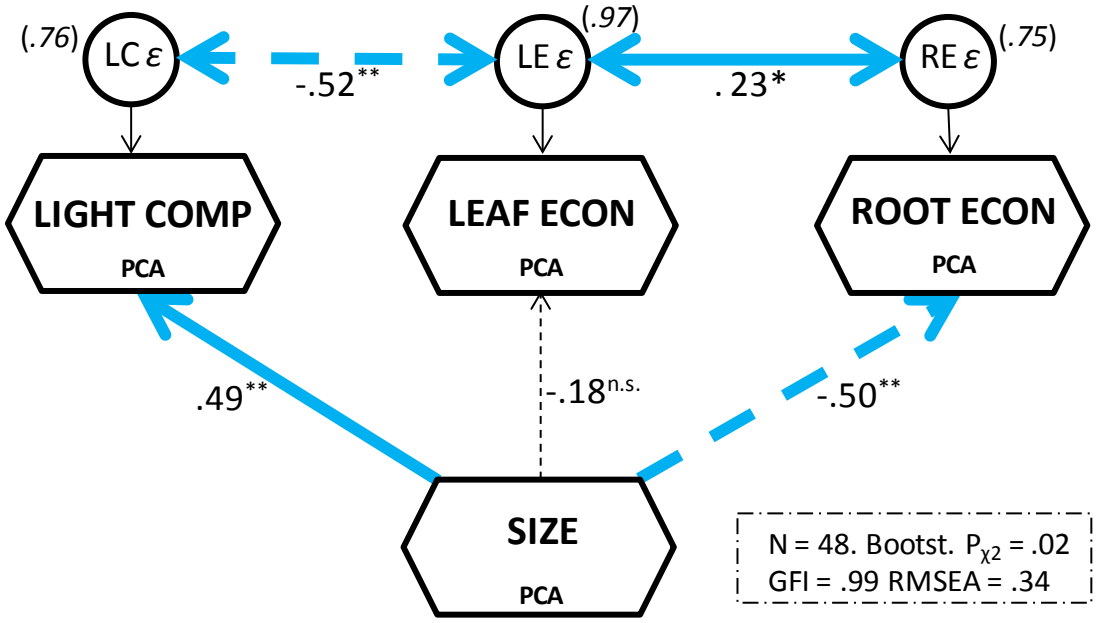
General validity of the inter-trait relationships depicted in Figure 2 of the main body of the paper

We based our approach to the analysis of coordinated evolution among traits with domestication on the conceptual model of among species inter-trait relationships that is described above and supported by literature and fit to our dataset (Figure 2 of the main body of the paper).

We tested the general validity of our general model making use of the only dataset that we found in the literature that was suitable for a direct, independent, validation of our model. This dataset was taken from Laughlin et al. Functional Ecology 24: 493–501. In that paper Laughlin and colleagues make use of a 133 herbaceous species dataset to test predictions of Leaf-Height-Seed (LHS) plant strategy scheme. Data on all our traits, except plant dry mass and density of fine roots, were available for the 133 herbaceous species. We processed this database using the same procedures as with our datasets. Data were log-scaled. Dimensionality was reduced through PCA for light competition and leaf economics traits (not for plant size and root economics, for which data were only available for one trait). Then, SEM fitting was performed as for SEM models in the main body of our paper. Results of model fit and parameter estimates are shown below. Model fit was very high, and this model closely resembles the pivotal structural lines of our model (Figure 2B). Size negatively impacts root economics, and increases light competitive ability. Leaf and root economics co-vary positively. However, the trade-off between light competitive ability and leaf economics does not receive support here. Given the differences between datasets in species and parameters measured, the resemblance between the two models is remarkable and indicates that the underlying model structure has broad applicability among species.



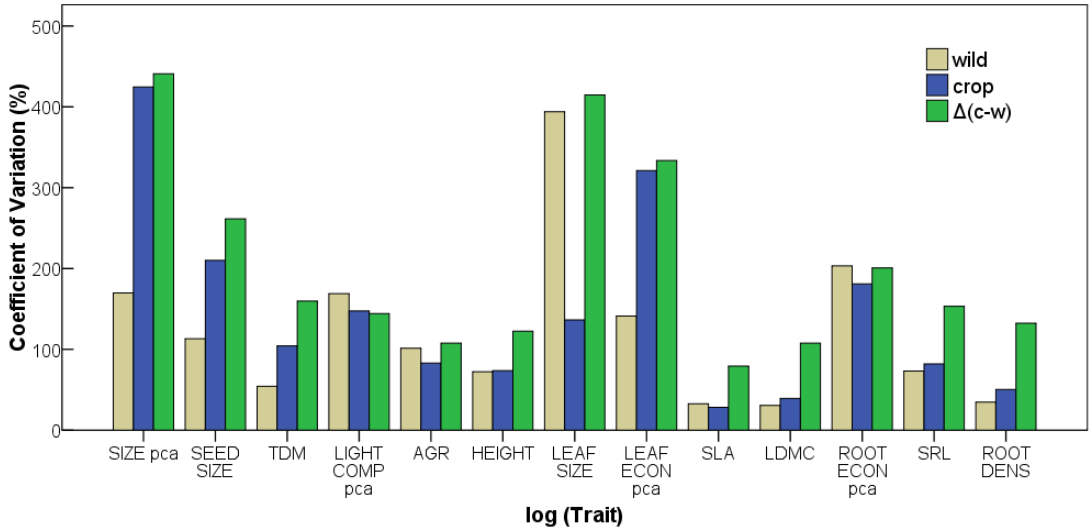
We further tested the soundness of our general model by including only accessions of our dataset that qualified as wilds, using the same procedures as with the complete dataset of wild species, landraces, and crops. Technical details on model fit followed the procedures described for SEM analyses in the main body of the paper and in Supplementary Full Materials and Methods. Inter-trait relationships among traits remained similar in magnitude, directionality and statistical significance to those depicted in the complete model using all accessions (Fig. 2). This reinforces the validity of that model, and its general insensitivity to the inclusion/exclusion of non-wild accessions. Overall model fit, in this case, was lower than for the complete dataset.



5) SUPPLEMENTARY PHENOTYPIC INTEGRATION ANALYSES

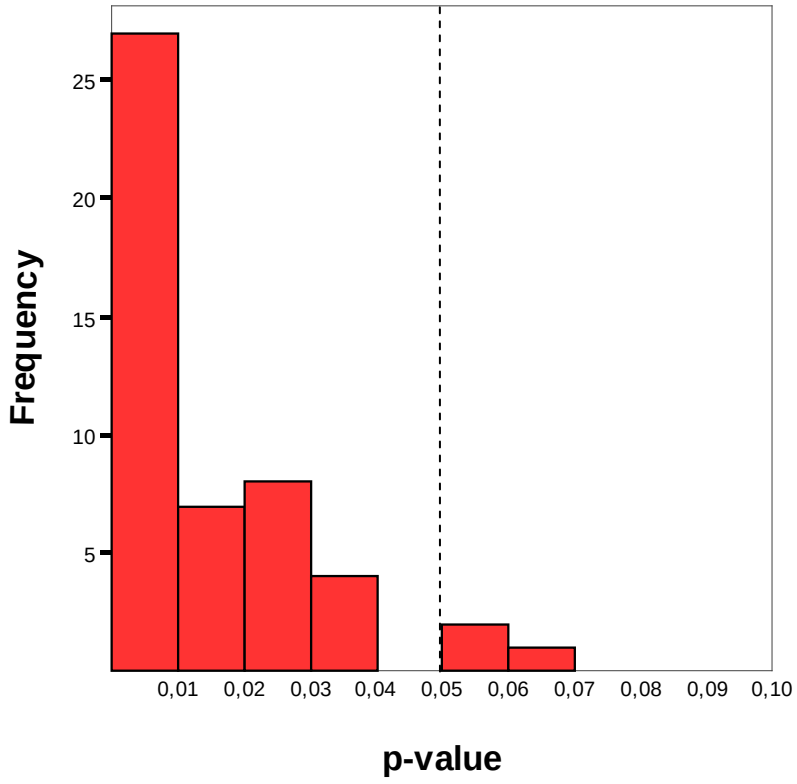
Addressing the problem of the effect of heterogeneous sample sizes on the number of significant correlations

The several models and correlation matrices evaluated in this paper differ in sample size. For instance, our "general" model and corresponding correlation matrix was built out of data for 114 accessions, whereas the "$\Delta_{C\text{-}W}$Trait", model resulted from 30 pairs of crop species and their wild relatives, and the "$\Delta_{LR\text{-}WI}$Trait" and "$\Delta_{IM\text{-}LR}$Trait" models were derived from a 54-accession database. The amount of paired data used to build a correlation-regression model affects several of the resulting statistics of model fit [40]. On the one hand, magnitude of slopes, or of correlation coefficients, does not consistently increase or decrease with varying sample size, but they become unstable at very low sample sizes, or when there are large variations in the ranges of focal variables [40,41]. The effects of low sample size on the magnitude of correlation coefficients are less predictable, but decreasing ranges of variation may consistently induce lower coefficients. In the Figure below we plot coefficients of variation for cultivated and wild accessions, and also for the within-crop difference between crops and wilds. Variation was of similar magnitude, or even higher, for $\Delta_{C\text{-}W}$Trait. Therefore, we reject the possibility that differences in the magnitude of correlation coefficients are due to contrasting degree of variability among datasets.



Regarding sample size and statistical significance tests of regression-correlation model parameters, tests of significance become vulnerable to type I errors as sample size increases [13]. We specifically included statistical significance as one of our components of phenotypic integration in Aster models (see Full Materials and Methods above). Thus, the differences between phenotypic integration levels among the two upper models shown in Figure 4 in the main body of the paper might arise, in part, from the effect of sample size on statistical significance of correlations.

To address this concern, we developed sets of comparisons of correlation matrices built from subsamples of equal sample size. For comparing the 114 "complete" dataset *vs* the 30 "$\Delta_{C\text{-}W}$Trait", we randomly generated fifty 30-sample sets from the 114-sample large dataset. Then, we fitted 50 Aster models to each of those 50 datasets and compared them through log-likelihood ratio tests, model-by-model, to the 30 "$\Delta_{C\text{-}W}$Trait" dataset. 45 out of 50 log-likelihood ratio tests remained significant when sample size was forced to be equal among the datasets under comparison (see Figure below). In light of those results, we assert that the differences among "Datasets" shown in the paper are not due to contrasting sample sizes.



6) REFERENCES CITED IN SUPPLEMENTARY MATERIAL

1. Hancock, J. 2004 *Plant evolution and the origin of crop species*. Cambridge, MA: CABI Publishing.

2. Hodgson, J., Wilson, P., Hund, R., Grime, J. & Thompson, K. 1999 Allocating C-S-R plant functional types: a soft approach to a hard problem. *Oikos* **85**, 283–294.

3. Pérez-Harguindeguy, N. et al. 2013 New handbook for standardised measurement of plant functional traits worldwide. *Aust. J. Bot.* **61**, 167–234.

4. Wilson, P., Thompson, K. & Hodgson, J. 1999 Specific leaf area and leaf dry matter content as alternative predictors of plant strategies. *New Phytol.* **143**, 155–162.

5. Birouste, M., Zamora-Ledezma, E., Bossard, C., Pérez-Ramos, I. M. & Roumet, C. 2013 Measurement of fine root tissue density: a comparison of three methods reveals the potential of root dry matter content. *Plant Soil* **374**, 299–313. (doi:10.1007/s11104-013-1874-y)

6. Westoby, M., Falster, D. S., Moles, A. T., Vesk, P. a. & Wright, I. J. 2002 PLANT ECOLOGICAL STRATEGIES: Some Leading Dimensions of Variation Between Species. *Annu. Rev. Ecol. Syst.* **33**, 125–159. (doi:10.1146/annurev.ecolsys.33.010802.150452)

7. Bouma, T. J., Nielsen, K. L. & Koutstaal, B. 2000 Sample preparation and scanning protocol for computerised analysis of root length and diameter. *Plant Soil* **218**, 185–196.

8. Himmelbauer, M. L., Loiskandl, W. & Kastanek, F. 2004 Estimating length, average diameter and surface area of roots using two different Image analyses systems. *Plant Soil* **260**, 111–120. (doi:10.1023/B:PLSO.0000030171.28821.55)

9. Nakagawa, S. & Freckleton, R. P. 2008 Missing inaction: the dangers of ignoring missing data. *Trends Ecol. Evol.* **23**, 592–6. (doi:10.1016/j.tree.2008.06.014)

10. Arbuckle, J. L. 2007 *Amos 18 User`s Guide.* Chicago, USA.:

11. Wright, I. J. et al. 2004 The worldwide leaf economics spectrum. *Nature* **428**, 821–7. (doi:10.1038/nature02403)

12. Anderson, M. J. & Ter Braak, C. J. F. 2003 Permutation tests for multi-factorial analysis of variance. *J. Stat. Comput. Simul.* **73**, 85–113.

13. Shipley, B. 2002 *Cause and correlation in Biology: A user's guide to path analysis, structural equations and causal inference.* Cambridge, UK.: Cambridge Univeristy Press.

14. Falster, D. S. & Westoby, M. 2003 Plant height and evolutionary games. *Trends Ecol. Evol.* **18**, 337–343. (doi:10.1016/S0169-5347(03)00061-2)

15. Vasseur, F., Violle, C., Enquist, B. J., Granier, C. & Vile, D. 2012 A common genetic basis to the origin of the leaf economics spectrum and metabolic scaling allometry. *Ecol. Lett.* **15**, 1149–1157. (doi:10.1111/j.1461-0248.2012.01839.x)

16. Tjoelker, M. G., Craine, J. M., Wedin, D., Reich, P. B. & Tilman, D. 2005 Linking leaf and root trait syndromes among 39 grassland and savannah species. *New Phytol.* **167**, 493–508. (doi:10.1111/j.1469-8137.2005.01428.x)

17. Ryser, P. 1996 The importance of tissue density for growth and lifespan of leaves and roots: a comparison of five ecologically contrasting grasses. *Funct. Ecol.* **10**, 717–723.

18. Jakobsson, A. & Eriksson, O. 2000 A comparative study of seed number, seed size, seedling size and recruitment in grassland plants. *Oikos* **88**, 494–502.

19. Moles, A. T., Ackerly, D. D., Webb, C. O., Tweddle, J. C., Dickie, J. B., Pitman, A. J. & Westoby, M. 2005 Factors that shape seed mass evolution. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10540–4. (doi:10.1073/pnas.0501473102)

20. West, G. B., Brown, J. H. & Enquist, B. J. 1997 A General Model for the Origin of Allometric Scaling Laws in Biology. *Science (80-.).* **276**, 122–126. (doi:10.1126/science.276.5309.122)

21. Niklas, K. J., Cobb, E. D., Niinemets, U., Reich, P. B., Sellin, A., Shipley, B. & Wright, I. J. 2007 "Diminishing returns" in the scaling of functional leaf traits across and within species groups. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 8891–6. (doi:10.1073/pnas.0701135104)

22. Grime, J. P. 2001 *Plant Strategies, Vegetation Processes and Ecosystem Properties*. 2nd edn. Chichester, UK: John Wiley & Sons.

23. Chapin, F. S. 1980 The mineral nutrition of wild plants. *Annu. Rev. Ecol. Syst.* **11**, 233–260.

24. Craine, J. M., Lee, W. G., Bond, W. J., Williams, R. J. & Johnson, L. C. 2005 Environmental contraints on a global relationship among leaf and root traits of grasses. *Ecology* **86**, 12–19.

25. Kembel, S. W. & Cahill, J. F. 2011 Independent evolution of leaf and root traits within and among temperate grassland plant communities. *PLoS One* **6**, e19992. (doi:10.1371/journal.pone.0019992)

26. Lloyd, J., Bloomfield, K., Domingues, T. & Farquhar, G. 2013 Photosynthetically relevant foliar traits correlating better on a mass vs an area basis : of ecophysiological relevance or just a case of mathematical imperatives and statistical quicksand? *New Phytol.* **199**, 311–321.

27. Grace, J. 2006 *Structural Equation Modeling and Natural Systems.* Cambridge, UK: Cambridge University Press.

28. Bollen, K. & Stine, R. 1992 Bootstrapping Goodness-of-Fit Measures in Structural Equation Models. *Sociol. Methods Res.* **21**, 205–229.

29. Bentler, P. 1989 *EQS structural equations program manual.* Los Angeles, USA: BMDP Statistical Software.

30. Schermelleh-engel, K., Moosbrugger, H. & Müller, H. 2003 Evaluating the Fit of Structural Equation Models : Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods Psychol. Res. Online* **8**, 23–74.

31. Grace, J. & Bollen, K. 2005 Interpreting the results from multiple regression and structural equation models. *Bull. Ecol. Soc. Am.* , 283–313.

32. Chreverud, J., Wagner, G. & Dow, M. 1989 Methods for the comparative analysis of variation patterns. *Syst. Zool.* **38**, 201–213.

33. Armbruster, W. S., Pélabon, C., Bolstad, G. H. & Hansen, T. F. 2014 Integrated phenotypes : understanding trait covariation in plants and animals. *Phil Trans R Soc* **369**, 20130245.

34. Schlichting, C. 1989 Phenotypic plasticity in Phlox. II. Plasticity of character correlations. *Oecologia* **78**, 496–501.

35. Pigliucci, M. & Marlow, E. 2001 Differentiation for flowering time and phenotypic integration in Arabidopsis thaliana in response to season length and vernalization. *Oecologia* **127**, 501–508.

36. Pigliucci, M. 2002 Touchy and Bushy : Phenotypic Plasticity and Integration in Response to Wind Stimulation in Arabidopsis thaliana. *Int. J. Plant Sci.* **163**, 399–408.

37. Shaw, R. G., Geyer, C. J., Wagenius, S., Hangelbroek, H. H. & Etterson, J. R. 2008 Unifying life-history analyses for inference of fitness and population growth. *Am. Nat.* **172**, E35–47. (doi:10.1086/588063)

38. Geyer, C. J., Wagenius, S. & Shaw, R. G. 2007 Aster models for life history analysis. *Biometrika* **94**, 415–426. (doi:10.1093/biomet/asm030)

39. R Core Team 2013 R: A Language and Environment for Statistical Computing.

40. Kelley, K. & Maxwell, S. E. 2003 Sample size for multiple regression: obtaining regression coefficients that are accurate, not simply significant. *Psychol. Methods* **8**, 305–21. (doi:10.1037/1082-989X.8.3.305)

41. Baumgartner, T. a. & Chung, H. 2001 Confidence Limits for Intraclass Reliability Coefficients. *Meas. Phys. Educ. Exerc. Sci.* **5**, 179–188. (doi:10.1207/S15327841MPEE0503_4)