Evolution of crop phenotypic spaces through domestication

Arthur Wojcik[1,2], Harry Belcram[1], Agnès Rousselet[1], Manon Bouët[2], Marie Brault[2], Pierre Serin[3], Augustin Desprez[1], Renaud Rincent[1], Andreas Peil[4], Karine Henry[5], Cécile Marchal[6], Thierry Lacombe[7], Sylvain Glémin[8], Karine Alix[1], Pierre R. Gérard[1], Domenica Manicacci[1], Yves Vigouroux[3], Catherine Dogimont[2], Maud I. Tenaillon[1]

[1] GQE - Le Moulon, Université Paris-Saclay, INRAE, CNRS, AgroParisTech, Gif-sur-Yvette, France

[2] GAFL - Génétique et Amélioration des Fruits et Légumes, INRAE, Avignon, France

[3] UMR DIADE, Université de Montpellier, IRD, CIRAD, Montpellier, France

[4] Julius Kühn-Institut, Pillnitz, Germany

[5] Florimond-Desprez, Cappelle-en-Pévèle, France

[6] Grapevine Biological Resources Center, INRAE, Unité Expérimentale Domaine de Vassal, University of Montpellier, 34340 Marseillan

[7] Institut Agro/Montpellier Supagro - UMR AGAP, Montpellier, France

[8] Laboratoire ECOBIO, UMR 6553 UR1-CNRS, Rennes, France

Corresponding author:

Maud I. Tenaillon, maud.tenaillon@inrae.fr, ORCID 0000-0002-0867-3678

Keywords: Domestication index, Divergent selection, Phenotypic convergence, Life-history traits, Near infrared spectra, Trait correlations

Summary

- We used domestication as an *in vivo* replicated experiment to investigate how divergent selection has shaped the evolution of multivariate phenotypic spaces.
- We measured 11 to 57 qualitative and quantitative traits in 13 species, either unique or shared between species, and established a framework for cross-species comparisons.
- Our results revealed significant convergence that translated into a cross-species domestication syndrome. Most species exhibited a reduction of the multivariate phenotypic space during domestication. We demonstrated that Near Infrared spectra measured on leaves evolve through processes unrelated with domestication, enabling its use as a control for sampling effects across species. Building on this, we developed a multivariate Phenotypic Divergence Index (mPDI) to rank species by the extent of phenotypic divergence under domestication. We found a high disjunction of wild and domestic phenotypic spaces in all species. Neither the mPDI nor the relative size of wild versus domestic multivariate phenotypic spaces was influenced by the domestication timing or mating system. Lastly, we observed a progressive decoupling of trait correlations with increasing time since domestication.
- In addition to introducing a new index that can be applied across species for comparative purposes, our study uncovers recurring patterns shared among species, pointing to general principles underlying plant domestication.

Introduction

Domestication is a process through which organisms evolve in response to human-altered habitats (Lord et al., 2025). One of the outcomes of this process is phenotypic divergence between the forms, resulting from disruptive selection between wild populations shaped by natural selection in their native habitats and domestic populations shaped by the combined effects of natural selection and adaptation to anthropogenic niches (Glémin & Bataillon, 2009; Ross-Ibarra et al., 2007). From an evolutionary standpoint, plant domestication is a recent process that began during the neolithic revolution (Weisdorf, 2005) with the earliest known domestications occurring ~12,000 years ago in the Fertile Crescent (Barker, 2006). Although debates persist about the intentionality of early selection (Abbo & Gopher, 2020), archaeological evidence from the few crop species for which extensive temporal series are available supports a protracted model of domestication, with the slow fixation of domestic traits over thousands of generations as shown in einkorn wheat, rice and barley (Purugganan & Fuller, 2011). After domestication, crop expanded through trade and seed exchange, leading to the emergence of a genetic structuring into groups, well described in e.g. maize or pearl millet (Burgarella et al., 2018; Tenaillon & Charcosset, 2011).

Over 2500 plants species belonging to 173 botanical families have undergone cultivation or domestication — providing an upper bound for the number of domesticated plants (reviewed in Dirzo & Raven, 2003). Human selection has targeted a diversity of organs (e.g., flower, fruit, seed, tuber) and usages (e.g. food, fiber, ornamentals, medicine). Phenotypic changes that often arise through parallel evolution across species under domestication are collectively referred to as the phenotypic domestication syndrome. These changes include traits that facilitate cultivation, harvesting, and consumption and differ between crop-types e.g., between herbs and woody plants or between vegetatively and sexually propagated crops, for example (Fuller et al., 2023). Domestication traits are well-documented in annual crops. In cereals, for example, they include reduced tillering, loss of seed shattering, and increased ear and seed size (Glémin & Bataillon, 2009; Tenaillon & Manicacci, 2011). More rarely, they involve modifications of life-history traits such as an increase in selfing rate in sunflower, cacao or grapevine (Dempewolf et al., 2012). Compared to annual species, perennial crops are characterized by extended juvenile phases, fewer sexual generations, and a reliance on clonal propagation, which together make them younger domesticates (Gaut et al., 2015). Just like for annual seed crops, domestication traits also point to simplified plant structure, increased edible organs, loss of bitter compounds and reduced sexual reproduction for vegetatively reproduced

crop (Denham et al., 2020; Miller & Gross, 2011). Hence, convergence of domestication traits across species within-crop types, well documented in cereals (Tenaillon & Manicacci, 2011), seems to extend to more distant taxa (Meyer et al., 2012) but has yet to be tested.

The phenotypic domestication syndrome is accompanied by a molecular domestication syndrome (Burban et al., 2022), primarily characterized by a loss of genetic diversity due to bottlenecks and linked selection at key domestication genes as exemplified in maize (Wright et al., 2005; Yamasaki et al., 2005) and other crops (reviewed in Gaut et al., 2015). Analysis of maize landraces suggests that these combined processes have resulted in a loss of genetic variance at traits targeted by selection (Yang et al., 2019). Thus, one might expect a reduction of the domestic phenotypic space— that is, a reduction in the range of observable trait variation expressed by cultivated forms—as divergence increases. This hypothesis requires formal validation through a comparative analysis across multiple species, based on rigorous sampling that limits the potential inflating effects of genetic structuring and local adaptation on domestic phenotypic spaces following agriculture diffusion and modern breeding. Indeed, crop expansion led to adaptation to novel environments, and selection for local preferences (Bradshaw, 2016; Meyer & Purugganan, 2013), and thus to a range of new phenotypes, not directly linked to domestication *per se*.

Another intriguing question arises from the concerted evolution of domestication traits whose determinism involves complex gene regulatory networks (Studer et al., 2017). Several studies indicate that domestication has led to a deep rewiring of these networks as shown by the reorganization of co-expression networks (e.g., in cotton (Rapp et al., 2010), common bean (Bellucci et al., 2014), and tomato (Sauvage et al., 2017)). Simulations suggest that the combined effects of demography and selection during domestication result in a redistribution of genetic correlations, leading to less distinct clusters and a moderate increase of these correlations (Burban et al., 2022). Studies on maize and teosintes indeed showed an overall change in the structure of the correlation matrix, but the findings were less clear for reproductive traits directly affected by domestication, with some analyses indicating conservation of correlations in the two forms. Contrary to the model predictions, teosintes generally exhibit more significant and stronger correlations than maize (Yang et al., 2019). Model predictions therefore need to be further tested across a wider range of species.

In sum, although individual domestication traits have now been described in many systems, our understanding of the evolution of phenotypic spaces through the domestication process, and the underlying key evolutionary mechanisms, remains superficial. This is at least partly due to a lack of unifying statistical framework able to integrate data from different

species, where various and often distinct traits are experimentally measured, including traits of different types (quantitative and qualitative) and scales. The approach we propose here is to build on multivariate analysis to transform measured traits into statistically comparable objects across species; and to use a control trait to account for sampling heterogeneity across species providing a baseline of wild and domestic phenotypic divergence largely independent of domestication. From there, we derived a multivariate Phenotypic Divergence Index (mPDI), which quantifies the degree of phenotypic divergence between wild and domestic forms in each species.

By applying this unifying framework to 13 wild/domesticated pairs (13 species) including a range of domestication dates from ancient to more recent, different mating systems - outcrossers and selfers, and various life cycles from annual to perennial, we addressed the following questions: (1) Is there a significant phenotypic convergence across species? (2) How did phenotypic spaces evolve during domestication? (3) Do domestication dates and mating systems influence the mPDI? (4) Are there consistent patterns in trait correlations throughout the domestication process?

Material and Methods

Plant Material

Our study focused on 13 domestic species and the wild ancestors from which they were derived through domestication: African rice (*Oryza glaberrima*), apple (*Malus domestica*), cabbage (*Brassica oleracea*), common bean (*Phaseolus vulgaris*), eggplant (*Solanum melongena*), einkorn wheat (*Triticum monoccocum*), foxtail millet (*Setaria italica*), grape (*Vitis vinifera*), maize (*Zea mays*), melon (*Cucumis melo*), pearl millet (*Pennisetum glaucum*), sugar beet (*Beta vulgaris*) and tomato (*Solanum lycopersicum*) spanning eight botanical families (Fig. 1a, Table S1). These 13 wild/domestic pairs were chosen because they were all diploids and span a broad timeline of domestication, from the most ancient domestication of einkorn wheat around 9,500 years before present (YBP) to the most recent domestication of eggplant 2000 YBP (Table S1) that were converted in generations (Fig. 2a). Note that for species where two domestication centers were recorded, we focused on a single domestication center and the corresponding wild gene pool: South America for tomato, Mesoamerica for common bean, and Sudan for melon. For beet, we focused on sugar beet whose domestication happened later than for leaf beet (Table S1).

These 13 species furthermore exhibit a diversity of life-cycles from annual, biennial (e.g., cabbage) to perennial (e.g., grapevine), with sugar beet transitioning from annual to biennial during domestication (Table S1, Fig. 2a). They also display a range of reproductive systems from highly self-fertilizing species (selfers) such as foxtail millet to obligate outcrossers such as apple tree (Table S1). Most species have hermaphroditic flowers, except for maize which is monoecious, grapevine that transitioned from dioecy to monoecy during domestication, and melon from monoecy to andromonoecy (Table S1).

To facilitate a comparative analysis across species, we aimed to apply consistent sampling criteria within each species. For wild/domestic pairs, we chose 17 to 20 wild populations (accessions) to represent the wild gene pool and 16 to 22 landraces (accessions) to represent the domestic gene pool (Table S2) except for melon and eggplant for which the wild sampling encompassed 8 and 13 populations, respectively. For the domestic gene pool, we focused on landraces in order to capture the domestication process itself while minimizing the effects of more recent breeding. In addition, we minimized genetic structuring among individuals to avoid confounding effects of genetic group in phenotypic analyses. To this end, we relied on established domestication centers for sampling both wild populations and landraces in nine of the systems (Table S1). Perennials were exceptions in this sampling design, as we relied on existing and accessible tree collections. For grapevine, whose domestication center is in South Caucasus, we focused on the western Mediterranean region, where a greater number of wild and domestic accessions were available. For apple, the domestic sample was confined to European traditional cultivars, while wild samples originate from Central Asia. For cabbage, where the exact European domestication center is still debated (Mabry et al., 2021), we focused on wild and domestic accessions from western Europe (France and United Kingdom).

Plant phenotyping

Except for seed germination rate which required to evaluate many seeds per accession, plant phenotyping was conducted using a three randomized-block design for all but four species (melon, cabbage, apple, grape), with each block containing a single individual from each wild population and each landrace for a total of maximum 120 individuals – 3x40 initial individuals (Table S3). For melon and cabbage, we measured a single individual per population/landrace (accession) in a randomized design (i.e., no blocks). For apples, we also evaluated a single individual per accession in a conservation orchard with trees of different ages. For grapes, one to five clones from a single individual per accession were phenotyped, and quantitative

186 variables were averaged across clones – qualitative variables being the same across clones. For
187 a given accession, the clones were planted in a single row, so there were no environmental
188 replicates (i.e, no blocks). Cabbage, grape, apple, einkorn wheat and foxtail millet were
189 phenotyped in the field, while all other species were phenotyped in a greenhouse as they
190 required day-length control for flowering (Table S3). Note that sugar beet was phenotyped in
191 controlled chamber where plant cycles were synchronized to ensure simultaneous flowering of
192 wild (annual) and domestic (biennial) individuals. All specific phenotyping conditions are
193 detailed in Table S3.
194 The number of traits measured within species varied widely, ranging from 11 traits in
195 apple to 57 traits in sugar beet (Fig. 1 and 3a), summing to 357 different traits across the 13
196 species (Fig. 3a). The traits are briefly described in Table S4 but a complete ontology can be
197 found at https://doi.org/10.57745/QWEKVK. Traits were evaluated from germination to
198 maturity and harvest. They included both common descriptors across most species, such as
199 plant height or grain size, as well as species-specific traits, such as the acidity of grape must or
200 the spiny or non-spiny leaves of eggplant (Fig. 1 and 3a).
201 The traits measured fell into three categories (1) continuous quantitative variables, such
202 as stem circumferences, pH and the Brix index (an estimate of the sugar content); (2) discrete
203 quantitative variables, such as the number of internodes or seeds; (3) qualitative variables, such
204 as flower colour or the relative position of styles within the flower. We used imaging software
205 to assess some of the traits including the number of pollen grains in grape flowers and several
206 root traits in sugar beet (Bucksch et al., 2014).
207
208 **NIRS acquisition and analysis**
209 We acquired Near InfraRed Spectra (NIRS) on leaves for all species except apple. In
210 addition, NIRS were collected on seeds for eight of the 13 species (Tables S4), specifically
211 those for which sufficient seed material was available. In brief, NIR spectroscopy works by
212 directing near-infrared light onto a sample and measuring the amount of light absorbed at
213 different wavelengths, returning spectra whose shape are influenced by the chemical
214 composition of tissues.
215 For NIRS phenotyping on seeds ($NIRS_{seed}$), we used a Perkin Elmer FT9700 NIR
216 spectrometer. For samples with a limited number of seeds, the light source was directed at the
217 seeds, and the acquisition was performed in complete darkness. To account for sample
218 heterogeneity within species, we performed four technical replicates per sample, mixing the

219 seeds between each measurement. Each spectrum was obtained as the mean of 32 scans, and
220 wavelengths ranged from 700 to 2500 nm with a step of 8 cm^{-1}, resulting on 1300 wavelengths.
221 For NIRS phenotyping of leaves, we harvested three leaves or leaf sections per
222 individual. We then oven-dried them for 48h at 40°C in coffee filters as drying does not
223 significantly affect spectra (Borraz-Martínez et al., 2019) and was more convenient for
224 transport. NIRS measurements were taken using a LapSpec 2500 NIR spectrometer. We
225 gathered three spectra per leaf, while carefully avoiding the main veins. Each spectrum was
226 obtained as the mean of 10 scans, and wavelengths ranged from 350 to 2500 nm with a step of
227 1 nm, resulting on 2150 wavelengths.
228 We prepared the data as follows, using recommended procedures in Rincent et al.
229 (2018) to control for additive and multiplicative effect of wavelength: (1) we loaded the spectra
230 using the *asdreader* R package (https://github.com/pierreroudier/asdreader); (2) we discarded
231 technical replicates from one individual whose deviation from the individual mean exceeded
232 twice the standard deviation over at least 30% of the spectrum – apparent outlier replicates
233 were also discarded; (3) we further normalized the spectra by sample, and computed their first
234 derivative using a Savitsky-Golay filter (Savitzky & Golay, 1964) as implemented in the *signal*
235 R package (https://cran.r-project.org/web/packages/signal/index.html) using a window size of
236 37 points and a filter order of two.
237 Because the wavelengths are numerous and not independent, we computed Principal
238 Coordinate Analyses (PCoA) for leaf and seed samples separately. We used the projections of
239 the individuals on the first two PCoA axes as two new independent traits. We evaluated
240 whether clustering obtained with NIRS effectively grouped the technical replicates of the same
241 individual. In order to do so, we conducted supervised hierarchical clustering of *g* groups,
242 where *g* represents the number of individuals. The grouping of technical replicates was
243 validated using the R package *flexclust* (Leisch, 2006) and the Rand index (Rand, 1971). The
244 Rand index, which measures clustering similarity, ranges from zero to one, with the latter
245 indicating perfect match between the observed and that predicted from individuals.
246
247 **Domestication syndrome characterization**
248 To describe the domestication syndrome in each species, we applied a trait-by-trait
249 approach. For quantitative traits, we tested differences in mean between forms using a mixed
250 linear model:
251 $X_{ijk} = \mu + B_i + F_j + A_k + (A \times F)_{jk} + \varepsilon_{ijk}$ (1)

where X_{ijk} represents the phenotypic value of one trait in block i, in form j and accession k; μ is the phenotypic mean; B_i is the fixed effect of block I; F_j is the fixed effect of form j; A_k is the random effect of accession k, with $A_k \sim \mathcal{N}(0, \sigma^2_k)$; $(A \times F)_{jk}$ is the random effect of the interaction between form j and accession k, with $(A \times F)_{jk} \sim \mathcal{N}(0, \sigma^2_j)$, σ^2_j being the specific variance for form j; ε_{ijk} is the residual error, with $\varepsilon_{ijk} \sim \mathcal{N}(0, \sigma 2_E)$.

When there was no replicate, we used a simple linear model:

$X_j = \mu + F_j + \varepsilon_j$ (2)

Where X_j represent the phenotypic value of a trait in form j; μ is the phenotypic mean; F_j is the fixed effect of form j; ε_j is the residual error, with $\varepsilon_j \sim \mathcal{N}(0, \sigma 2_E)$. We used Student's t-tests as implemented in the *lmertest* R package (Kuznetsova et al., 2017) to assess the significance of fixed effects. For qualitative traits we used Fisher exact tests to assess significant differences in the distribution. We calculated species-specific false discovery rate (FDR) q-values using the Benjamini–Hochberg procedure (Benjamini & Hochberg, 1995) based on the p-values for the form effect, applying a 5% threshold. We defined domestication-associated traits as those showing a significant main effect of form, i.e. traits that differed between domestic and wild forms. For graphical representation of trait values in each species, we constructed radarplot using the R packages *ggplot2* (Wickham, 2009) and *ggradar* (https://github.com/ricardo-bion/ggradar).

Domestication syndrome convergence

To test for convergence, we assessed whether species shared more domestication-associated traits than expected by chance. First, we determined the observed number of domestication-associated traits unique to a species, or shared by two, or three or more species. Second, we applied a resampling procedure to generate a null distribution under the hypothesis of random sharing of domestication-associated traits (i.e., no convergence). To account for the structure of our data, we resampled traits using a design that mimicked both the number of traits measured per species and the pattern of trait sharing across species. Of the 357 measured traits, we identified *n* domestication-associated traits. Our procedure involved performing 10,000 resampling iterations in which *n* traits were randomly drawn from the *t* traits and designated as domestication-associated traits. For each resampling, we generated a simulated distribution that we compared to the observed values to compute p-values. Convergence was supported if the number of traits shared across three or more species exceeded that expected by chance, as determined by the simulated distribution. Note that because NIRS relies on the

first two axes of a PCoA, with potentially different wavelengths contributing unequally to each axis across species, these axes are not homologous among species and were excluded from the convergence analysis.

Computation of phenotypic spaces disjunction

We measured the extent of disjunction between the phenotypic spaces of wild and domestic forms using the Pillai trace (Pillai, 1955) which quantifies how well grouping factors separate observations in multivariate data. We first performed a Factorial Analysis of Mixed Data (FAMD) on each dataset; we further considered all resulting axes as new multivariate variables. Pillai trace was then calculated in a MANOVA using the form (wild or domestic) as the explanatory variable and the FAMD axes as response variables. The resulting value ranged from zero, indicating complete overlap of the spaces, to one, indicating complete disjunction between the two forms.

Sampling heterogeneity between forms (wild and domestic) across species can bias the estimates of the disjunction between phenotypic spaces. To account for sampling heterogeneity—referring here to the fact that in some species wild and domesticated material may have diverged more strongly than in others due to factors such as distance from the domestication center (drift) or differing levels of crop improvement (selection) for example—we used $NIRS_{leaf}$ as a control. $NIRS_{leaf}$ is assumed to capture variation primarily unrelated to domestication. While certain wavelength may be shaped by selection during domestication, using the full spectrum of wavelengths collectively provides a baseline for processes not directly associated with domestication.

By computing Pillai trace on the two first axes of PCoA computed on $NIRS_{leaf}$, we were able to estimate the disjunction unrelated with domestication ($Pillai_{control}$). We then corrected the disjunction of other traits by subtracting $Pillai_{control}$ to obtain a corrected space disjunction. In the following, we refer to the resulting index as the multivariate Phenotypic Divergence Index (mPDI). We also computed the mPDI per unit of disjunction caused by processes unrelated with domestication, by dividing mPDI by $Pillai_{control}$.

We assessed correlations between mPDI and archaeological estimates of the domestication timing in generations (Table S1) using Kendall and Spearman correlation tests and tested the effect of mating system on the mPDI using a Student t-test.

319 **Computation of the of the multivariate phenotypic space**

320 We aimed to compare the relative size of the multivariate phenotypic space between
321 wild and domestic forms across all species pairs. As the determinant of the variance-covariance
322 matrix describing a space is related to its size, we used it as a measurement of the phenotypic
323 space hypervolume. To compute the variance-covariance matrix in each species, we (1)
324 normalized the quantitative traits across all individuals; (2) transformed qualitative traits by
325 decomposing them into their constituent levels (e.g., a trait with five levels became five binary
326 traits), scoring each as present or absent. Considering each form (wild and domestic) within
327 species, we further eliminated traits that were direct linear combinations of others. Finally, we
328 computed the determinant of the variance-covariance matrix of traits for each form.

329 Since we removed linear combination of traits separately for wild and domestic forms,
330 the sizes of the matrices were not necessarily equal, making direct comparisons of determinants
331 impossible. To address this, we normalized the determinants by taking the *r*-root, where *r* is
332 the highest dimension between the wild and domestic matrices. We then calculated the log ratio
333 of the size of the multivariate domesticated phenotypic space to that of the wild multivariate
334 phenotypic space. A ratio greater than one indicated that domesticated forms explore a broader
335 phenotypic space than their wild counterparts, while a ratio smaller than one indicated the
336 reverse. We tested the influence of the number of samples and the number of traits on the
337 hypervolume size by using Kendall and Spearman correlation tests. We used the same tests to
338 assess correlations between the log ratio of the phenotypic spaces and both, domestication
339 timing in generations and mPDI (Table S1).

340 Because we assumed $NIRS_{leaf}$ primarily captures variation not driven by domestication,
341 we hypothesized that the log ratio of domestic to wild space size based on $NIRS_{leaf}$ to correlate
342 with the corresponding ratio of neutral genomic diversity. To test this hypothesis, we
343 determined wild and domestic multivariate phenotypic spaces on $NIRS_{leaf}$ using all the axes of
344 the corresponding PCoA as variables. We then tested correlations with ratio of genomic
345 diversity taken from literature (Table S5) using Spearman and Kendall correlation tests.

346

347 **Similarity of trait matrices between wild and domestic forms**

348 For each form, we first calculated the mean value of each trait by averaging across
349 replicates (when multiple replicates were available, Table S3). Second, we computed the
350 pairwise phenotypic correlation matrices between traits using Spearman correlation
351 coefficients and, phenotypic variance-covariance matrices. We used a Student t-test to compare
352 the absolute mean values of correlations. We assessed the similarity of variance-covariance

353 matrices between forms using the random skewers method (Cheverud & Marroig, 2007), as
354 implemented in the *EvolQG* R package (Melo et al., 2016). This method is based on Lande's
355 equation (Lande, 1979), and evaluates how similarly two matrices respond to deformation. In
356 short, each matrix is multiplied by one random scaled deforming vector, producing two
357 matrices whose deformation can be described by a response vector. The correlation between
358 the two response vectors is measured. This process is repeated a large number of times (10,000
359 in our case), yielding the average response vector correlation. The advantage of the random
360 skewers method over a Mantel test, is that it provides a quantitative measure of the similarity
361 between the variance-covariance matrices of wild and domestic traits, enabling comparisons
362 across species.
363 We assessed correlations between the similarity of the matrices and archaeological
364 estimates of domestication timing in generations (Table S1) and mPDI with Kendall and
365 Spearman correlation tests.
366

367 ## Results

368 We measured 357 traits across 13 species, 189 of which were species-specific while 40
369 were shared by 2 to 10 species (Fig. 1a, Table S4). Such a diversity of species and traits poses
370 a statistical challenge for performing comparisons across species. In addition to trait-by-trait
371 analysis within species, we therefore developed a unifying statistical framework that includes
372 the establishment of a multivariate phenotypic divergence index (mPDI) and the use of specific
373 tools to compare the evolution of phenotypic spaces and trait correlations across species.
374

375 **Domestication syndrome**

376 We characterized the phenotypic domestication syndrome within species.
377 Domestication-associated traits were identified based on statistically significant differences in
378 trait values between wild and domestic forms. For example, in foxtail millet, we measured 15
379 traits, 12 of which were domestication-associated traits that differentiated wild and domestic
380 forms (Fig. 4a). Across systems, the percent of domestication-associated traits among all traits
381 ranged from 26% in melon and 35% in grapevine to 81% in maize and 89.7% in common bean
382 (Table S4; Figs S1-12).
383 In addition, as illustrated in foxtail millet for 8 out of 13 domestication-associated traits
384 (Fig. 4a), the domestic forms tended to display greater values than the wild forms, e.g., larger
385 leaves and spikes. This pattern was repeatedly observed across species extending to

386 domestication-associated traits related to plant and organ (leaf/spike/fruit/seed) size (Fig. 4a,
387 Figs S1-S12). We also found a number of qualitative domestication-associated traits such as in
388 common bean, where the colour, pattern and tendency for shattering in pods differed between
389 wild and domestic forms (Fig. S4, FrCol, FrPat and FrTension). Our findings also highlighted
390 new interesting domestication-associated traits, such as the length of flower petals in eggplant
391 or the increase in the number of secondary stems in common bean (Figs S1-S12). We
392 investigated seed/grain properties using NIRS, hypothesizing that domestication may have
393 influenced their composition. Interestingly, we found significant differences in NIRS profiles
394 between wild and domestic forms (either on the first axis of the PCoA – SNIRS1, or the second
395 – SNIRS 2 or both) in melon, einkorn wheat, foxtail millet, cabbage, maize, African rice,
396 common bean and pearl millet (Fig. 1b, Fig. 4b, Figs S1-12).
397
398 **Convergence of domestication-associated traits**
399 Among 78 species pairs, we found that 82% had at least one shared domestication-
400 associated trait (Fig. 3b). The average number of pairwise shared traits across all species was
401 2.6 and ranged from 0 – between African rice and eggplant or African rice and grapevine and
402 cabbage, to 7 – between maize and tomato or foxtail millet and maize (Fig. 3b). We formally
403 tested phenotypic convergence for domestication-associated traits. While the number of unique
404 domestication-associated traits to a given species did not significantly differ from expectation,
405 we found a significant enrichment of domestication-associated traits (convergence) shared by
406 three or more species (P-value=0.014, Fig. S13, Fig. S14) indicating convergence. Conversely,
407 this was accompanied by a deficit of traits shared by exactly two species (P-value=0.014, Fig.
408 S13, Fig. S14), consistent with the fixed total number of traits—where enrichment in one
409 category entails a corresponding reduction in another.
410 Our dataset was strongly biased toward *Poaceae* (5 of the 13 species), a family known
411 for phenotypic convergence of domestication-associated traits (Glémin & Bataillon, 2009). To
412 ensure that observed convergence was not driven by this bias, we repeated the analysis
413 including only one of the *Poaceae* species at a time. The results remained consistent (P-
414 values<0.01) with an enrichment of traits shared by three or more species irrespective of the
415 *Poaceae* species sampled. Altogether, our results indicate convergence in trait selection at a
416 broad evolutionary scale, spanning multiple botanical families, including both monocots and
417 dicots (Fig. 1).
418 The domestication traits displaying the greatest convergence are leaf size, leaf width
419 and seed area shared respectively by 8, 7 and 9 species (Fig. 1b, LeafLen, LeafWid and SArea).

420 Notably the *Poaceae* family encompassed an important number of shared domestication-
421 associated traits (Fig. 1b), including grain width (SWid), the first axis of the PCoA based on
422 NIRS (SNIRS1) and the number of tillers (TllNo).
423
424 **Index of phenotypic divergence**
425 Previous studies have shown that NIR spectra are able to discriminate between species
426 (Lang et al., 2017; Meder et al., 2014) or varieties (Barbin et al., 2018), and can be used to
427 capture genetic similarity between genotypes (Rincent et al., 2018). Reasoning that leaf
428 chemical composition was less likely targeted by selection during domestication (except
429 perhaps in leaf crops such as cabbage), we proposed to use NIRS profiling on leaves ($NIRS_{leaf}$)
430 as a control to estimate the disjunction between wild and domestic forms that is unrelated with
431 domestication. To verify this assumption, we first verified that $NIRS_{leaf}$ captured a biological
432 signal. Rand index values ranged from 0.92 in common bean to 0.96 in pearl millet confirming
433 that $NIRS_{leaf}$ effectively captured phenotypic diversity among individuals, with high
434 consistency across replicates. Second, we verified that the values of the Pillai trace obtained
435 from $NIRS_{leaf}$ ($Pillai_{control}$) were much lower than the ones computed on other traits (≥0.71 on
436 other traits versus ≤0.16 for $NIRS_{leaf}$ excluding apple, Table S5). Consistently, we found no
437 distinction between wild and domesticated forms along the axes of the PCoA (Fig. 4c) in
438 contrast with what we obtained with NIRS profiling on seeds (SNIRS1 and SNIRS2) as
439 illustrated in foxtail millet (Fig. 4a-b) and all other species except apple (Figs S1, S3–S12). In
440 apple, $Pillai_{control}$ was high (0.70, Table S5) reflecting our sampling of distant Asiatic wild
441 accessions and European domestic landraces. Finally, we verified that the mPDI values were
442 not influenced by the number of traits measured in each species (Spearman: P-value= 0.57;
443 Kendall: P-value=0.56). Based on these results, we treated $NIRS_{leaf}$ as a control trait. We
444 further used $NIRS_{leaf}$ to calculate a multivariate phenotypic divergence index (mPDI). Because
445 the mPDI of apple was very close to zero (0.03) and the disjunction of the phenotypic space
446 was almost entirely due to divergence between forms unrelated to domestication, we decided
447 to exclude it from the rest of our analysis.
448 mPDI values ranged from 0.69 for sugar beet to 0.97 for melon with an average of 0.81
449 (Fig. 2b, Table S5). Note that mPDI values obtained for maize (0.77) were likely
450 underestimated due to suboptimal greenhouse conditions (Table S3), which hindered data
451 collection as many ears failed to produce seeds, resulting in poor characterization of the
452 reproductive organs. All together our results indicate that wild and domestic forms occupy
453 distinct phenotypic spaces with some variation across species: overlap between phenotypic

454 spaces was greater in sugar beet, tomato or pearl millet than in grape, common bean or melon
455 (Fig. 2b). We found neither an effect of mating system on mPDI (Wilcoxon-Mann-Whitney:
456 P-value= 0.76), nor a correlation between mPDI and the timing of domestication (Spearman:
457 P-value= 0.64; Kendall: P-value= 0.58). We further tested the effect of mating system on the
458 mPDI per unit of disjunction unrelated with domestication and found a tendency for selfers to
459 be less disjoint than outcrossers (Fig. 2c, Wilcoxon-Mann-Whitney: P-value= 0.15).
460
461 **Comparative analysis of the multivariate phenotypic spaces**
462 Within each species, we calculated the log ratio of the determinant of the domesticated
463 variance-covariance matrix over that of the wild matrix using all traits to quantify the relative
464 expansion or contraction of multivariate phenotypic diversity in domesticated forms compared
465 to wild forms. We did similarly for the control trait ($NIRS_{leaf}$). We verified that this ratio was
466 neither influenced by the sampling size of the domestic (Spearman, P-value = 0.69; Kendall,
467 P-value = 0.58) and the wild (Spearman, P-value = 0.69; Kendall, P-value = 0.58) forms, nor
468 by the difference between the two (Spearman, P-value = 0.69; Kendall, P-value = 0.58).
469 Three out of the thirteen pairs —namely einkorn wheat, apple and maize— presented
470 positive log ratio values, indicating a larger phenotypic space in domestic forms (Fig. 5). In
471 most of the pairs (10 out of 13), the domestic forms exhibited smaller phenotypic spaces than
472 their wild counterparts, indicating a tendency for shrinkage during domestication. The log ratio
473 was neither affected by mating system (Wilcoxon-Mann-Whitney, P-value = 0.83), nor by the
474 timing of domestication (Spearman, P-value = 0.69; Kendall, P-value = 0.58). When looking
475 at the absolute values of the log ratio, the pairs covered a range from 0.07 in sugar beet, to 2.2
476 in grape that displayed the largest shrinkage of phenotypic space.
477 Finally, we computed the log ratio of phenotypic spaces using size of $NIRS_{leaf}$ and
478 found a positive correlation (Spearman, P-value = 0.04; Kendall, P-value = 0.04, Fig S15) with
479 the ratio of estimates of nucleotide diversity from the literature (Table S5) indicating that the
480 evolution of the phenotypic space of $NIRS_{leaf}$ was mainly driven by genetic drift between wild
481 and domesticated forms, further reinforcing its use as a control trait.
482
483 **Comparative analysis of trait correlations**
484 Common bean and tomato displayed significantly higher mean absolute correlations
485 between traits than their wild counterparts. Conversely, cabbage, eggplant, einkorn wheat,
486 foxtail millet, grape, melon and sugar beet displayed lower mean absolute correlation values in
487 domestic forms than in wild forms. In the other five wild/domestic pairs — African rice, apple,

maize and pearl millet — we detected no difference in the average absolute correlation values between domestic and wild form (Table S6).

Random skewers method was used to compute the similarity between domestic and wild variance-covariance matrices of traits (Fig. 6). Similarity values ranged from 0.36 in tomato to 0.99 in grape, spanning a continuous gradient. We found no impact of the mating system (Wilcoxon-Mann-Whitney: P-value=0.09), and matrices similarity did not correlate with phenotypic disjunction (Spearman: P-value= 0.72; Kendall: P-value= 0.85). However, it did correlate negatively with domestication timing (Fig. 6; Spearman: P-value= 0.004; Kendall: P-value= 0.01).

Discussion

Since Darwin (1868), domestication has served as an *in vivo* experiment for studying phenotypic evolution, offering replicated experiments to explore how divergent selection has shaped traits over time. However, most studies have focused within species, aiming to identify domestication traits and their underlying genetic changes to infer mechanisms of adaptation to human-made environments. Surprisingly, broader patterns of phenotypic shifts and the key parameters driving their evolution remain poorly understood.

A major limitation has been the absence of a unified experimental and statistical framework. Domestication traits vary across species, and, in the absence of standardized trait ontologies, are often assessed using inconsistent methodologies. Additionally, standardizing sampling across species poses significant challenges. To address these biases, we propose an experimental design in which wild populations and landraces are sampled near the center of domestication to minimize genetic structuring and avoid the confounding effects of recent breeding for the domestic form. We phenotyped these landraces/populations using standardized measurements across a diverse set of traits, representing both species-specific and shared traits. As most of the traits were quantitative, we sought to minimize environmental effects by using common gardens and standardized cultivation conditions (Table S3). Although these conditions may not always optimal (e.g., for maize), plastic responses to cultivation conditions is part of the domestication syndrome (Burban et al., 2022). In addition, we customized a statistical framework to perform cross-species comparisons, and used NIRS measured on leaves as a control trait to establish a multivariate Phenotypic Divergence Index (mPDI).

521 **Convergence of domestication-associated traits**

522 The convergence of domestication-associated traits has been well documented in
523 cereals for traits such as seed shattering. It has revealed notable cases of underlying genetic
524 convergence where orthologous genes contribute to variation of the same trait across different
525 species (Patterson et al., 2023). Recently, studies have explored convergence over deep
526 evolutionary timescales in crops from distantly related families, once again highlighting cases
527 of genetic convergence. An example is the *G* gene, which influences seed dormancy variation
528 in rice, tomato, soybean, sorghum and maize (Rendón-Anaya & Herrera-Estrella, 2018). At the
529 phenotypic level, however, meta-analyses of convergence across species struggle with the
530 diversity of experimental designs and disparity of measurement methods. By grouping
531 domestication syndrome traits of 203 crops in nine classes associated to function, Meyer,
532 DuVal, et al., (2012) have hinted convergence in some functions like the production of
533 secondary metabolites (66% of crops). Nevertheless, the statistical significance of these
534 observations was not tested and remains to be established. This necessitates a customized test,
535 as the number of measurable common traits between species decreases with increasing
536 divergence across taxa.

537 By relying on a specific sampling strategy across species, the use of trait ontology in
538 the MIAPPE format (Papoutsoglou et al., 2020) and a resampling test procedure, we
539 demonstrated that when domestication-associated traits were shared, they tended to be shared
540 by more than two species (Fig. S13). We were thus able to formally test for convergence of
541 domestication-associated traits.

542 Convergence was found for a diversity of traits. In *Poaceae*, for example, we formally
543 established the convergence observed for reduced tillering, spike and seed size increase
544 (Glémin & Bataillon, 2009). Convergence extended to more distant taxa like common bean or
545 melon as previously reported (Meyer et al., 2012). We showed increases in leaf length and
546 plant height across eight and seven species, respectively (Fig. 1b), consistent with the well-
547 documented trend of gigantism in crops (Milla & Matesanz, 2017). Noteworthy, seeds
548 increased in size in nine species (Fig. 1b, SArea), a consequence of domestication known to
549 extend to most crops. Indeed, seed enlargement during domestication is thought to have arisen
550 initially as an indirect consequence of selection for deeper burial and increased interplant
551 competition (Kluyver et al., 2013, 2017), and linked to the general increase of organ size by
552 allometry (Milla & Matesanz, 2017). Thus, even though seeds are not always the consumed
553 organ, convergence for this trait is found across taxa.

554 Interestingly, NIRS measured on seeds (SNIRS1, SNIRS2, Fig. 1b) was found to be a
555 domestication-associated trait for all systems on which it was measured, even on systems where
556 seeds were likely to have not been selected, e.g., cabbage and apple. For those crops, given that
557 NIRS is a predictor of grain quality, it implies that the changes observed following
558 domestication are not only relative to allometry but also to quality for a variety of species, as
559 previously described by chemical composition analysis in wheat or maize (Fang et al., 2020;
560 Zeibig et al., 2022). These changes may be related to the (partial) loss of seed dormancy during
561 domestication, as it is linked to changes in seed chemical composition (Penfield, 2017;
562 Prudente & Paiva, 2018).
563
564 **The multivariate phenotypic divergence index**
565 In order to better understand the evolutionary mechanisms driving phenotypic
566 divergence between wild and domesticated forms, it is essential to quantify this divergence
567 using a measure that allows cross-species comparisons. A previous attempt proposed a semi-
568 qualitative index, in which species were ranked according to their level of domestication from
569 low (1) to high (5) (Dempewolf et al., 2008). Here, we propose a multivariate divergence
570 quantitative index, the mPDI. Our results suggest that the mPDI is a reliable index. For
571 instance, the common bean, which is known for its significant phenotypic disjunction (García
572 et al., 1997; Peña-Valdivia et al., 2011) exhibits high mPDI (Fig. 2b). Conversely, sugar beet
573 a very recent to domestication displays one of the lowest mPDI value (Fig. 2b). Note that in
574 apple, mPDI was very close to zero, which can be explained by our sampling in this perennial
575 species. Indeed, our domestic landraces originated from West Europe, whereas wild
576 populations were sampled in Central Asia. The Pillai value computed on all traits was thus
577 almost equal to the value of the $Pillai_{control}$ obtained with $NIRS_{leaf}$, in line with the drift observed
578 during apple expansion to the west (Cornille et al., 2014).
579 Overall, we found a high disjunction of phenotypic spaces between wild and
580 domesticated forms across all species except apple (Pillai ≥ 0.69, Fig. 2b). Previous studies
581 focusing on two traits had proposed rather contradictory hypotheses, suggesting either a
582 disjunction (Glémin & Bataillon, 2009) or that domesticated forms represented a subset of the
583 wild phenotypic space (Milla et al., 2015). Our characterization, extended to multivariate
584 spaces, clearly supports the first hypothesis, confirming an expected outcome of divergent
585 selection.
586 Phenotypic divergence is influenced by both genetic drift and selection. We therefore
587 hypothesized that if it was primarily driven by drift, phenotypic divergence should increase

588 with the time since domestication. While one of the youngest domesticates, sugar beet, and the
589 oldest domesticate, common bean, were indeed found at the opposite extremes of mPDI values
590 (Fig. 2b), we did not find a significant trend. This suggests that selection dominates the
591 observed pattern. This conclusion is supported by the lack of a relationship between the
592 phenotypic divergence measured in our study and estimates of genomic differentiation between
593 wild and domestic forms, estimated from previous studies (Table S5). This result contrasted to
594 what we obtained from the control trait, for which Pillai values correlated with genomic
595 differentiation (Fig. S15) reinforcing its use as a trait whose evolution was dominated by drift.
596 Note that phenotypic divergence also reflects traits evolvability as measured by standing
597 genetic variance (Opedal et al., 2023). We are thus capturing multiple effects with the mPDI,
598 which are hard to decipher.

599 Because selfers can adapt faster than outcrossers (Ronfort & Glemin, 2013) potentially
600 accelerating phenotypic divergence, we tested whether the mating system had an effect on the
601 mPDI. We found no effect, and when the mPDI was expressed per unit of disjunction estimated
602 from the control trait, we observed the opposite trend, namely that selfers tended to fall in the
603 lower range of values (Fig. 2c). However, this result should be interpreted with caution due to
604 the binary classification of species as either selfers or outcrossers. In practice, many selfing
605 crops exhibit low levels of outcrossing, which can be sufficient to mix alleles and phenotypes
606 (Morjan & Rieseberg, 2004). Conversely, in outcrossing species, barriers to gene flow between
607 wild and domesticates exist – such as the *GA1* and *GA2* loci that prevent maize pollen from
608 fertilizing teosintes (Chen et al., 2022; Evans & Kermicle, 2001) among other examples.
609 Quantitative variation in the intensity of gene flow, rather than the mating system *per se*, may
610 play a more critical role in determining mPDI rankings, but gene flow remains challenging to
611 quantify.

612 As a measure of phenotypic divergence between wild and domestic forms, and thus of
613 adaptation to anthropogenetic niches, mPDI could be correlated with the fitness of these forms
614 in human-altered habitats or with their occurrence across environments differently affected by
615 human activity. Following the domestication framework proposed by Lord et al. (2025),
616 species with higher mPDI values are expected to be better adapted to human-mediated
617 environments. Unlike other indices of phenotypic divergence, such as Darwin's (Haldane,
618 1949), Haldane's (Gingerich, 1993) or Qst (Spitze, 1993), mDPI offers a multivariate
619 quantitative measure that allows cross-species comparisons and can be readily applied beyond
620 domestication contexts—for instance, to assess local adaptation in populations (Fustier et al.,
621 2019) or to study phenotypic evolution in experimental evolution assays (Desbiez-Piat et al.,

2023). Note that while easily applicable in distantly related plant taxa, the use of NIRS on specific homologous organ as a control trait may be impractical for animal species.

Evolution of the size of multivariate phenotypic spaces

Domestication has led to a loss of genomic diversity due to both demographic and selective processes (Wright et al., 2005; Yamasaki et al., 2005) that have reduced effective population sizes, thereby decreasing standing variation. One consequence of this loss is a depletion of genetic variance of traits under selection, as shown in maize (Yang et al., 2019), a pattern that extends to traits that are not directly targeted by selection (Rolhauser et al., 2024). Our results support the predicted loss of phenotypic variance, as evidenced by shrinkage of the multivariate phenotypic space in domesticated forms compared to their wild counterparts in 10 of the 13 species. However, we found no correlation between this ratio and the corresponding ratio of genomic diversity (Table S6; Spearman P-value=0.79; Kendall P-value=0.86), consistent with studies indicating that domestication affects only a small fraction of the genome—for example, 7.60% in maize (Hufford et al., 2012). Note that we found no effect of the mating system on the ratio of phenotypic diversity between domestic and wild forms, consistent with the hypothesis that selfing does not impair the response to directional selection (Noël et al., 2017).

Reorchestration of correlations between traits

The concept of domestication syndrome not only refers to a series of traits that are different between forms, but also to the correlations between them (Milla et al., 2015). The impact of phenotypic shifts across multiple traits on the evolution of phenotypic spaces is still not fully understood. Phenotypic evolution follows a genetic line of least resistance corresponding to traits with high genetic variance (Mallard, Afonso, et al., 2023; Opedal et al., 2023). The rate of this evolution is determined by the alignment between this line and the direction of selection (Mallard, Afonso, et al., 2023), and therefore on the correlation between traits. Classically, a genetic matrix of traits correlations (G-matrix) is computed. Here we relied on the phenotypic matrix of traits correlations as a proxy for the G-matrix (Cheverud, 1988).

We found that traits correlations changed during domestication, although with no clear pattern for increase or decrease across species. Previous studies predicted an increase in genetic correlations between traits in domestic compared to wild forms (Burban et al., 2022), notably because of the loss of genetic diversity in the domestic compartment and pleiotropic effect of genes, e.g. *GL4* in African rice for reduced grain size and non-shattering (Wu et al., 2017) and

656 *TB1* in maize for modification of plant and root architecture and spikelet development (Stitzer
657 & Ross-Ibarra, 2018). However, traits like seed shattering and grain size have been shown to
658 evolve concomitantly but at different rate in cereals (Fuller et al., 2012). Domestication have
659 also led to a decoupling of certain functions, such as root and leaf resource acquisition (Roucou
660 et al., 2018).

661 Interestingly, we showed that the more time has passed the more the uncoupling of
662 variance-covariance matrices, as indicated by the loss of similarity between wild and domestic
663 matrices (Fig. 6). This suggests that a general hallmark of domestication may involve a
664 reorganization of trait correlations, a pattern previously identified in different species (various
665 herbaceous crops, Milla et al., 2014; wheat, Roucou et al., 2018; maize, Yang et al., 2019). As
666 time passes, the interplay of mutation and selection reshapes the structure of the G-matrix
667 (Dugand et al., 2021; Houle et al., 2017; Mallard, Noble, et al., 2023), potentially leading to an
668 alignment to the adaptative landscape (Hohenlohe & Arnold, 2008; Hunt et al., 2007).
669 Consequently, as time since domestication increases, the G-matrices of wild and domestic
670 forms may diverge as they become aligned with their respective axes of selection. Altogether,
671 these observations echoes findings that recombination tends to increase during domestication
672 (Ross-Ibarra, 2004) or, at minimum, that the recombination landscape is reshaped along the
673 genome (Bursell et al., 2025). Such novel recombination dynamics can influence phenotypic
674 variation (Pan et al., 2016; Wagner, 2011) by breaking correlation between traits, potentially
675 accelerating G-matrix alignment with the direction of selection in the anthropogenic niche.
676

677 **Conclusion**

678 In this study, we present the first comparative analysis of plant domestication across
679 multiple species. We introduce a novel phenotypic divergence index (mPDI), applicable across
680 all domesticated species, providing a standardized framework for cross-species comparisons.
681 By characterizing the domestication syndrome in 13 species, we demonstrated that near-
682 infrared spectroscopy (NIRS) measurements on grains contribute to the domestication
683 syndrome in many species, highlighting the potential of the NIRS technology for both
684 calculating the mPDI and characterizing domestication-related changes.

685 Our findings support several key hypotheses of domestication: trait convergence,
686 disjunction between wild and domestic phenotypes, a relative reduction of the phenotypic space
687 in domesticated forms, and a reorganization of trait correlations. Importantly, we reveal the

emergence of shared patterns across species, independent of domestication timing or life-history traits, indicating the existence of general principles underlying plant domestication.

Acknowledgments

We are very grateful to Roberto Papa and Elena Bitocchi (Università Politecnica delle Marche, Ancona, Italy) for providing common bean seeds; Alain Label-Richardson (Université Rennes, France), Lorenzo Maggioni (European Cooperative Program for Plant Genetic Resources, Roma, Italy), Roland Von Bothmer (Swedish University of Agricultural Sciences, Alnarp, Sweden), and Charlotte Allender & Nick Fenby (University of Warwick, UK) for providing cabbage seeds; Monika Höfer (Julius Kühn Institute - Federal Research Institute for Cultivated Plants) for managing apple genetic ressource and to Amandine Cornille (NYU Abu Dhabi) for discussing the sampling design of apple trees. We thank the technical staff Virginie Heraudet, Nathalie Galic, Cécilia Zokly, Arnaud Leclerc, Victoria Rozon and Florie Legentil of IDEEV greenhouse for the maintenance of plants. We also thank Yaël Bourgain, Dominique Emmanuel, Laura Courreyan, Martin Simonoviez, Lucie Guerrin, Violette van der Horst and Theophile Thomas for their contribution to data collection. We thank Leila Zekraoui at IRD for her contribution to the discussion on the phenotyping strategy. We thank Martin Ecarnot at INRAE for the access and support at the NIRS plateform ARCAD in Montpellier. We thank Betty Gimenez, Odile Peyron, Sébastien Joannin and Laurent Bouby for the counting of grapevine pollen grains. We thank Alain Ghesquière for the photography of African rice (© IRD). We also deeply thank Mark Chapman and one anonymous reviewer for their insightful comments on the manuscript.

Competing interests

The authors declare no conflicts of interest

Data availability

Phenotypic data and NIR spectra are available on https://doi.org/10.57745/QWEKVK. R scripts are available on the INRAE Forge at https://forge.inrae.fr/gqe-gevad/domisol_phenotypic_spaces. Ontology of measured traits at the MIAPPE format is available at https://doi.org/10.57745/QWEKVK.

References

Abbo, S., & Gopher, A. (2020). Plant domestication in the Neolithic Near East: The humans-plants liaison. *Quaternary Science Reviews*, *242*, 106412. https://doi.org/10.1016/j.quascirev.2020.106412

Ahmed, H. I., Heuberger, M., Schoen, A., Koo, D.-H., Quiroz-Chavez, J., Adhikari, L., Raupp, J., Cauet, S., Rodde, N., Cravero, C., Callot, C., Lazo, G. R., Kathiresan, N., Sharma, P. K., Moot, I., Yadav, I. S., Singh, L., Saripalli, G., Rawat, N., … Krattinger, S. G. (2023). Einkorn genomics sheds light on history of the oldest domesticated wheat. *Nature*, *620*(7975), 830–838. https://doi.org/10.1038/s41586-023-06389-7

Arnoux, S., Fraïsse, C., & Sauvage, C. (2021). Genomic inference of complex domestication histories in three Solanaceae species. *Journal of Evolutionary Biology*, *34*(2), 270–283. https://doi.org/10.1111/jeb.13723

Barbin, D. F., Maciel, L. F., Bazoni, C. H. V., Ribeiro, M. da S., Carvalho, R. D. S., Bispo, E. da S., Miranda, M. da P. S., & Hirooka, E. Y. (2018). Classification and compositional characterization of different varieties of cocoa beans by near infrared spectroscopy and multivariate statistical analyses. *Journal of Food Science and Technology*, *55*(7), 2457–2466. https://doi.org/10.1007/s13197-018-3163-5

Barker, G. (2006). *The Agricultural Revolution in Prehistory: Why did Foragers become Farmers?* Oxford University Press. https://doi.org/10.1093/oso/9780199281091.001.0001

Bartsch, D., Lehnen, M., Clegg, J., Pohl-Orf, M., Schuphan, I., & Ellstrand, N. C. (1999). Impact of gene flow from cultivated beet on genetic diversity of wild sea beet populations. *Molecular Ecology*, *8*(10), 1733–1741. https://doi.org/10.1046/j.1365-294x.1999.00769.x

Bellucci, E., Bitocchi, E., Ferrarini, A., Benazzo, A., Biagetti, E., Klie, S., Minio, A., Rau, D., Rodriguez, M., Panziera, A., Venturini, L., Attene, G., Albertini, E., Jackson, S. A., Nanni, L., Fernie, A. R., Nikoloski, Z., Bertorelle, G., Delledonne, M., & Papa, R. (2014). Decreased Nucleotide and Expression Diversity and Modified Coexpression Patterns Characterize Domestication in the Common Bean. *The Plant Cell*, *26*(5), 1901–1912. https://doi.org/10.1105/tpc.114.124040

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bhishagratna, K. (1911). *An English translation of The Sushruta Samhita: Based on original Sanskrit text* (Vol. 2).

Borraz-Martínez, S., Boqué, R., Simó, J., Mestre, M., & Gras, A. (2019). Development of a methodology to analyze leaves from *Prunus dulcis* varieties using near infrared spectroscopy. *Talanta*, *204*, 320–328. https://doi.org/10.1016/j.talanta.2019.05.105

Bradshaw, J. E. (2016). Domestication, Dispersion, Selection and Hybridization of Cultivated Plants. In J. E. Bradshaw (Ed.), *Plant Breeding: Past, Present and Future* (pp. 3–38). Springer International Publishing. https://doi.org/10.1007/978-3-319-23285-0_1

Bucksch, A., Burridge, J., York, L. M., Das, A., Nord, E., Weitz, J. S., & Lynch, J. P. (2014). Image-Based High-Throughput Field Phenotyping of Crop Roots. *Plant Physiology*, *166*(2), 470–486. https://doi.org/10.1104/pp.114.243519

Burban, E., Tenaillon, M. I., & Le Rouzic, A. (2022). Gene network simulations provide testable predictions for the molecular domestication syndrome. *Genetics*, *220*(2), iyab214. https://doi.org/10.1093/genetics/iyab214

Burgarella, C., Cubry, P., Kane, N. A., Varshney, R. K., Mariac, C., Liu, X., Shi, C., Thudi, M., Couderc, M., Xu, X., Chitikineni, A., Scarcelli, N., Barnaud, A., Rhoné, B., Dupuy, C., François, O., Berthouly-Salazar, C., & Vigouroux, Y. (2018). A western Sahara centre of domestication inferred from pearl millet genomes. *Nature Ecology & Evolution*, *2*(9), 1377–1380. https://doi.org/10.1038/s41559-018-0643-y

Bursell, M., Rohilla, M., Ramirez, L., Cheng, Y., Schwarzkopf, E. J., Guerrero, R. F., & Smukowski Heil, C. (2025). Mixed Outcomes in Recombination Rates After Domestication: Revisiting Theory and Data. *Molecular Ecology*, *n/a*(n/a), e17773. https://doi.org/10.1111/mec.17773

Cai, C., Bucher, J., Bakker, F. T., & Bonnema, G. (2022). Evidence for two domestication lineages supporting a middle-eastern origin for Brassica oleracea crops from diversified kale populations. *Horticulture Research*, *9*, uhac033. https://doi.org/10.1093/hr/uhac033

Chen, Z., Zhang, Z., Zhang, H., Li, K., Cai, D., Zhao, L., Liu, J., & Chen, H. (2022). A pair of non-Mendelian genes at the Ga2 locus confer unilateral cross-incompatibility in maize. *Nature Communications*, *13*(1), 1993. https://doi.org/10.1038/s41467-022-29729-z

Cheverud, J. M. (1988). A Comparison of Genetic and Phenotypic Correlations. *Evolution*, *42*(5), 958–968. https://doi.org/10.1111/j.1558-5646.1988.tb02514.x

Cheverud, J. M., & Marroig, G. (2007). Research Article Comparing covariance matrices: Random skewers method compared to the common principal components model. *Genetics and Molecular Biology*, *30*, 461–469. https://doi.org/10.1590/S1415-47572007000300027

Choi, J. Y., Zaidem, M., Gutaker, R., Dorph, K., Singh, R. K., & Purugganan, M. D. (2019). The complex geography of domestication of the African rice Oryza glaberrima. *PLOS Genetics*, *15*(3), e1007414. https://doi.org/10.1371/journal.pgen.1007414

Cornille, A., Giraud, T., Smulders, M. J. M., Roldán-Ruiz, I., & Gladieux, P. (2014). The domestication and evolutionary ecology of apples. *Trends in Genetics*, *30*(2), 57–65. https://doi.org/10.1016/j.tig.2013.10.002

Darwin, C. (1868). *The Variation of Animals and Plants Under Domestication* (H. Ritvo, Ed.; Vol. 1). Johns Hopkins University Press. https://doi.org/10.1017/CBO9780511709500

de Bock, T. S. M. (1986). The genus Beta: Domestication, taxonomy and interspecific hybridization for plant breeding. *Acta Horticulturae*, *38*(182), 335–343.

Dempewolf, H., Hodgins, K. A., Rummell, S. E., Ellstrand, N. C., & Rieseberg, L. H. (2012). Reproductive Isolation during Domestication. *The Plant Cell*, *24*(7), 2710–2717. https://doi.org/10.1105/tpc.112.100115

Dempewolf, H., Rieseberg, L. H., & Cronk, Q. C. (2008). Crop domestication in the Compositae: A family-wide trait assessment. *Genetic Resources and Crop Evolution*, *55*(8), 1141–1157. https://doi.org/10.1007/s10722-008-9315-0

Denham, T., Barton, H., Castillo, C., Crowther, A., Dotte-Sarout, E., Florin, S. A., Pritchard, J., Barron, A., Zhang, Y., & Fuller, D. Q. (2020). The domestication syndrome in vegetatively propagated field crops. *Annals of Botany*, *125*(4), 581–597. https://doi.org/10.1093/aob/mcz212

Desbiez-Piat, A., Ressayre, A., Marchadier, E., Noly, A., Remoué, C., Vitte, C., Belcram, H., Bourgais, A., Galic, N., Le Guilloux, M., Tenaillon, M. I., & Dillmann, C. (2023). Pervasive G × E interactions shape adaptive trajectories and the exploration of the phenotypic space in artificial selection experiments. *Genetics*, *225*(4), iyad186. https://doi.org/10.1093/genetics/iyad186

Dirzo, R., & Raven, P. H. (2003). Global State of Biodiversity and Loss. *Annual Review of Environment and Resources*, *28*(Volume 28, 2003), 137–167. https://doi.org/10.1146/annurev.energy.28.050302.105532

Duan, N., Bai, Y., Sun, H., Wang, N., Ma, Y., Li, M., Wang, X., Jiao, C., Legall, N., Mao, L., Wan, S., Wang, K., He, T., Feng, S., Zhang, Z., Mao, Z., Shen, X., Chen, X., Jiang, Y., … Chen, X. (2017). Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. *Nature Communications*, *8*(1), 249. https://doi.org/10.1038/s41467-017-00336-7

Dugand, R. J., Aguirre, J. D., Hine, E., Blows, M. W., & McGuigan, K. (2021). The contribution of mutation and selection to multivariate quantitative genetic variance in an outbred population of Drosophila serrata. *Proceedings of the National Academy of Sciences*, *118*(31), e2026217118. https://doi.org/10.1073/pnas.2026217118

Dussert, Y., Snirc, A., & Robert, T. (2015). Inference of domestication history and differentiation between early- and late-flowering varieties in pearl millet. *Molecular Ecology*, *24*(7), 1387–1402. https://doi.org/10.1111/mec.13119

Evans, M. M. S., & Kermicle, J. L. (2001). Teosinte crossing barrier1, a locus governing hybridization of teosinte with maize. *Theoretical and Applied Genetics*, *103*(2), 259–265. https://doi.org/10.1007/s001220100549

Fang, H., Fu, X., Wang, Y., Xu, J., Feng, H., Li, W., Xu, J., Jittham, O., Zhang, X., Zhang, L., Yang, N., Xu, G., Wang, M., Li, X., Li, J., Yan, J., & Yang, X. (2020). Genetic basis of kernel nutritional traits during maize domestication and improvement. *The Plant Journal*, *101*(2), 278–292. https://doi.org/10.1111/tpj.14539

Fietta, A., Caramelli, D., Schlumbaum, A., & Vernesi, C. (2014). EvolvApple. Studying apple domestication through the genome scan of ancient seeds. *6th International Symposium on Biomolecular Archaeology*.

Fuller, D. Q., Asouti, E., & Purugganan, M. D. (2012). Cultivation as slow evolutionary entanglement: Comparative data on rate and sequence of domestication. *Vegetation History and Archaeobotany*, *21*(2), 131–145. https://doi.org/10.1007/s00334-011-0329-8

Fuller, D. Q., Denham, T., & Allaby, R. (2023). Plant domestication and agricultural ecologies. *Current Biology*, *33*(11), R636–R649. https://doi.org/10.1016/j.cub.2023.04.038

Fustier, M.-A., Martínez-Ainsworth, N. E., Aguirre-Liguori, J. A., Venon, A., Corti, H., Rousselet, A., Dumas, F., Dittberner, H., Camarena, M. G., Grimanelli, D., Ovaskainen, O., Falque, M., Moreau, L., Meaux, J. de, Montes-Hernández, S., Eguiarte, L. E., Vigouroux, Y., Manicacci, D., & Tenaillon, M. I. (2019). Common gardens in teosintes reveal the establishment of a syndrome of adaptation to altitude. *PLOS Genetics*, *15*(12), e1008512. https://doi.org/10.1371/journal.pgen.1008512

García, E. H., Peña-valdivia, C. B., Rogelio aguirre, J. R., & Muruaga, J. S. M. (1997). Morphological and Agronomic Traits of a Wild Population and an Improved Cultivar of Common Bean (Phaseolus vulgarisL.). *Annals of Botany*, *79*(2), 207–213. https://doi.org/10.1006/anbo.1996.0329

Gaut, B. S., Díez, C. M., & Morrell, P. L. (2015). Genomics and the Contrasting Dynamics of Annual and Perennial Domestication. *Trends in Genetics*, *31*(12), 709–719. https://doi.org/10.1016/j.tig.2015.10.002

Gingerich, P. D. (1993). Quantification and comparison of evolutionary rates. *American Journal of Science*, *293*(A), 453–478.

Glémin, S., & Bataillon, T. (2009). A comparative view of the evolution of grasses under domestication. *New Phytologist*, *183*(2), 273–290. https://doi.org/10.1111/j.1469-8137.2009.02884.x

Gross, B. L., Henk, A. D., Richards, C. M., Fazio, G., & Volk, G. M. (2014). Genetic diversity in Malus ×domestica (Rosaceae) through time in response to domestication. *American Journal of Botany*, *101*(10), 1770–1779. https://doi.org/10.3732/ajb.1400297

Haldane, J. B. S. (1949). Suggestions as to Quantitative Measurement of Rates of Evolution. *Evolution*, *3*(1), 51–56. https://doi.org/10.2307/2405451

Hohenlohe, P. A., & Arnold, S. J. (2008). MIPoD: A Hypothesis-Testing Framework for Microevolutionary Inference from Patterns of Divergence. *The American Naturalist*, *171*(3), 366–385. https://doi.org/10.1086/527498

Houle, D., Bolstad, G. H., van der Linde, K., & Hansen, T. F. (2017). Mutation predicts 40 million years of fly wing evolution. *Nature*, *548*(7668), 447–450. https://doi.org/10.1038/nature23473

Hufford, M. B., Xu, X., van Heerwaarden, J., Pyhäjärvi, T., Chia, J.-M., Cartwright, R. A., Elshire, R. J., Glaubitz, J. C., Guill, K. E., Kaeppler, S. M., Lai, J., Morrell, P. L., Shannon, L. M., Song, C., Springer, N. M., Swanson-Wagner, R. A., Tiffin, P., Wang, J., Zhang, G., … Ross-Ibarra, J. (2012). Comparative population genomics of maize domestication and improvement. *Nature Genetics*, *44*(7), 808–811. https://doi.org/10.1038/ng.2309

Hunt, J., Blows, M. W., Zajitschek, F., Jennions, M. D., & Brooks, R. (2007). Reconciling Strong Stabilizing Selection with the Maintenance of Genetic Variation in a Natural Population of Black Field Crickets (Teleogryllus commodus). *Genetics*, *177*(2), 875–880. https://doi.org/10.1534/genetics.107.077057

Kaplan, L. (1981). What is the origin of the common bean? *Economic Botany*, *35*(2), 240–254. https://doi.org/10.1007/BF02858692

Kluyver, T. A., Charles, M., Jones, G., Rees, M., & Osborne, C. P. (2013). Did greater burial depth increase the seed size of domesticated legumes? *Journal of Experimental Botany*, *64*(13), 4101–4108. https://doi.org/10.1093/jxb/ert304

Kluyver, T. A., Jones, G., Pujol, B., Bennett, C., Mockford, E. J., Charles, M., Rees, M., & Osborne, C. P. (2017). Unconscious selection drove seed enlargement in vegetable crops. *Evolution Letters*, *1*(2), 64–72. https://doi.org/10.1002/evl3.6

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lande, R. (1979). Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry. *Evolution*, *33*(1), 402–416. https://doi.org/10.2307/2407630

Lang, C., Almeida, D. R. A., & Costa, F. R. C. (2017). Discrimination of taxonomic identity at species, genus and family levels using Fourier Transformed Near-Infrared Spectroscopy (FT-NIR). *Forest Ecology and Management*, *406*, 219–227. https://doi.org/10.1016/j.foreco.2017.09.003

Leisch, F. (2006). A toolbox for K-centroids cluster analysis. *Computational Statistics & Data Analysis*, *51*(2), 526–544. https://doi.org/10.1016/j.csda.2005.10.006

Lord, K. A., Larson, G., Allaby, R. G., & Karlsson, E. K. (2025). A universally applicable definition for domestication. *Proceedings of the National Academy of Sciences*, *122*(22), e2413207122. https://doi.org/10.1073/pnas.2413207122

Mabry, M. E., Turner-Hissong, S. D., Gallagher, E. Y., McAlvay, A. C., An, H., Edger, P. P., Moore, J. D., Pink, D. A. C., Teakle, G. R., Stevens, C. J., Barker, G., Labate, J., Fuller, D. Q., Allaby, R. G., Beissinger, T., Decker, J. E., Gore, M. A., & Pires, J. C. (2021). The Evolutionary History of Wild, Domesticated, and Feral Brassica oleracea

(Brassicaceae). *Molecular Biology and Evolution*, *38*(10), 4419–4434. https://doi.org/10.1093/molbev/msab183

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., Ormanbekova, D., Lux, T., Prade, V. M., Milner, S. G., Himmelbach, A., Mascher, M., Bagnaresi, P., Faccioli, P., Cozzi, P., Lauria, M., Lazzari, B., Stella, A., Manconi, A., … Cattivelli, L. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nature Genetics*, *51*(5), 885–895. https://doi.org/10.1038/s41588-019-0381-3

Maggioni, L., von Bothmer, R., Poulsen, G., & Härnström Aloisi, K. (2020). Survey and genetic diversity of wild Brassica oleracea L. germplasm on the Atlantic coast of France. *Genetic Resources and Crop Evolution*, *67*(7), 1853–1866. https://doi.org/10.1007/s10722-020-00945-0

Mallard, F., Afonso, B., & Teotónio, H. (2023). Selection and the direction of phenotypic evolution. *eLife*, *12*, e80993. https://doi.org/10.7554/eLife.80993

Mallard, F., Noble, L., Baer, C. F., & Teotónio, H. (2023). Variation in mutational (co)variances. *G3 Genes|Genomes|Genetics*, *13*(2), jkac335. https://doi.org/10.1093/g3journal/jkac335

Manning, K., Pelling, R., Higham, T., Schwenniger, J.-L., & Fuller, D. Q. (2011). 4500-Year old domesticated pearl millet (*Pennisetum glaucum*) from the Tilemsi Valley, Mali: New insights into an alternative cereal domestication pathway. *Journal of Archaeological Science*, *38*(2), 312–322. https://doi.org/10.1016/j.jas.2010.09.007

Marrano, A., Micheletti, D., Lorenzi, S., Neale, D., & Grando, M. S. (2018). Genomic signatures of different adaptations to environmental stimuli between wild and cultivated Vitis vinifera L. *Horticulture Research*, *5*, 34. https://doi.org/10.1038/s41438-018-0041-2

McGovern, P., Jalabadze, M., Batiuk, S., Callahan, M. P., Smith, K. E., Hall, G. R., Kvavadze, E., Maghradze, D., Rusishvili, N., Bouby, L., Failla, O., Cola, G., Mariani, L., Boaretto, E., Bacilieri, R., This, P., Wales, N., & Lordkipanidze, D. (2017). Early Neolithic wine of Georgia in the South Caucasus. *Proceedings of the National Academy of Sciences*, *114*(48), E10309–E10318. https://doi.org/10.1073/pnas.1714728114

Meder, R., Kain, D., Ebdon, N., Macdonell, P., & Brawner, J. T. (2014). Identifying Hybridisation in Pinus Species Using near Infrared Spectroscopy of Foliage. *Journal of Near Infrared Spectroscopy*, *22*(5), 337–345.

Melo, D., Garcia, G., Hubbe, A., Assis, A. P., & Marroig, G. (2016). EvolQG - An R package for evolutionary quantitative genetics. *F1000Research*, *4*, 925. https://doi.org/10.12688/f1000research.7082.3

Meyer, R. S., DuVal, A. E., & Jensen, H. R. (2012). Patterns and processes in crop domestication: An historical review and quantitative analysis of 203 global food crops. *New Phytologist*, *196*(1), 29–48. https://doi.org/10.1111/j.1469-8137.2012.04253.x

Meyer, R. S., & Purugganan, M. D. (2013). Evolution of crop species: Genetics of domestication and diversification. *Nature Reviews Genetics*, *14*(12), 840–852. https://doi.org/10.1038/nrg3605

Milla, R. (2020). Crop Origins and Phylo Food: A database and a phylogenetic tree to stimulate comparative analyses on the origins of food crops. *Global Ecology and Biogeography*, *29*(4), 606–614. https://doi.org/10.1111/geb.13057

Milla, R., & Matesanz, S. (2017). Growing larger with domestication: A matter of physiology, morphology or allocation? *Plant Biology*, *19*(3), 475–483. https://doi.org/10.1111/plb.12545

Milla, R., Morente-López, J., Alonso-Rodrigo, J. M., Martín-Robles, N., & Stuart Chapin, F. (2014). Shifts and disruptions in resource-use trait syndromes during the evolution of herbaceous crops. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1793), 20141429. https://doi.org/10.1098/rspb.2014.1429

Milla, R., Osborne, C. P., Turcotte, M. M., & Violle, C. (2015). Plant domestication through an ecological lens. *Trends in Ecology & Evolution*, *30*(8), 463–469. https://doi.org/10.1016/j.tree.2015.06.006

Miller, A. J., & Gross, B. L. (2011). From forest to field: Perennial fruit crop domestication. *American Journal of Botany*, *98*(9), 1389–1414. https://doi.org/10.3732/ajb.1000522

Morjan, C. L., & Rieseberg, L. H. (2004). How species evolve collectively: Implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology*, *13*(6), 1341–1356. https://doi.org/10.1111/j.1365-294X.2004.02164.x

Murray, S. S. (2004). Searching for the origins of African rice domestication. *Antiquity Project Gallery78*, *300*.

Noël, E., Jarne, P., Glémin, S., MacKenzie, A., Segard, A., Sarda, V., & David, P. (2017). Experimental Evidence for the Negative Effects of Self-Fertilization on the Adaptive Potential of Populations. *Current Biology*, *27*(2), 237–242. https://doi.org/10.1016/j.cub.2016.11.015

Opedal, Ø. H., Armbruster, W. S., Hansen, T. F., Holstad, A., Pélabon, C., Andersson, S., Campbell, D. R., Caruso, C. M., Delph, L. F., Eckert, C. G., Lankinen, Å., Walter, G. M., Ågren, J., & Bolstad, G. H. (2023). Evolvability and trait function predict phenotypic divergence of plant populations. *Proceedings of the National Academy of Sciences*, *120*(1), e2203228120. https://doi.org/10.1073/pnas.2203228120

Pan, Q., Li, L., Yang, X., Tong, H., Xu, S., Li, Z., Li, W., Muehlbauer, G. J., Li, J., & Yan, J. (2016). Genome-wide recombination dynamics are associated with phenotypic

variation in maize. *New Phytologist*, *210*(3), 1083–1094. https://doi.org/10.1111/nph.13810

Papoutsoglou, E. A., Faria, D., Arend, D., Arnaud, E., Athanasiadis, I. N., Chaves, I., Coppens, F., Cornut, G., Costa, B. V., Ćwiek-Kupczyńska, H., Droesbeke, B., Finkers, R., Gruden, K., Junker, A., King, G. J., Krajewski, P., Lange, M., Laporte, M.-A., Michotey, C., … Pommier, C. (2020). Enabling reusability of plant phenomic datasets with MIAPPE 1.1. *New Phytologist*, *227*(1), 260–273. https://doi.org/10.1111/nph.16544

Patterson, E. L., Richardson, A., & Bartlett, M. (2023). Pushing the boundaries of organ identity: Homology of the grass lemma. *American Journal of Botany*, *110*(4), e16161. https://doi.org/10.1002/ajb2.16161

Peña-Valdivia, C. B., García-Nava, J. R., Aguirre R, J. R., Ybarra-Moncada, Ma. C., & López H, M. (2011). Variation in Physical and Chemical Characteristics of Common Bean (Phaseolus vulgaris L.) Grain along a Domestication Gradient. *Chemistry & Biodiversity*, *8*(12), 2211–2225. https://doi.org/10.1002/cbdv.201100102

Penfield, S. (2017). Seed dormancy and germination. *Current Biology*, *27*(17), R874–R878. https://doi.org/10.1016/j.cub.2017.05.050

Pillai, K. C. S. (1955). Some New Test Criteria in Multivariate Analysis. *The Annals of Mathematical Statistics*, *26*(1), 117–121.

Piperno, D. R., & Flannery, K. V. (2001). The earliest archaeological maize (Zea mays L.) from highland Mexico: New accelerator mass spectrometry dates and their implications. *Proceedings of the National Academy of Sciences*, *98*(4), 2101–2103. https://doi.org/10.1073/pnas.98.4.2101

Prudente, D. de O., & Paiva, R. (2018). Seed Dormancy and Germination: Physiological Considerations. *Journal of Cell and Developmental Biology*, *2*(1).

https://www.imedpub.com/abstract/seed-dormancy-and-germination-physiological-considerations-22078.html

Purugganan, M. D., & Fuller, D. Q. (2011). Archeological data reveal slow rates of evolution during plant domestication. *Evolution*, *65*(1), 171-183. https://doi.org/10.1111/j.1558-5646.2010.01093.x

Rand, W. M. (1971). Objective Criteria for the Evaluation of Clustering Methods. *Journal of the American Statistical Association*, *66*(336), 846–850. https://doi.org/10.2307/2284239

Rapp, R. A., Haigler, C. H., Flagel, L., Hovav, R. H., Udall, J. A., & Wendel, J. F. (2010). Gene expression in developing fibres of Upland cotton (Gossypium hirsutum L.) was massively altered by domestication. *BMC Biology*, *8*(1), 139. https://doi.org/10.1186/1741-7007-8-139

Razifard, H., Ramos, A., Della Valle, A. L., Bodary, C., Goetz, E., Manser, E. J., Li, X., Zhang, L., Visa, S., Tieman, D., van der Knaap, E., & Caicedo, A. L. (2020). Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. *Molecular Biology and Evolution*, *37*(4), 1118–1132. https://doi.org/10.1093/molbev/msz297

Rendón-Anaya, M., & Herrera-Estrella, A. (2018). The advantage of parallel selection of domestication genes to accelerate crop improvement. *Genome Biology*, *19*(1), 147. https://doi.org/10.1186/s13059-018-1537-7

Riaz, S., De Lorenzis, G., Velasco, D., Koehmstedt, A., Maghradze, D., Bobokashvili, Z., Musayev, M., Zdunic, G., Laucou, V., Andrew Walker, M., Failla, O., Preece, J. E., Aradhya, M., & Arroyo-Garcia, R. (2018). Genetic diversity analysis of cultivated and wild grapevine (Vitis vinifera L.) accessions around the Mediterranean basin and

Central Asia. *BMC Plant Biology*, *18*(1), 137. https://doi.org/10.1186/s12870-018-1351-0

Rincent, R., Charpentier, J.-P., Faivre-Rampant, P., Paux, E., Le Gouis, J., Bastien, C., & Segura, V. (2018). Phenomic Selection Is a Low-Cost and High-Throughput Method Based on Indirect Predictions: Proof of Concept on Wheat and Poplar. *G3 Genes|Genomes|Genetics*, *8*(12), 3961–3972. https://doi.org/10.1534/g3.118.200760

Rolhauser, A. G., Isaac, M. E., Violle, C., Martin, A. R., Vasseur, F., Lemoine, T., Mahaut, L., Fort, F., Rotundo, J. L., & Vile, D. (2024). Phenotypic limits of crop diversity: A data exploration of functional trait space. *New Phytologist*, *244*(2), 708–718. https://doi.org/10.1111/nph.20050

Ronfort, J., & Glemin, S. (2013). Mating system, Haldane's sieve, and the domestication process. *Evolution*, *67*(5), 1518–1526. https://doi.org/10.1111/evo.12025

Ross-Ibarra, J. (2004). The Evolution of Recombination under Domestication: A Test of Two Hypotheses. *The American Naturalist*, *163*(1), 105–112. https://doi.org/10.1086/380606

Ross-Ibarra, J., Morrell, P. L., & Gaut, B. S. (2007). Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proceedings of the National Academy of Sciences*, *104*(suppl_1), 8641–8648. https://doi.org/10.1073/pnas.0700643104

Roucou, A., Violle, C., Fort, F., Roumet, P., Ecarnot, M., & Vile, D. (2018). Shifts in plant functional strategies over the course of wheat domestication. *Journal of Applied Ecology*, *55*(1), 25–37. https://doi.org/10.1111/1365-2664.13029

Sauvage, C., Rau, A., Aichholz, C., Chadoeuf, J., Sarah, G., Ruiz, M., Santoni, S., Causse, M., David, J., & Glémin, S. (2017). Domestication rewired gene expression and

nucleotide diversity patterns in tomato. *The Plant Journal*, *91*(4), 631–645. https://doi.org/10.1111/tpj.13592

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, *36*, 1627–1639. https://doi.org/doi.org/10.1021/ac60319a045

Schmutz, J., McClean, P. E., Mamidi, S., Wu, G. A., Cannon, S. B., Grimwood, J., Jenkins, J., Shu, S., Song, Q., Chavarro, C., Torres-Torres, M., Geffroy, V., Moghaddam, S. M., Gao, D., Abernathy, B., Barry, K., Blair, M., Brick, M. A., Chovatia, M., … Jackson, S. A. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. *Nature Genetics*, *46*(7), 707–713. https://doi.org/10.1038/ng.3008

Smith, B. D. (1998). *Emergence of Agriculture*.

Spitze, K. (1993). Population Structure in Daphnia Obtusa: Quantitative Genetic and Allozymic Variation. *Genetics*, *135*(2), 367–374. https://doi.org/10.1093/genetics/135.2.367

Stitzer, M. C., & Ross-Ibarra, J. (2018). Maize domestication and gene interaction. *New Phytologist*, *220*(2), 395–408. https://doi.org/10.1111/nph.15350

Studer, A. J., Wang, H., & Doebley, J. F. (2017). Selection During Maize Domestication Targeted a Gene Network Controlling Plant and Inflorescence Architecture. *Genetics*, *207*(2), 755–765. https://doi.org/10.1534/genetics.117.300071

Tehseen, M. M., Wyatt, N. A., Bolton, M. D., Fugate, K. K., Preister, L. S., Yang, S., Ramachandran, V., Li, X., & Chu, C. (2024). Genetic drift, historic migration, and limited gene flow contributing to the subpopulation divergence in wild sea beet (Beta vulgaris ssp. Maritima (L.) Arcang). *PLOS ONE*, *19*(9), e0308626. https://doi.org/10.1371/journal.pone.0308626

Tenaillon, M. I., & Charcosset, A. (2011). A European perspective on maize history. *Comptes Rendus. Biologies*, *334*(3), 221–228. https://doi.org/10.1016/j.crvi.2010.12.015

Tenaillon, M. I., & Manicacci, D. (2011). Maize origins: An old question under the spotlights. In *Advances in Maize* (Vol. 3, pp. 89–110). The Society for Experimental Biology.

van Zeist, W., & de Roller, G. J. (1993). Plant remains from Maadi, a predynastic site in lower Egypt. *Vegetation History and Archaeobotany*, *2*(1), 1–14. https://doi.org/10.1007/BF00191701

Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., Zhang, H., Zhao, Y., Wang, X., Rathore, A., Srivastava, R. K., Chitikineni, A., Fan, G., Bajaj, P., Punnuri, S., Gupta, S. K., Wang, H., Jiang, Y., Couderc, M., … Xu, X. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nature Biotechnology*, *35*(10), 969–976. https://doi.org/10.1038/nbt.3943

Vavilov, N. (1926). *Studies on the origin of cultivated plants* (Vol. 16).

Vavilov, N. I. (1951). The Origin, Variation, Immunity and Breeding of Cultivated Plants. *Soil Science*, *72*(6), 482.

Veltman, M. A., Flowers, J. M., Andel, T. R. van, & Schranz, M. E. (2019). Origins and geographic diversification of African rice (Oryza glaberrima). *PLOS ONE*, *14*(3), e0203508. https://doi.org/10.1371/journal.pone.0203508

Wagner, A. (2011). The low cost of recombination in creating novel phenotypes. *BioEssays*, *33*(8), 636–646. https://doi.org/10.1002/bies.201100027

Wang, C., Chen, J., Zhi, H., Yang, L., Li, W., Wang, Y., Li, H., Zhao, B., Chen, M., & Diao, X. (2010). Population genetics of foxtail millet and its wild ancestor. *BMC Genetics*, *11*(1), 90. https://doi.org/10.1186/1471-2156-11-90

Weisdorf, J. L. (2005). From Foraging To Farming: Explaining The Neolithic Revolution. *Journal of Economic Surveys*, *19*(4), 561–586. https://doi.org/10.1111/j.0950-0804.2005.00259.x

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer. https://doi.org/10.1007/978-0-387-98141-3

Wright, S. I., Bi, I. V., Schroeder, S. G., Yamasaki, M., Doebley, J. F., McMullen, M. D., & Gaut, B. S. (2005). The Effects of Artificial Selection on the Maize Genome. *Science*, *308*(5726), 1310–1314. https://doi.org/10.1126/science.1107891

Wu, W., Liu, X., Wang, M., Meyer, R. S., Luo, X., Ndjiondjop, M.-N., Tan, L., Zhang, J., Wu, J., Cai, H., Sun, C., Wang, X., Wing, R. A., & Zhu, Z. (2017). A single-nucleotide polymorphism causes smaller grain size and loss of seed shattering during African rice domestication. *Nature Plants*, *3*(6), 1–7. https://doi.org/10.1038/nplants.2017.64

Yamasaki, M., Tenaillon, M. I., Vroh Bi, I., Schroeder, S. G., Sanchez-Villeda, H., Doebley, J. F., Gaut, B. S., & McMullen, M. D. (2005). A Large-Scale Screen for Artificial Selection in Maize Identifies Candidate Agronomic Loci for Domestication and Crop Improvement. *The Plant Cell*, *17*(11), 2859–2872. https://doi.org/10.1105/tpc.105.037242

Yang, C. J., Samayoa, L. F., Bradbury, P. J., Olukolu, B. A., Xue, W., York, A. M., Tuholski, M. R., Wang, W., Daskalska, L. L., Neumeyer, M. A., Sanchez-Gonzalez, J. de J., Romay, M. C., Glaubitz, J. C., Sun, Q., Buckler, E. S., Holland, J. B., & Doebley, J. F. (2019). The genetic architecture of teosinte catalyzed and constrained maize domestication. *Proceedings of the National Academy of Sciences*, *116*(12), 5643–5652. https://doi.org/10.1073/pnas.1820997116

Zeibig, F., Kilian, B., & Frei, M. (2022). The grain quality of wheat wild relatives in the evolutionary context. *Theoretical and Applied Genetics*, *135*(11), 4029–4048. https://doi.org/10.1007/s00122-021-04013-8

Zhao, G., Lian, Q., Zhang, Z., Fu, Q., He, Y., Ma, S., Ruggieri, V., Monforte, A. J., Wang, P., Julca, I., Wang, H., Liu, J., Xu, Y., Wang, R., Ji, J., Xu, Z., Kong, W., Zhong, Y., Shang, J., … Huang, S. (2019). A comprehensive genome variation map of melon identifies multiple domestication events and loci influencing agronomic traits. *Nature Genetics*, *51*(11), 1607–1615. https://doi.org/10.1038/s41588-019-0522-8

Zhao Z. (2014). The process of origin agriculture in China: Archeological evidence from flotation results. *Quaternary Sciences*, *34*(1), 73–84. https://doi.org/10.3969/j.issn.1001-7410.2014.10

Zohary, D., & Hopf, M. (1998). *Domestication of plants in the Old World. The origin and spread of cultivated plants in West Asia, Europe and the Nile Valley.* https://doi.org/10.5555/20013014838

Tables and Figures

Figure 1: Domestication-associated traits shared by multiple species. Phylogenetic relationships of the 13 species (adapted from Milla et al., 2020) (a) and shared domestication-associated traits in at least three species (b). The length of the tree branches is not to scale. *: SNIRS1 and SNIRS2 cannot be considered homologous traits among species because NIRS relies on the first two axes of a PCoA, with potentially different wavelengths contributing unequally to each axis across species. Trait abbreviations can be found in Table S3.

Figure 2: Domestication timing and multivariate phenotypic divergence index (mPDI) for the 13 species. Domestication timing (Table S1) is indicated in generations (a). Species are ranked by their mPDI (b) and the mPDI/$Pillai_{control}$ (c).

1162 Figure 3: Convergence in domestication-associated traits. The number of traits measured per
1163 species and shared between species is shown (a) along with the subset of domestication-
1164 associated traits (b).
1165
1166 Figure 4: Domestication syndrome in foxtail millet. Spider-plot for quantitative traits with
1167 significance (*) between wild and domestic forms at an FDR of 5% (a). First two axes of the
1168 PCoA computed from $NIRS_{seed}$ (b) and $NIRS_{leaf}$ (c). In (a) the dots represent the mean values,
1169 with the shaded area showing the standard error. Abbreviations for traits can be found in Table
1170 S3.
1171
1172 Figure 5: Relative size of domestic and wild multivariate phenotypic space. Log ratio of
1173 domestic over wild phenotypic space with positive/negative values indicating larger/smaller
1174 phenotypic space in domestic than in wild samples.
1175
1176 Figure 6: Correlation between similarity of wild and domestic variance-covariance matrices
1177 and domestication timing (Table S1). A matrix similarity of 1 indicates colinear responses to
1178 deformation.
1179

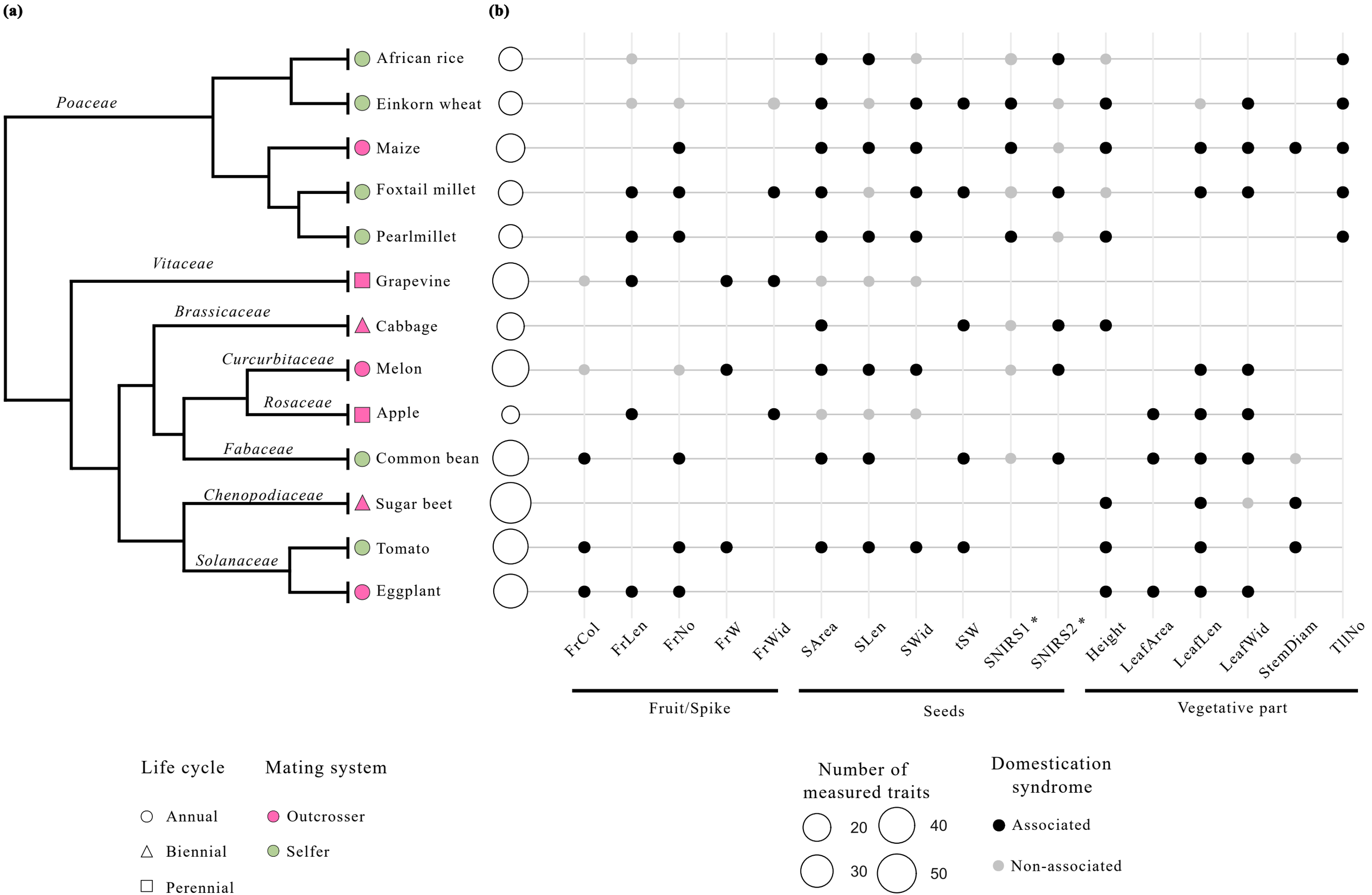
1180 **Supporting Information**

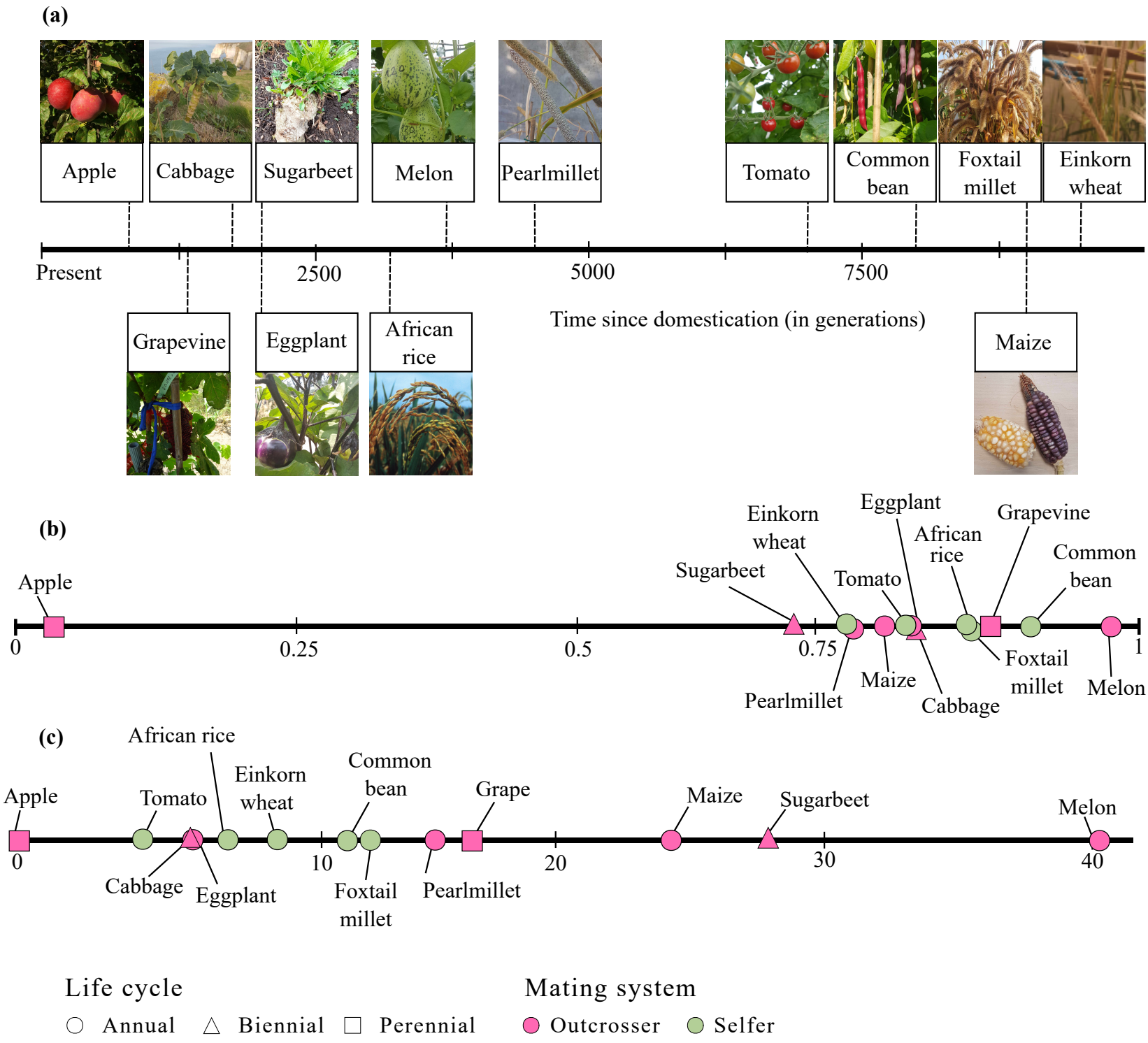
1181
1182 Figure S1: Domestication syndrome in African rice. Spider-plot for quantitative traits with
1183 significance (*) between wild and domestic forms at an FDR of 5% (a). Cleveland plot of
1184 qualitative traits (b). First two axes of the PCoA computed from $NIRS_{seed}$ (c) and $NIRS_{leaf}$ (d).
1185 In (a) the dots represent the mean values, with the shaded area showing the standard error.
1186 Abbreviation meaning of traits can be found in table S3.
1187
1188 Figure S2: Domestication syndrome of apple. Spider-plot for quantitative traits with
1189 significance (*) between wild and domestic forms at an FDR of 5% (a). First two axes of the
1190 PCoA computed from $NIRS_{leaf}$ (b). In (a) the dots represent the mean values and the shaded
1191 areas denote standard error. Abbreviation meaning of traits can be found in table S3.
1192
1193 Figure S3: Domestication syndrome of cabbage. Spider-plot for quantitative traits with
1194 significance (*) between wild and domestic forms at an FDR of 5% (a). Cleveland plot of

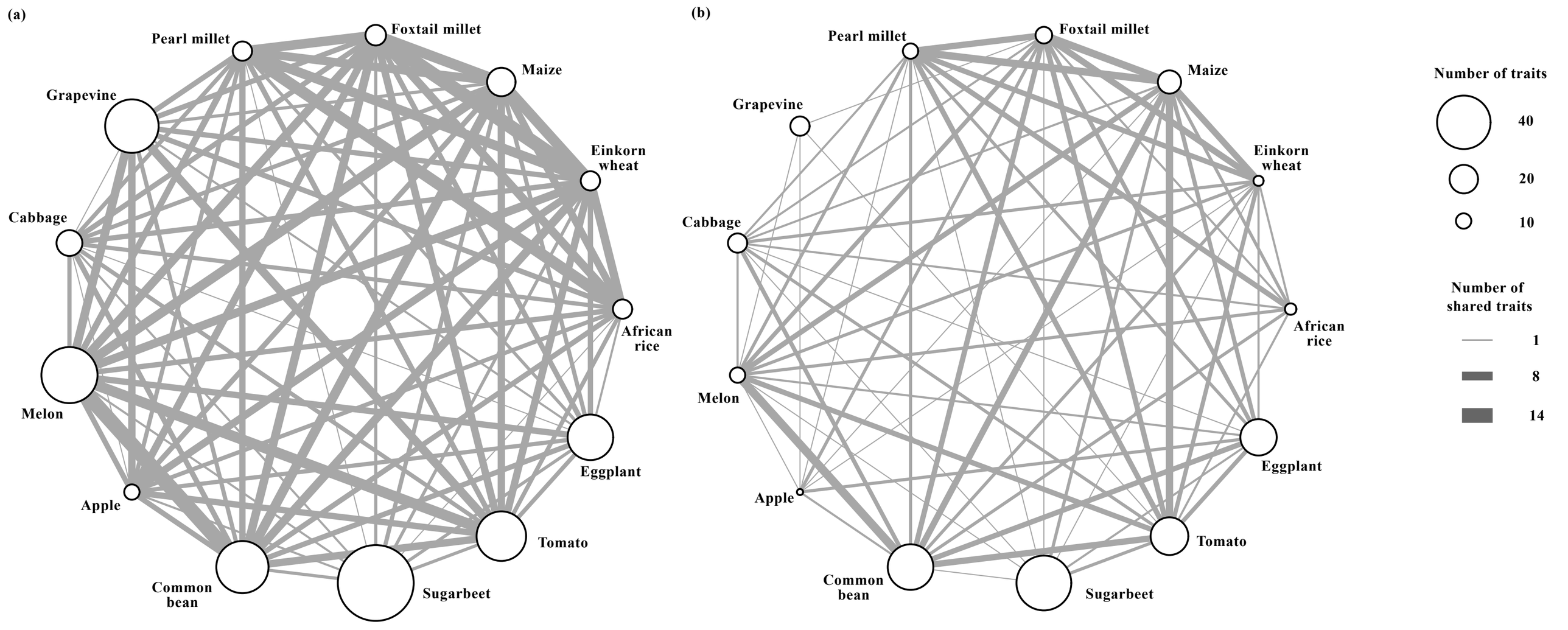
1195 qualitative traits (b). First two axes of the PCoA computed from $NIRS_{seed}$ (c) and $NIRS_{leaf}$ (d).
1196 In (a), the dots represent the mean values and the shaded areas denote standard error.
1197 Abbreviation meaning of traits can be found in table S3.
1198
1199 Figure S4: Domestication syndrome of common bean. Spider-plot for quantitative traits with
1200 significance (*) between wild and domestic forms at an FDR of 5% (a). Cleveland plot of
1201 qualitative traits (b). First two axes of the PCoA computed from $NIRS_{seed}$ (c) and $NIRS_{leaf}$ (d).
1202 In (a), the dots represent the mean values and the shaded areas denote standard error.
1203 Abbreviation meaning of traits can be found in table S3.
1204
1205 Figure S5: Domestication syndrome of eggplant. Spider-plot for quantitative traits with
1206 significance (*) between wild and domestic forms at an FDR of 5% (a). Cleveland plot of
1207 qualitative traits (b). First two axes of the PCoA computed from $NIRS_{leaf}$ (c). In (a), the dots
1208 represent the mean values and the shaded areas denote standard error. Abbreviation meaning
1209 of traits can be found in table S3.
1210
1211 Figure S6: Domestication syndrome of einkorn wheat. Spider-plot for quantitative traits with
1212 significance (*) between wild and domestic forms at an FDR of 5% (a). First two axes of the
1213 PCoA computed from $NIRS_{seed}$ (b) and $NIRS_{leaf}$ (c). In (a), the dots represent the mean values
1214 and the shaded areas denote standard error. Abbreviation meaning of traits can be found in
1215 table S3.
1216
1217 Figure S7: Domestication syndrome of melon. Spider-plot for quantitative traits with
1218 significance (*) between wild and domestic forms at an FDR of 5% (a). Cleveland plot of
1219 qualitative traits (b). First two axes of the PCoA computed from $NIRS_{seed}$ (c) and $NIRS_{leaf}$ (d).
1220 In (a), the dots represent the mean values and the shaded areas denote standard error.
1221 Abbreviation meaning of traits can be found in table S3.
1222
1223 Figure S8: Domestication syndrome of grapevine. Spider-plot for quantitative traits with
1224 significance (*) between wild and domestic forms at an FDR of 5% (a). Cleveland plot of
1225 qualitative traits (b). First two axes of the PCoA computed from $NIRS_{leaf}$ (c). In (a), the dots
1226 represent the mean values and the shaded areas denote standard error. Abbreviation meaning
1227 of traits can be found in table S3.
1228

1229 Figure S9: Domestication syndrome of maize. Spider-plot for quantitative traits with
1230 significance (*) between wild and domestic forms at an FDR of 5% (a). Cleveland plot of
1231 qualitative traits (b). First two axes of the PCoA computed from $NIRS_{seed}$ (c) and $NIRS_{leaf}$ (d).
1232 In (a), the dots represent the mean values and the shaded areas denote standard error.
1233 Abbreviation meaning of traits can be found in table S3.
1234
1235 Figure S10: Domestication syndrome of pearl millet. Spider-plot for quantitative traits with
1236 significance (*) between wild and domestic forms at an FDR of 5% (a). Cleveland plot of
1237 qualitative traits (b). First two axes of the PCoA computed from $NIRS_{seed}$ (c) and $NIRS_{leaf}$ (d).
1238 In (a), the dots represent the mean values and the shaded areas denote standard error.
1239 Abbreviation meaning of traits can be found in table S3.
1240
1241 Figure S11: Domestication syndrome of sugar beet. Spider-plot for quantitative traits with
1242 significance (*) between wild and domestic forms at an FDR of 5% (a). Cleveland plot of
1243 qualitative traits (b). First two axes of the PCoA computed from $NIRS_{leaf}$ (c). In (a), the dots
1244 represent the mean values and the shaded areas denote standard error. Abbreviation meaning
1245 of traits can be found in table S3.
1246
1247 Figure S12: Domestication syndrome of tomato. Spider-plot for quantitative traits with
1248 significance (*) between wild and domestic forms at an FDR of 5% (a). Cleveland plot of
1249 qualitative traits (b). First two axes of the PCoA computed from $NIRS_{leaf}$ (c). In (a), the dots
1250 represent the mean values and the shaded areas denote standard error. Abbreviation meaning
1251 of traits can be found in table S3.
1252
1253 Figure S13: Testing for convergence of domestication-associated traits. Observed number of
1254 domestication traits either unique or shared by two, or three or more species are shown with
1255 the expected number of shared domestication traits as established by a resampling procedure.
1256 Significance was tested by Student t-test (P-values are indicated when significant; * $P < 0.05$;
1257 ** $P < 0.01$).
1258
1259 Figure S14: Convergence of domestication traits. Pseudo-observed distribution and observed
1260 values for domestication-associated traits in our dataset (blue line), for traits either unique, or
1261 shared by two, or three or more species.
1262

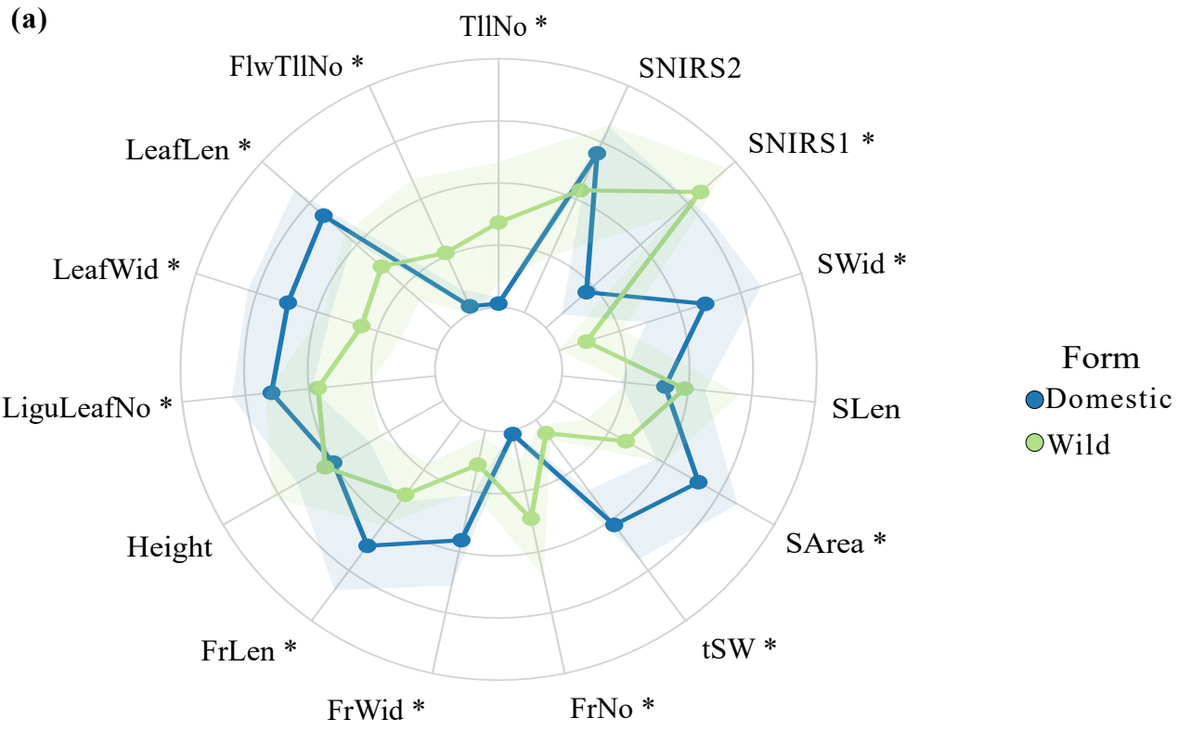
1263 Figure S15: Correlation between the log ratio of domestic to wild $NIRS_{leaf}$ phenotypic space
1264 size with the ratio of domestic to wild genomic diversity (values from the literature, Table S5).
1265
1266 Table S1: Description of the 13 pairs of species with estimate of the domestication timing,
1267 domestication center used in this study, mating system and life cycle.
1268
1269 Table S2: Passport data of sampled accessions.
1270
1271 Table S3: Table S3: Number of traits measured and growing conditions during phenotyping in
1272 each species.
1273
1274 Table S4: List of measured traits for each species, and domestication-associated (DA) traits
1275 with corresponding q-values and percentage of DA-traits in each species.
1276
1277 Table S5: Ratio of domestic (D) over wild (W) multivariate phenotypic space, mDPI,
1278 Pillaicontrol, and estimates from the literature of the ratio (D/W) of genomic diversity and Fst.
1279
1280 Table S6: Average absolute pairwise correlation between traits, computed from all pairwise
1281 correlations in wild and domestic forms. P-values from Student t-tests.

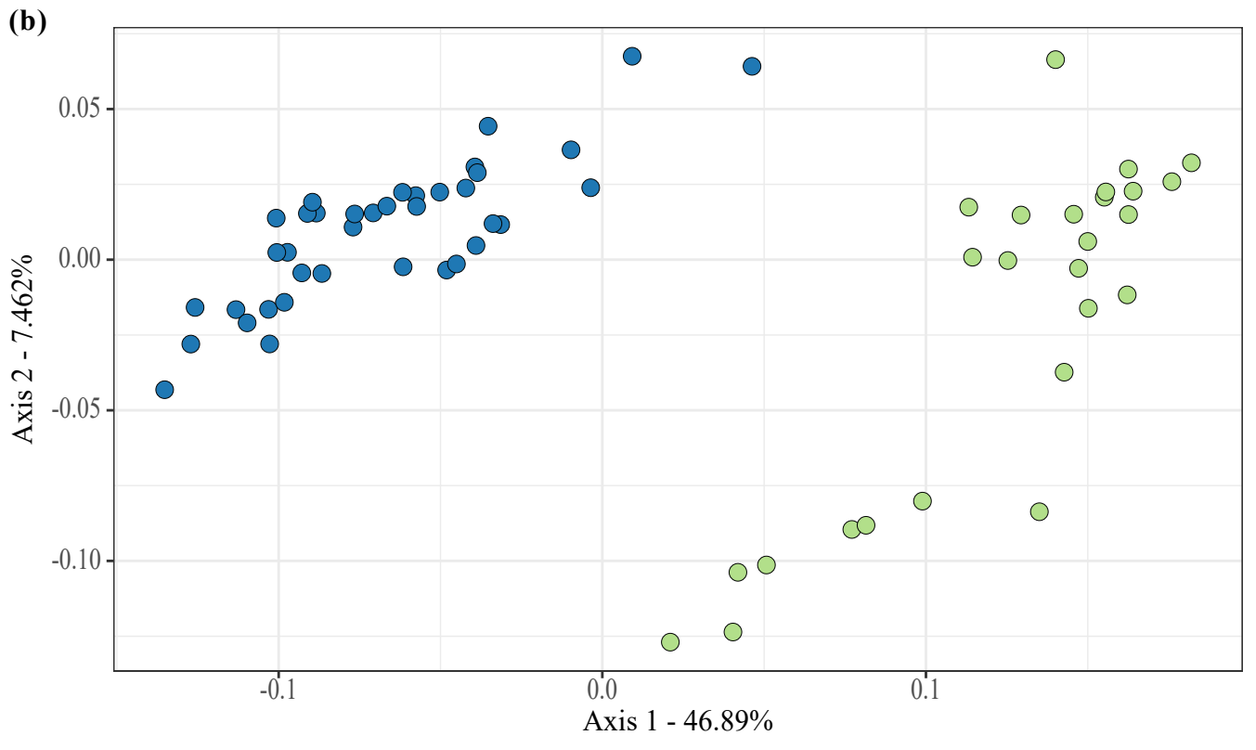

(a)
(b)
Poaceae
Vitaceae
Brassicaceae
Curcurbitaceae
Rosaceae
Fabaceae
Chenopodiaceae
Solanaceae
African rice
Einkorn wheat
Maize
Foxtail millet
Pearlmillet
Grapevine
Cabbage
Melon
Apple
Common bean
Sugar beet
Tomato
Eggplant
FrCol
FrLen
FrNo
FrW
FrWid
SArea
SLen
SWid
tSW
SNIRS1 *
SNIRS2 *
Height
LeafArea
LeafLen
LeafWid
StemDiam
TllNo
Fruit/Spike
Seeds
Vegetative part
Life cycle
Annual
Biennial
Perennial
Mating system
Outcrosser
Selfer
Number of measured traits
20
30
40
50
Domestication syndrome
Associated
Non-associated

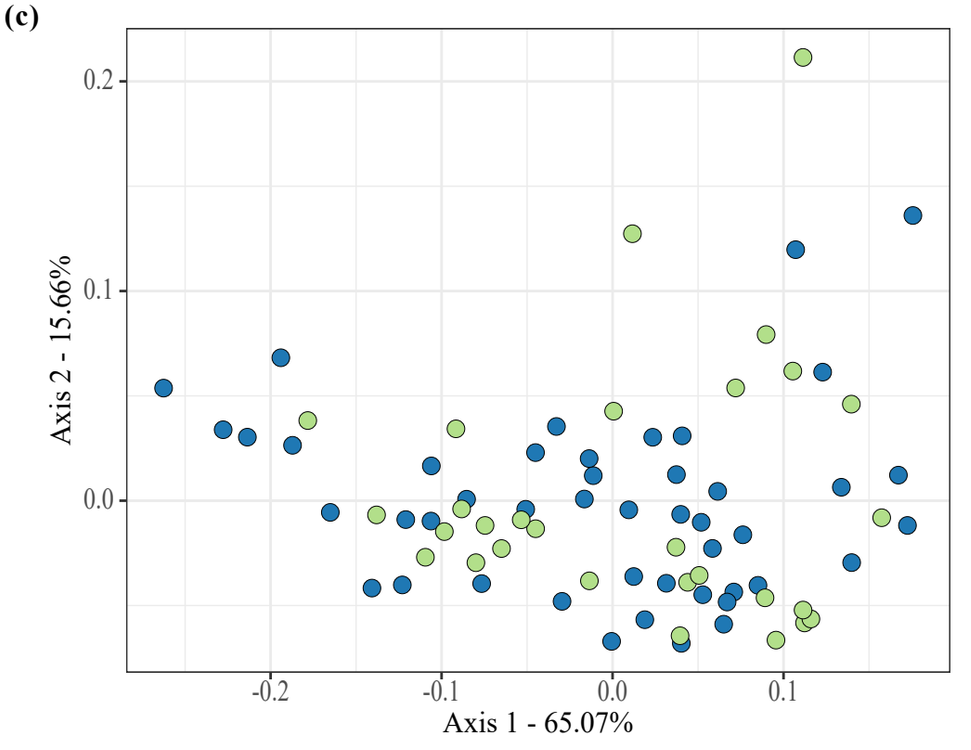

(a)
Apple
Cabbage
Sugarbeet
Melon
Pearlmillet
Tomato
Common bean
Foxtail millet
Einkorn wheat
Present
2500
5000
7500
Time since domestication (in generations)
Grapevine
Eggplant
African rice
Maize
(b)
Apple
Sugarbeet
Einkorn wheat
Tomato
Eggplant
African rice
Grapevine
Common bean
0
0.25
0.5
0.75
1
Pearlmillet
Maize
Cabbage
Foxtail millet
Melon
(c)
Apple
Tomato
African rice
Einkorn wheat
Common bean
Grape
Maize
Sugarbeet
Melon
0
10
20
30
40
Cabbage
Eggplant
Foxtail millet
Pearlmillet
Life cycle
Annual
Biennial
Perennial
Mating system
Outcrosser
Selfer

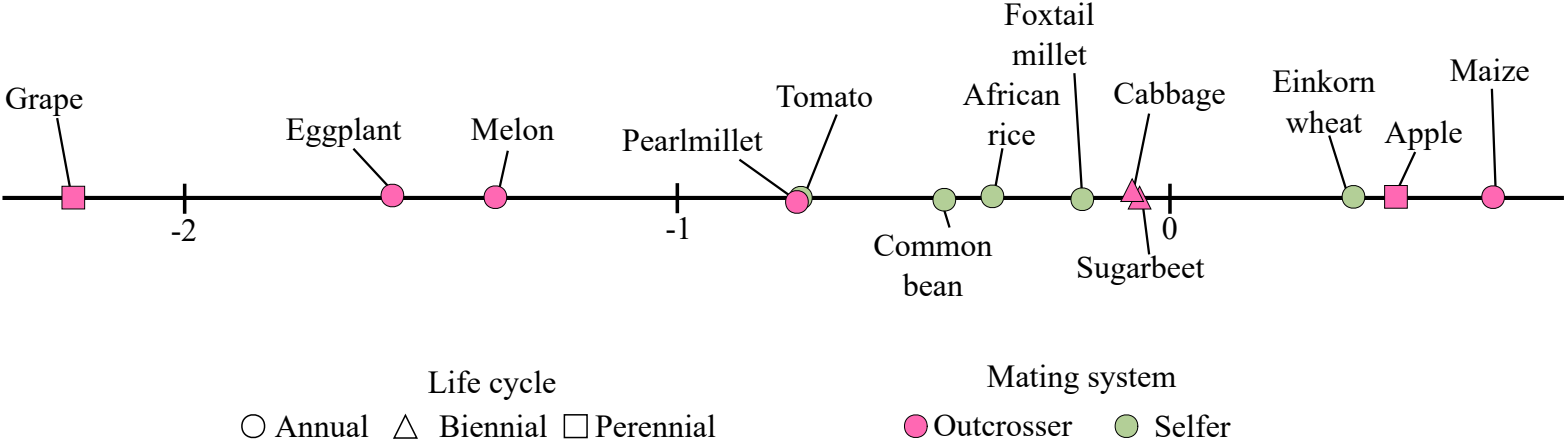

(a)
Pearl millet
Foxtail millet
Maize
Grapevine
Einkorn wheat
Cabbage
African rice
Melon
Eggplant
Apple
Tomato
Common bean
Sugarbeet
(b)
Pearl millet
Foxtail millet
Maize
Grapevine
Einkorn wheat
Cabbage
African rice
Melon
Eggplant
Apple
Tomato
Common bean
Sugarbeet
Number of traits
40
20
10
Number of shared traits
1
8
14

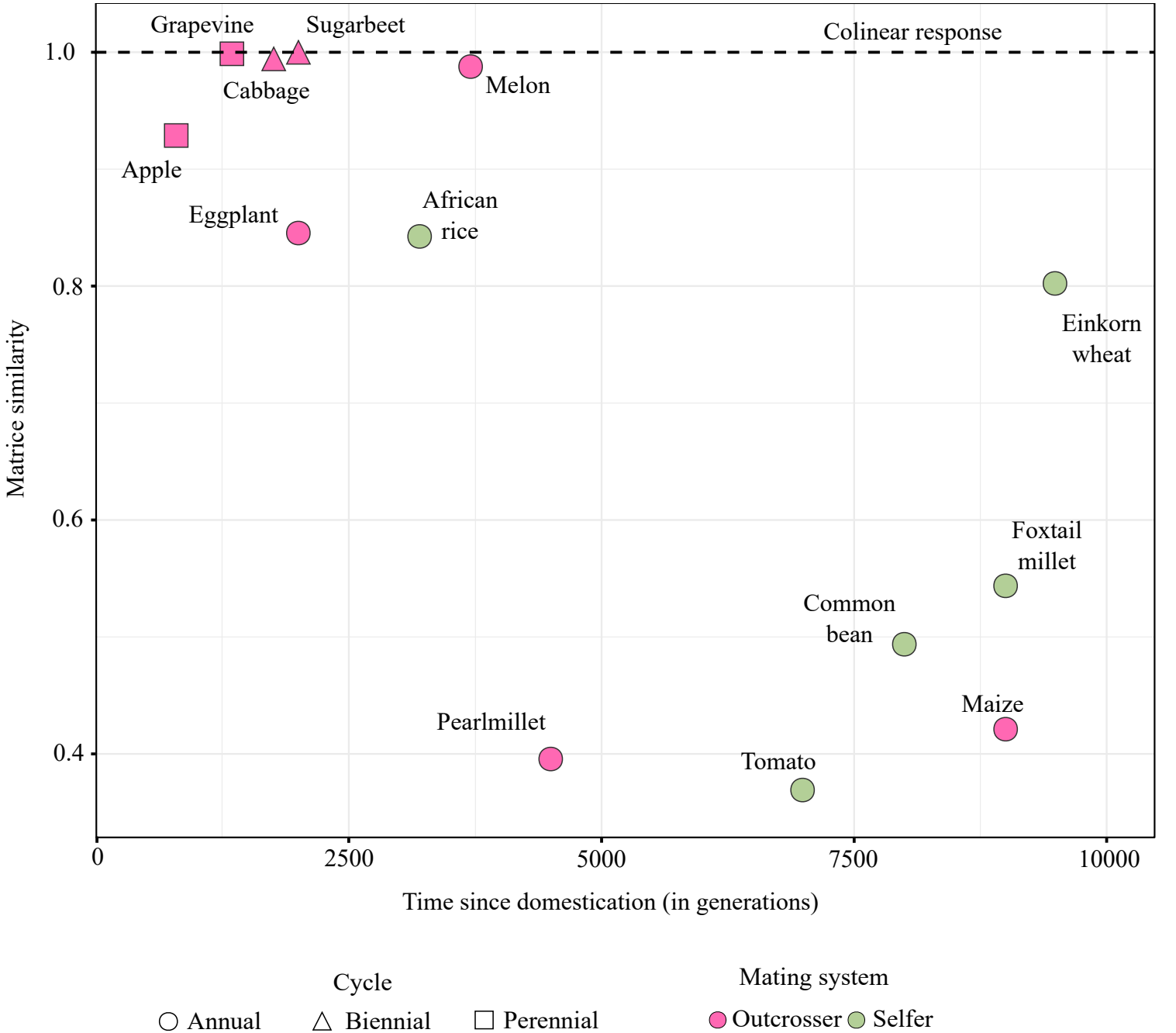
Foxtail millet








Grape
Eggplant
Melon
Pearlmillet
Tomato
African rice
Foxtail millet
Cabbage
Einkorn wheat
Apple
Maize
-2
-1
0
Common bean
Sugarbeet
Life cycle
Annual
Biennial
Perennial
Mating system
Outcrosser
Selfer


Grapevine
Sugarbeet
Colinear response
1.0
Cabbage
Melon
Apple
African
rice
Eggplant
0.8
Einkorn
wheat
Matrice similarity
0.6
Foxtail
millet
Common
bean
Maize
Pearlmillet
0.4
Tomato
0
2500
5000
7500
10000
Time since domestication (in generations)
Cycle
Annual
Biennial
Perennial
Mating system
Outcrosser
Selfer