

**Title:** Common gardens in teosintes reveal the establishment of a syndrome of adaptation to altitude

**Short Title:** Altitudinal syndrome in teosintes

M-A. Fustier<sup>1, ☉</sup>, N.E. Martínez-Ainsworth<sup>1, ☉</sup>, J.A. Aguirre-Liguori<sup>2</sup>, A. Venon<sup>1</sup>, H. Corti<sup>1</sup>, A. Rousselet<sup>1</sup>, F. Dumas<sup>1</sup>, H. Dittberner<sup>3</sup>, M.G. Camarena<sup>4</sup>, D. Grimanelli<sup>5</sup>, O. Ovaskainen<sup>6,7</sup>, M. Falque<sup>1</sup>, L. Moreau<sup>1</sup>, J. de Meaux<sup>3</sup>, S. Montes<sup>4</sup>, L.E. Eguiarte<sup>2</sup>, Y. Vigouroux<sup>5</sup>, D. Manicacci<sup>1,\*</sup>, M.I. Tenaillon<sup>1,\*</sup>

<sup>1</sup>: Génétique Quantitative et Evolution – Le Moulon, Institut National de la Recherche Agronomique, Université Paris-Sud, Centre National de la Recherche Scientifique, AgroParisTech, Université Paris-Saclay, France

<sup>2</sup>: Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Ciudad de México, Mexico

<sup>3</sup>: Institute of Botany, University of Cologne Biocenter, 47b Cologne, Germany

<sup>4</sup>: Campo Experimental Bajío, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Celaya, Mexico

<sup>5</sup>: Université de Montpellier, Institut de Recherche pour le développement, UMR Diversité, Adaptation et Développement des plantes, Montpellier, France

<sup>6</sup>: Organismal and Evolutionary Biology Research Programme, PO Box, 65, FI-00014 University of Helsinki, Finland

<sup>7</sup>: Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

☉: The authors equally contributed to this work

\* co-corresponding authors

E-mail: [domenica.manicacci@inra.fr](mailto:domenica.manicacci@inra.fr), [maud.tenaillon@inra.fr](mailto:maud.tenaillon@inra.fr)

## 1 **Abstract**

2

3           In plants, local adaptation across species range is frequent. Yet, much has to  
4 be discovered on its environmental drivers, the underlying functional traits and their  
5 molecular determinants. Genome scans are popular to uncover outlier loci potentially  
6 involved in the genetic architecture of local adaptation, however links between  
7 outliers and phenotypic variation are rarely addressed. Here we focused on adaptation  
8 of teosinte populations along two elevation gradients in Mexico that display  
9 continuous environmental changes at a short geographical scale. We used two  
10 common gardens, and phenotyped 18 traits in 1664 plants from 11 populations of  
11 annual teosintes. In parallel, we genotyped these plants for 38 microsatellite markers  
12 as well as for 171 outlier single nucleotide polymorphisms (SNPs) that displayed  
13 excess of allele differentiation between pairs of lowland and highland populations  
14 and/or correlation with environmental variables. Our results revealed that phenotypic  
15 differentiation at 10 out of the 18 traits was driven by local selection. Trait covariation  
16 along the elevation gradient indicated that adaptation to altitude results from the  
17 assembly of multiple co-adapted traits into a complex syndrome: as elevation  
18 increases, plants flower earlier, produce less tillers, display lower stomata density and  
19 carry larger, longer and heavier grains. The proportion of outlier SNPs associating  
20 with phenotypic variation, however, largely depended on whether we considered a  
21 neutral structure with 5 genetic groups (73.7%) or 11 populations (13.5%), indicating  
22 that population stratification greatly affected our results. Finally, chromosomal  
23 inversions were enriched for both SNPs whose allele frequencies shifted along  
24 elevation as well as phenotypically-associated SNPs. Altogether, our results are  
25 consistent with the establishment of an altitudinal syndrome promoted by local

26 selective forces in teosinte populations in spite of detectable gene flow. Because  
27 elevation mimics climate change through space, SNPs that we found underlying  
28 phenotypic variation at adaptive traits may be relevant for future maize breeding.

29

30 **Keywords:** spatially-varying selection;  $F_{ST}$ -scan; association mapping; altitudinal  
31 syndrome; pleiotropy; chromosomal inversions.

32

### 33 **Author summary**

34 Across their native range species encounter a diversity of habitats promoting local  
35 adaptation of geographically distributed populations. While local adaptation is  
36 widespread, much has yet to be discovered about the conditions of its emergence, the  
37 targeted traits, their molecular determinants and the underlying ecological drivers.  
38 Here we employed a reverse ecology approach, combining phenotypes and genotypes,  
39 to mine the determinants of local adaptation of teosinte populations distributed along  
40 two steep altitudinal gradients in Mexico. Evaluation of 11 populations in two  
41 common gardens located at mid-elevation pointed to adaptation via an altitudinal  
42 multivariate syndrome, in spite of gene flow. We scanned genomes to identify loci  
43 with allele frequencies shifts along elevation, a subset of which associated to trait  
44 variation. Because elevation mimics climate change through space, these  
45 polymorphisms may be relevant for future maize breeding.

## 46 **Introduction**

47

48           Local adaptation is key for the preservation of ecologically useful genetic  
49 variation [1]. The conditions for its emergence and maintenance have been the focus  
50 of a long-standing debate nourished by ample theoretical work [2-9]. On the one  
51 hand, spatially-varying selection promotes the evolution of local adaptation, provided  
52 that there is genetic diversity underlying the variance of fitness-related traits [10]. On  
53 the other hand, opposing forces such as neutral genetic drift, temporal fluctuations of  
54 natural selection, recurrent introduction of maladaptive alleles via migration and  
55 homogenizing gene flow may hamper local adaptation (reviewed in [11])□. Meta-  
56 analyzes indicate that local adaptation is pervasive in plants, with evidence of native-  
57 site fitness advantage in reciprocal transplants detected in 45% to 71% of populations  
58 [12, 13].

59           While local adaptation is widespread, much has yet to be discovered about the  
60 traits affected by spatially-varying selection, their molecular determinants and the  
61 underlying ecological drivers [14]. Local adaptation is predicted to increase with  
62 phenotypic, genotypic and environmental divergence among populations [6, 15, 16].  
63 Comparisons of the quantitative genetic divergence of a trait ( $Q_{ST}$ ) with the neutral  
64 genetic differentiation ( $F_{ST}$ ) can provide hints on whether trait divergence is driven by  
65 spatially-divergent selection [17-20]. Striking examples of divergent selection include  
66 developmental rate in the common toad [21], drought and frost tolerance in alpine  
67 populations of the European silver fir [22], and traits related to plant phenology, size  
68 and floral display among populations of *Helianthus* species [23, 24]. These studies  
69 have reported covariation of physiological, morphological and/or life-history traits  
70 across environmental gradients which collectively define adaptive syndromes. Such

71 syndromes may result from several non-exclusive mechanisms: plastic responses,  
72 pleiotropy, non-adaptive genetic correlations among traits (constraints), and joint  
73 selection of traits encoded by different sets of genes resulting in adaptive correlations.  
74 In some cases, the latter mechanism may involve selection and rapid spread of  
75 chromosomal inversions that happen to capture multiple locally favored alleles [25] as  
76 exemplified in *Mimulus guttatus* [26]. While distinction between these mechanisms is  
77 key to decipher the evolvability of traits, empirical data on the genetic bases of  
78 correlated traits are currently lacking [27].

79         The genes mediating local adaptation are usually revealed by genomic regions  
80 harboring population-specific signatures of selection. These signatures include alleles  
81 displaying greater-than-expected differentiation among populations [28] and can be  
82 identified through  $F_{ST}$ -scans [29-35]. However,  $F_{ST}$ -scans and its derivative methods  
83 [28] suffer from a number of limitations, among them a high number of false positives  
84 (reviewed in [36, 37]) and the lack of power to detect true positives [38]. Despite  
85 these caveats,  $F_{ST}$ -outlier approaches have helped in the discovery of emblematic  
86 adaptive alleles such as those segregating at the *EPASI* locus in Tibetan human  
87 populations adapted to high altitude [39]. An alternative to detect locally adaptive loci  
88 is to test for genotype-environment correlations [35, 40-45]. Correlation-based  
89 methods can be more powerful than differentiation-based methods [46], but spatial  
90 autocorrelation of population structure and environmental variables can lead to  
91 spurious signatures of selection [47].

92         Ultimately, to identify the outlier loci that have truly contributed to improve  
93 local fitness, a link between outliers and phenotypic variation needs to be established.  
94 The most common approach is to undertake association mapping. However, recent  
95 literature in humans has questioned our ability to control for sample stratification in

96 such approach [48]. Detecting polymorphisms responsible for trait variation is  
97 particularly challenging when trait variation and demographic history follow parallel  
98 environmental (geographic) clines. Plants however benefit from the possibility of  
99 conducting replicated phenotypic measurements in common gardens, where  
100 environmental variation is controlled. Hence association mapping has been  
101 successfully employed in the model plant species *Arabidopsis thaliana*, where  
102 broadly distributed ecotypes evaluated in replicated common gardens have shown that  
103 fitness-associated alleles display geographic and climatic patterns indicative of  
104 selection [49]. Furthermore, the relative fitness of *A. thaliana* ecotypes in a given  
105 environment could be predicted from climate-associated SNPs [50]. While climatic  
106 selection over broad latitudinal scales produces genomic and phenotypic patterns of  
107 local adaptation in the selfer plant *A. thaliana*, whether similar patterns exist at shorter  
108 spatial scale in outcrossing species remains to be elucidated.

109         We focused here on a well-established outcrossing plant system, the teosintes,  
110 to investigate the relationship of molecular, environmental, and phenotypic variation  
111 in populations sampled across two elevation gradients in Mexico. The gradients  
112 covered a relatively short yet climatically diverse, spatial scale. They encompassed  
113 populations of two teosinte subspecies that are the closest wild relatives of maize, *Zea*  
114 *mays* ssp. *parviglumis* (hereafter *parviglumis*) and *Z. mays* ssp. *mexicana* (hereafter  
115 *mexicana*). The two subspecies display large effective population sizes [51], and span  
116 a diversity of climatic conditions, from warm and mesic conditions below 1800 m for  
117 *parviglumis*, to drier and cooler conditions up to 3000 m for *mexicana* [52]. Previous  
118 studies have discovered potential determinants of local adaptation in these systems.  
119 At a genome-wide scale, decrease in genome size correlates with increasing altitude,  
120 which likely results from the action of natural selection on life cycle duration [53, 54].

121 More modest structural changes include megabase-scale inversions that harbor  
122 clusters of SNPs whose frequencies are associated with environmental variation [55,  
123 56]. Also, differentiation- and correlation-based genome scans in teosinte populations  
124 succeeded in finding outlier SNPs potentially involved in local adaptation [57, 58].  
125 But a link with phenotypic variation has yet to be established.

126 In this paper, we genotyped a subset of these outlier SNPs on a broad sample  
127 of 28 teosinte populations, for which a set of neutral SNPs was also available; as well  
128 as on an association panel encompassing 11 populations. We set up common gardens  
129 in two locations to evaluate the association panel for 18 phenotypic traits over two  
130 consecutive years. Individuals from this association panel were also genotyped at 38  
131 microsatellite markers to enable associating genotypic to phenotypic variation while  
132 controlling for sample structure and kinship among individuals. We addressed three  
133 main questions: What is the extent of phenotypic variation within and among  
134 populations? Can we define a set of locally-selected traits that constitute a syndrome  
135 of adaptation to altitude? What are the genetic bases of such syndrome? We further  
136 discuss the challenges of detecting phenotypically-associated SNPs when trait and  
137 genetic differentiation parallel environmental clines.

138

## 139 **Results**

140

### 141 **Trait-by-trait analysis of phenotypic variation within and among populations.**

142 In order to investigate phenotypic variation, we set up two common garden  
143 experiments located in Mexico to evaluate individuals from 11 teosinte populations  
144 (Fig 1). The two experimental fields were chosen because they were located at  
145 intermediate altitudes (S1 Fig). Although natural teosinte populations are not typically

146 encountered around these locations [52], we verified that environmental conditions  
147 were compatible with both subspecies (S2 Fig). The 11 populations were sampled  
148 among 37 populations (S1 Table) distributed along two altitudinal gradients that range  
149 from 504 to 2176 m in altitude over ~460 kms for gradient *a*, and from 342 to 2581m  
150 in altitude over ~350 kms for gradient *b* (S1 Fig). Lowland populations of the  
151 subspecies *parviglumis* (n=8) and highland populations of the subspecies *mexicana*  
152 (n=3) were climatically contrasted as can be appreciated in the Principal Component  
153 Analysis (PCA) computed on 19 environmental variables (S2 Fig). The corresponding  
154 set of individuals grown from seeds sampled from the 11 populations formed the  
155 association panel.

156

157 **Figure 1. Geographical location of sampled populations and experimental fields.**

158 The entire set of 37 Mexican teosinte populations is shown with *parviglumis* (circles)  
159 and *mexicana* (triangles) populations sampled along gradient *a* (white) and gradient *b*  
160 (black). The 11 populations indicated with a purple outline constituted the association  
161 panel. This panel was evaluated in a four-block design over two years in two  
162 experimental fields located at mid-elevation, SENGUA and CEBAJ. Two major cities  
163 (Mexico City and Guadalajara) are also indicated. Topographic surfaces have been  
164 obtained from International Centre for Tropical Agriculture (Jarvis A., H.I. Reuter, A.  
165 Nelson, E. Guevara, 2008, Hole-filled seamless SRTM data V4, International Centre  
166 for Tropical Agriculture (CIAT), available from <http://srtm.csi.cgiar.org>).

167

168 We gathered phenotypic data during two consecutive years (2013 and 2014).  
169 We targeted 18 phenotypic traits that included six traits related to plant architecture,  
170 three traits related to leaves, three traits related to reproduction, five traits related to  
171 grains, and one trait related to stomata (S2 Table). Each of the four experimental



172 assays (year-field combinations) encompassed four blocks. In each block, we  
173 evaluated one offspring (half-sibs) of ~15 mother plants from each of the 11 teosinte  
174 populations using a semi-randomized design. After filtering for missing data, the  
175 association panel included 1664 teosinte individuals. We found significant effects of  
176 Field, Year and/or their interaction for most traits, and a highly significant Population  
177 effect for all of them (model M1, S3 Table).

178 We investigated the influence of altitude on each trait independently. All traits,  
179 except for the number of nodes with ears (NoE), exhibited a significant effect of  
180 altitude (S3 Table, M3 model). Note that after accounting for elevation, the  
181 population effect remained significant for all traits, suggesting that factors other than  
182 altitude contributed to shape phenotypic variation among populations. Traits related to  
183 flowering time and tillering displayed a continuous decrease with elevation, and traits  
184 related to grain size increased with elevation (Fig 2 & S3 Fig). Stomata density also  
185 diminished with altitude (Fig 2). In contrast, plant height, height of the highest ear,  
186 number of nodes with ear in the main tiller displayed maximum values at intermediate  
187 altitudes (highland *parviglumis* and lowland *mexicana*) (S3 Fig).

188 We estimated narrow-sense heritabilities (additive genotypic effect) per  
189 population for all traits using a mixed animal model. Average per-trait heritability  
190 ranged from 0.150 for tassel branching to 0.664 for female flowering time, albeit with  
191 large standard errors (S2 Table). We obtained higher heritability for grain related  
192 traits when mother plant measurements were included in the model with 0.631 ( $sd =$   
193 0.246), 0.511 ( $sd = 0.043$ ) and 0.274 ( $sd = 0.160$ ) for grain length, weight and width,  
194 respectively, suggesting that heritability was under-estimated for other traits where  
195 mother plant values were not available.

196

197 **Figure 2: Population-level box-plots of adjusted means for four traits.** Traits are  
198 female flowering time (A), male flowering time (B), grain length (C) and stomata  
199 density (D). Populations are ranked by altitude. *Parviglumis* populations are shown in  
200 green and *mexicana* in red, lighter colors are used for gradient ‘a’ and darker colors  
201 for gradient ‘b’. In the case of male and female flowering time, we report data for 9  
202 out of 11 populations because most individuals from the two lowland populations  
203 (P1a and P1b) did not flower in our common gardens. Covariation with elevation was  
204 significant for the four traits. Corrections for experimental setting are detailed in the  
205 material and methods section (Model M’1).

206

207 **Multivariate analysis of phenotypic variation and correlation between traits.**

208 Principal component analysis including all phenotypic measurements  
209 highlighted that 21.26% of the phenotypic variation scaled along PC1 (Fig 3A), a PC  
210 axis that is strongly collinear with altitude (Fig 3B). Although populations partly  
211 overlapped along PC1, we observed a consistent tendency for population phenotypic  
212 differentiation along altitude irrespective of the gradient (Fig 3C). Traits that  
213 correlated the most to PC1 were related to grain characteristics, tillering, flowering  
214 and to a lesser extent to stomata density (Fig 3B). PC2 correlated with traits  
215 exhibiting a trend toward increase-with-elevation within *parviglumis*, but decrease-  
216 with-elevation within *mexicana* (Fig 3D). Those traits were mainly related to  
217 vegetative growth (Fig 3B). Together, both axes explained 37% of the phenotypic  
218 variation.

219

220 **Figure 3: Principal Component Analysis on phenotypic values corrected for the**  
221 **experimental setting.** Individuals factor map (A) and corresponding correlation

222 circle (B) on the first two principal components with altitude (Alt) added as a  
223 supplementary variable (in blue). Individual phenotypic values on PC1 (C) and PC2  
224 (D) are plotted against population ranked by altitude and color-coded following A.  
225 For populations from the two subspecies, *parviglumis* (circles) and *mexicana*  
226 (triangles), color intensity indicates ascending elevation in green for *parviglumis* and  
227 red for *mexicana*. Corrections for experimental setting are detailed in the material and  
228 methods (Model M2).

229

230 We assessed more formally pairwise-correlations between traits after  
231 correcting for experimental design and population structure ( $K=5$ ). We found 82  
232 (54%) significant correlations among 153 tested pairs of traits. The following pairs of  
233 traits had the strongest positive correlations: male and female flowering time, plant  
234 height and height of the highest ear, height of the highest and lowest ear, grain length  
235 with grain weight and width (S4 Fig). The correlation between flowering time (female  
236 or male) with grain weight and length were among the strongest negative correlations  
237 (S4 Fig).

238

### 239 **Neutral structuring of the association panel.**

240 We characterized the genetic structure of the association panel using SSRs.  
241 The highest likelihood from Bayesian classification was obtained at  $K=2$  and  $K=5$   
242 clusters (S5 Fig). At  $K=2$ , the clustering separated the lowland of gradient *a* from the  
243 rest of the populations. From  $K=3$  to  $K=5$ , a clear separation between the eight  
244 *parviglumis* and the three *mexicana* populations emerged. Increasing  $K$  values finally  
245 split the association panel into the 11 populations it encompassed (S6 Fig). The  $K=5$   
246 structure associated to both altitude (lowland *parviglumis* versus highland *mexicana*)

247 and gradients *a* and *b* (Fig 4A & B). TreeMix analysis for a subset of 10 of these  
248 populations confirmed those results with an early split separating the lowlands from  
249 gradient *a* (cf.  $K=2$ , S6 Fig) followed by the separation of the three *mexicana* from the  
250 remaining populations (Fig 4C). TreeMix further supported three migration edges, a  
251 model that explained 98.75% of the variance and represented a significant  
252 improvement over a model without admixture (95.7%, Figure S7). This admixture  
253 model was consistent with gene flow between distant lowland *parviglumis*  
254 populations from gradient *a* and *b*, as well as between *parviglumis* and *mexicana*  
255 populations (Fig 4C). Likewise, structure analysis also suggested admixture among  
256 some of the lowland populations, and to a lesser extent between the two subspecies  
257 (Fig 4B).

258

259 **Figure 4: Genetic clustering, historical splits and admixture among populations**  
260 **of the association panel.** Genetic clustering visualization based on 38 SSRs is shown  
261 for  $K=5$  (A). Colors represent the  $K$  clusters. Individuals (vertical lines) are  
262 partitioned into colored segments whose length represents the membership  
263 proportions to the  $K$  clusters. Populations (named after the subspecies M: *mexicana*,  
264 P: *parviglumis* and gradient ‘a’ or ‘b’) are ranked by altitude indicated in meters  
265 above sea level. The corresponding geographic distribution of populations along with  
266 their average membership probabilities are plotted (B). Historical splits and  
267 admixtures between populations were inferred from SNP data for a subset of 10  
268 populations of the association panel (C). Admixtures are colored according to their  
269 weight.

270

271

272 **Identification of traits evolving under spatially-varying selection.**

273 We estimated the posterior mean (and 95% credibility interval) of genetic  
274 differentiation ( $F_{ST}$ ) among the 11 populations of the association panel using  
275 DRIFTSEL. Considering 1125 plants for which we had both individual phenotypes and  
276 individual genotypes for 38 SSRs (S4 Table), we estimated the mean  $F_{ST}$  to 0.22  
277 (0.21-0.23). Note that we found a similar estimate on a subset of 10 of these  
278 populations with 1000 neutral SNPs ( $F_{ST}$  (CI)=0.26 (0.25-0.27)). To identify traits  
279 whose variation among populations was driven primarily by local selection, we  
280 employed the Bayesian method implemented in DRIFTSEL, that infers additive genetic  
281 values of traits from a model of population divergence under drift [59]. Selection was  
282 inferred when observed phenotypic differentiation exceeded neutral expectations for  
283 phenotypic differentiation under random genetic drift. Single-trait analyses revealed  
284 evidence for spatially-varying selection at 12 traits, with high consistency between  
285 SSRs and neutral SNPs (Table 1). Another method that contrasted genetic and  
286 phenotypic differentiation ( $Q_{ST}$ -  $F_{ST}$ ) uncovered a large overlap with nine out of the  
287 12 traits significantly deviating from the neutral model (Table 1) and one of the  
288 remaining ones displaying borderline significance (Plant height=PL, S8 Fig).  
289 Together, these two methods indicated that phenotypic divergence among populations  
290 was driven by local selective forces.

**Table 1.** Signals of selection (posterior probability S) for each trait considering SSR markers (11 populations) or SNPs (10 populations).

Traits <sup>a</sup>	SSR <sup>b</sup>	SNP <sup>b</sup>
<b><u>Plant height</u></b>	0.995	0.972
<b><u>Height of the lowest ear</u>*</b>	0.950	0.959
<b>Height of the highest ear</b>	0.982	0.966
<b><u>Number of tillers</u>*</b>	1.000	1.000
<b><u>Number of lateral branches</u>*</b>	1.000	0.990
Number of nodes with ears	0.682	0.699
Leaf length	0.888	0.875
<b>Leaf width</b>	0.999	0.996
Leaf color	0.633	0.583
<b><u>Female flowering time</u>*</b>	1.000	1.000
<b><u>Male flowering time</u>*</b>	1.000	1.000
Tassel branching*	0.925	0.908
Number of grains per ear	0.832	0.622
<b><u>Grain length</u>*</b>	1.000	1.000
<b><u>Grain width</u>*</b>	0.995	0.984
<b><u>Grain weight</u>*</b>	1.000	0.999
Grain color	0.717	0.689
<b><u>Stomata density</u>*</b>	0.999	0.999

<sup>a</sup>: Traits displaying signal of selection (spatially-varying traits,  $S > 0.95$ ) are indicated in bold, and marked by an asterisk when significant in  $Q_{ST}-F_{ST}Comp$  analysis. We considered the underlined traits as spatially varying. For a detailed description of traits see S2 Table.

<sup>b</sup>: Values reported correspond to  $S$  from DRIFTSEL.  $S$  is the posterior probability that divergence among populations was not driven by drift only. Following [60], we used here a conservative credibility value of  $S > 0.95$  to declare divergent selection.

291           Altogether, evidence of spatially varying selection at 10 traits (Table 1) as  
292 well as continuous variation of a subset of traits across populations in both elevation  
293 gradients (Fig 2, S3 Fig) was consistent with a syndrome where populations produced  
294 less tillers, flowered earlier, displayed lower stomata density and carried larger,  
295 longer and heavier grains with increasing elevation.

296

#### 297 **Outlier detection and correlation with environmental variables.**

298           We successfully genotyped 218 (~81%) out of 270 outlier SNPs on a broad set  
299 of 28 populations, of which 141 were previously detected in candidate regions for  
300 local adaptation [58]. Candidate regions were originally identified from re-sequencing  
301 data of only six teosinte populations (S1 Table) following an approach that included  
302 high differentiation between highlands and lowlands, environmental correlation, and  
303 in some cases their intersection with genomic regions involved in quantitative trait  
304 variation in maize. The remaining outlier SNPs (77) were discovered in the present  
305 study by performing  $F_{ST}$ -scans on the same re-sequencing data (S5 Table). We  
306 selected outlier SNPs that were both highly differentiated between highland and  
307 lowland populations within gradient (high/low in gradient  $a$  or  $b$  or both), and  
308 between highland and lowland populations within subspecies in gradient  $b$  (high/low  
309 within *parviglumis*, *mexicana* or both).  $F_{ST}$ -scans pinpointed three genomic regions of  
310 particularly high differentiation (S9 Fig) that corresponded to previously described

311 inversions [55, 56]: one inversion on chromosome 1 (*Inv1n*), one on chromosome 4  
312 (*Inv4m*) and one on the far end of chromosome 9 (*Inv9e*).

313 A substantial proportion of outlier SNPs was chosen based on their significant  
314 correlation among six populations between variation of allele frequency and their  
315 coordinate on the first environmental principal component [58]. We extended  
316 environmental analyses to 171 outlier SNPs (MAF>5%) on a broader sample of 28  
317 populations (S2 Fig) and used the two first components (PCenv1 and PCenv2) to  
318 summarize environmental information. The first component, that explained 56% of  
319 the variation, correlated with altitude but displayed no correlation to either latitude or  
320 longitude. PCenv1 was defined both by temperature- and precipitation- related  
321 variables (S2 B Fig) including Minimum Temperature of Coldest Month (T6), Mean  
322 Temperature of Driest and Coldest Quarter (T9 and T11) and Precipitation of Driest  
323 Month and Quarter (P14 and P17). The second PC explained 20.5% of the variation  
324 and was mainly defined (S2 B Fig) by Isothermality (T3), Temperature Seasonality  
325 (T4) and Temperature Annual Range (T7).

326 We first employed multiple regression to test for each SNP, whether the  
327 pairwise  $F_{ST}$  matrix across 28 populations correlated to the environmental (distance  
328 along PCenv1) and/or the geographical distance. As expected, we found a  
329 significantly greater proportion of environmentally-correlated SNPs among outliers  
330 compared with neutral SNPs ( $\chi^2 = 264.07$ , P-value =  $2.2 \cdot 10^{-16}$ ), a pattern not seen with  
331 geographically-correlated SNPs. That outlier SNPs displayed a greater isolation-by-  
332 environment than isolation-by-distance, indicated that patterns of allele frequency  
333 differentiation among populations were primarily driven by adaptive processes. We  
334 further tested correlations between allele frequencies and environmental variation.  
335 Roughly 60.8% (104) of the 171 outlier SNPs associated with at least one of the two



336 first PCenvs, with 87 and 33 associated with PCenv1 and PCenv2, respectively, and  
337 little overlap (S5 Table). As expected, the principal component driven by altitude  
338 (PCenv1) correlated to allele frequency for a greater fraction of SNPs than the second  
339 orthogonal component. Interestingly, we found enrichment of environmentally-  
340 associated SNPs within inversions both for PCenv1 ( $\chi^2 = 14.63$ , P-value= $1.30 \cdot 10^{-4}$ )  
341 and PCenv2 ( $\chi^2 = 33.77$ , P-value= $6.22 \cdot 10^{-9}$ ).

342

### 343 **Associating genotypic variation to phenotypic variation.**

344 We tested the association between phenotypes and 171 of the outlier SNPs  
345 (MAF>5%) using the association panel. For each SNP-trait combination, the sample  
346 size ranged from 264 to 1068, with a median of 1004 individuals (S6 Table). We used  
347 SSRs to correct for both structure (at  $K=5$ ) and kinship among individual genotypes.  
348 This model (M5) resulted in a uniform distribution of P-values when testing the  
349 association between genotypic variation at SSRs and phenotypic trait variation (S10  
350 Fig). Under this model, we found that 126 outlier SNPs (73.7%) associated to at least  
351 one trait (Fig 5 and S11 Fig) at an FDR of 10%. The number of associated SNPs per  
352 trait varied from 0 for leaf and grain coloration, to 55 SNPs for grain length, with an  
353 average of 22.6 SNPs per trait (S5 Table). Ninety-three (73.8%) out of the 126  
354 associated SNPs were common to at least two traits, and the remaining 33 SNPs were  
355 associated to a single trait (S5 Table). The ten traits displaying evidence of spatially  
356 varying selection in the  $Q_{ST}-F_{ST}$  analyses displayed more associated SNPs per trait  
357 (30.5 on average), than the non-spatially varying traits (12.75 on average).

358 A growing body of literature stresses that incomplete control of population  
359 stratification may lead to spurious associations [61]. Hence, highly differentiated  
360 traits along environmental gradients are expected to co-vary with any variant whose

361 allele frequency is differentiated along the same gradients, without underlying causal  
362 link. We therefore expected false positives in our setting where both phenotypic traits  
363 and outlier SNPs varied with altitude. We indeed found a slightly significant  
364 correlation ( $r=0.5$ ,  $P\text{-value}=0.03$ ) between the strength of the population effect for  
365 each trait – a measure of trait differentiation (S3 Table) – and its number of associated  
366 SNPs (S5 Table).

367 To verify that additional layers of structuring among populations did not cause  
368 an excess of associations, we repeated the association analyzes considering a  
369 structuring with 11 populations (instead of  $K=5$ ) as covariate (M5'), a proxy of the  
370 structuring revealed at  $K=11$  (S6 Fig). With this level of structuring, we retrieved  
371 much less associated SNPs (S5 Table). Among the 126 SNPs associating with at least  
372 one trait at  $K=5$ , only 22 were recovered considering 11 populations. An additional  
373 SNP was detected with structuring at 11 populations that was absent at  $K=5$ . Eight  
374 traits displayed no association, and the remaining traits varied from a single  
375 associated SNP (Leaf length – LeL and the number of tillers – Til) to 8 associated  
376 SNPs for grain weight (S5 Table). For instance, traits such as female or male  
377 flowering time that displayed 45 and 43 associated SNPs at  $K=5$ , now displayed only  
378 4 and 3 associated SNPs, respectively (Fig 5). Note that one trait (Leaf color)  
379 associated with 4 SNPs considering 11 populations while displaying no association at  
380  $K=5$ . Significant genetic associations were therefore highly contingent on the  
381 population structure. Noteworthy, traits under spatially varying selection still  
382 associated with more SNPs (2.00 on average) than those with no spatially varying  
383 selection (1.25 SNPs on average).

384

385

386 **Figure 5: Manhattan plots of associations between 171 outlier SNPs and 6**  
387 **phenotypic traits.** X-axis indicates the positions of outlier SNPs on chromosomes 1  
388 to 10, black and gray colors alternating per chromosome. Plotted on the Y-axis are the  
389 negative  $\text{Log}_{10}$ -transformed  $P$  values obtained for the  $K=5$  model. Significant  
390 associations (10% FDR) are indicated considering either a structure matrix at  $K=5$   
391 (pink dots), 11 populations (blue dots) or both  $K=5$  and 11 populations (purple dots)  
392 models.

393

394 Altogether the 23 SNPs recovered considering a neutral genetic structure with  
395 11 populations corresponded to 30 associations, 7 of the SNPs being associated to  
396 more than one trait (S5 Table). For all these 30 associations except in two cases (FFT  
397 with SNP\_7, and MFT with SNP\_28), the SNP effect did not vary among populations  
398 (non-significant SNP-by-population interaction in model M5' when we included the  
399 SNP interactions with year\*field and population). For a subset of two SNPs, we  
400 illustrated the regression between the trait value and the shift of allele frequencies  
401 with altitude (Fig 6 A&B). We estimated corresponding additive and dominance  
402 effects (S7 Table). In some cases, the intra-population effect corroborated the inter-  
403 population variation with relatively large additive effects of the same sign (Fig 6  
404 C&D). Note that in both examples shown in Fig 6, one or the other allele was  
405 dominant. In other cases, the results were more difficult to interpret with negligible  
406 additive effect but extremely strong dominance (S7 Table, SNP\_210 for instance).

407

408

409

410 **Figure 6: Regression of phenotypic average value on SNP allele frequency across**  
411 **populations, and within-population average phenotypic value for each SNP**  
412 **genotype.** Per-population phenotypic average values of traits are regressed on alleles  
413 frequencies at SNP\_149 (A) and SNP\_179 (B) with corresponding within-population  
414 average phenotypic value per genotype (C & D). In A and B, the 11 populations of  
415 the association panel are shown with *parviglumis* (green circles) and *mexicana* (red  
416 triangles) populations sampled along gradient *a* and gradient *b*. Phenotypic average  
417 values were corrected for the experimental design (calculated as the residues of model  
418 M2). Pval refers to the P-value of the linear regression represented in blue. In C and  
419 D, genotypic effects from model M5' are expressed as the average phenotypic value  
420 of heterozygotes (1) and homozygotes for the reference (0) and the alternative allele  
421 (2). FDR values are obtained from the association analysis on 171 SNPs with  
422 correction for genetic structure using 11 population.

423

#### 424 **Independence of SNPs associated to phenotypes.**

425 We computed the pairwise linkage disequilibrium (LD) as measured by  $r^2$   
426 between the 171 outlier SNPs using the R package LDcorSV [62]. Because we were  
427 specifically interested by LD pattern between phenotypically-associated SNPs, as for  
428 the association analyses we accounted for structure and kinship computed from SSRs  
429 while estimating LD [63]. The 171 outlier SNPs were distributed along the 10  
430 chromosomes of maize, and exhibited low level of linkage disequilibrium (LD),  
431 except for SNPs located on chromosomes eight, nine, and a cluster of SNPs located  
432 on chromosome 4 (S12 Fig).

433 Among the 171, the subset of 23 phenotypically-associated SNPs (detected  
434 when considering the 11-population structure) displayed an excess of elevated LD

435 values – out of 47 pairs of SNPs phenotypically-associated to a same trait, 16 pairs  
436 were contained in the 5% higher values of the LD distribution of all outlier SNP pairs.  
437 Twelve out of the 16 pairs related to grain weight, the remaining four to leaf  
438 coloration, and one pair of SNPs was associated to both traits. Noteworthy was that  
439 inversions on chromosomes 1, 4, and 9, taken together, were enriched for  
440 phenotypically-associated SNPs ( $\chi^2 = 8.95$ , P-value=0.0028). We recovered a  
441 borderline significant enrichment with the correction  $K=5$  ( $\chi^2 = 3.82$ , P-value=0.051).

442 Finally, we asked whether multiple SNPs contributed independently to the  
443 phenotypic variation of a single trait. We tested a multiple SNP model where SNPs  
444 were added incrementally when significantly associated (FDR < 0.10). We found 2, 3  
445 and 2 SNPs for female, male flowering time and height of the highest ear,  
446 respectively (S5 Table). Except for the latter trait, the SNPs were located on different  
447 chromosomes.

448

## 449 **Discussion**

450

451 Plants are excellent systems to study local adaptation. First, owing to their  
452 sessile nature, local adaptation of plant populations is pervasive [13]. Second,  
453 environmental effects can be efficiently controlled in common garden experiments,  
454 facilitating the identification of the physiological, morphological and phenological  
455 traits influenced by spatially-variable selection [64]. Identification of the determinants  
456 of complex trait variation and their covariation in natural populations is however  
457 challenging [65]. While population genomics has brought a flurry of tools to detect  
458 footprints of local adaptation, their reliability remains questioned [61]. In addition,  
459 local adaptation and demographic history frequently follow the same geographic route,

460 making the disentangling of trait, molecular, and environmental variation, particularly  
461 arduous. Here we investigated those links on a well-established outcrossing system,  
462 the closest wild relatives of maize, along altitudinal gradients that display  
463 considerable environmental shifts over short geographical scales.

464

465 **The syndrome of altitudinal adaptation results from selection at multiple co-**  
466 **adapted traits.**

467 Common garden studies along elevation gradients have been conducted in  
468 European and North American plants species [66]. Together with other studies, they  
469 have revealed that adaptive responses to altitude are multifarious [67]. They include  
470 physiological responses such as high photosynthetic rates [68], tolerance to frost [69],  
471 biosynthesis of UV-induced phenolic components [70]; morphological responses with  
472 reduced stature [71, 72], modification of leaf surface [73], increase in leaf non-  
473 glandular trichomes [74], modification of stomata density; and phenological  
474 responses with variation in flowering time [75], and reduced growth period [76].

475 Our multivariate analysis of teosinte phenotypic variation revealed a marked  
476 differentiation between teosinte subspecies along an axis of variation (21.26% of the  
477 total variation) that also discriminated populations by altitude (Fig 2A & B). The  
478 combined effects of assortative mating and environmental elevation variation may  
479 generate, in certain conditions, trait differentiation along gradients without underlying  
480 divergent selection [77]. While we did not measure flowering time differences among  
481 populations *in situ*, we did find evidence for long distance gene flow between  
482 gradients and subspecies (Fig 4 A & C). In addition, several lines of arguments  
483 suggest that the observed clinal patterns result from selection at independent traits and  
484 is not solely driven by differences in flowering time among populations. First, two

485 distinct methods accounting for shared population history concur with signals of  
486 spatially-varying selection at ten out of the 18 traits (Table 1). Nine of them exhibited  
487 a clinal trend of increase/decrease of population phenotypic values with elevation (S3  
488 Fig) within at least one of the two subspecies. This number is actually conservative,  
489 because these approaches disregard the impact of selective constraints which in fact  
490 tend to decrease inter-population differences in phenotypes. Second, while male and  
491 female flowering times were positively correlated, they displayed only subtle  
492 correlations ( $|r| < 0.16$ ) with other spatially-varying traits except for grain weight and  
493 length ( $|r| < 0.33$ ). Third, we observed convergence at multiple phenotypes between  
494 the lowland populations from the two gradients that occurred despite their  
495 geographical and genetical distance (Fig 4) again arguing that local adaptation drives  
496 the underlying patterns.

497       Spatially-varying traits that displayed altitudinal trends, collectively defined a  
498 teosinte altitudinal syndrome of adaptation characterized by early-flowering,  
499 production of few tillers albeit numerous lateral branches, production of heavy, long  
500 and large grains, and decrease in stomata density. We also observed increased leaf  
501 pigmentation with elevation, although with a less significant signal (S3 Table),  
502 consistent with the pronounced difference in sheath color reported between  
503 *parviglumis* and *mexicana* [78, 79]. Because seeds were collected from wild  
504 populations, a potential limitation of our experimental setting is the confusion  
505 between genetic and environmental maternal effects. Environmental maternal effects  
506 could bias upward our heritability estimates. However, our results corroborate  
507 previous findings of reduced number of tillers and increased grain weight in *mexicana*  
508 compared with *parviglumis* [80]. Thus although maternal effects could not be fully  
509 discarded, we believe they were likely to be weak.

510           The trend towards depleted stomata density at high altitudes (S3 Fig) could  
511   arguably represent a physiological adaptation as stomata influence components of  
512   plant fitness through their control of transpiration and photosynthetic rate [81]. Indeed,  
513   in natural accessions of *A. thaliana*, stomatal traits showed signatures of local  
514   adaptation and were associated with both climatic conditions and water-use efficiency  
515   [82]. Furthermore, previous work has shown that in arid and hot highland  
516   environments, densely-packed stomata may promote increased leaf cooling in  
517   response to desiccation [83] and may also counteract limited photosynthetic rate with  
518   decreasing pCO<sub>2</sub> [84]. Accordingly, increased stomata density at high elevation sites  
519   has been reported in alpine species such as the European beech [85] as well as in  
520   populations of *Mimulus guttatus* subjected to higher precipitations in the Sierra  
521   Nevada [86]. In our case, higher elevations display both *arid* environment and *cooler*  
522   temperatures during the growing season, features perhaps more comparable to other  
523   tropical mountains for which a diversity of patterns in stomatal density variation with  
524   altitude has been reported [87]. Further work will be needed to decipher the  
525   mechanisms driving the pattern of declining stomata density with altitude in teosintes.  
526   Altogether, the altitudinal syndrome was consistent with natural selection for rapid  
527   life-cycle shift, with early-flowering in the shorter growing season of the highlands  
528   and production of larger propagules than in the lowlands. This altitudinal syndrome  
529   evolved in spite of detectable gene flow.

530           Although we did not formally measure biomass production, the lower number  
531   of tillers and higher amount and size of grains in the highlands when compared with  
532   the lowlands may reflect trade-offs between allocation to grain production and  
533   vegetative growth [88]. Because grains fell at maturity and a single teosinte individual  
534   produces hundreds of ears, we were unable to provide a proxy for total grain



535 production. The existence of fitness-related trade-offs therefore still needs to be  
536 formally addressed.

537         Beyond trade-offs, our results more generally question the extent of  
538 correlations between traits. In maize, for instance, we know that female and male  
539 flowering time are positively correlated and that their genetic control is in part  
540 determined by a common set of genes [89]. They themselves further increase with  
541 yield-related traits [90]. Response to selection for late-flowering also led to a  
542 correlated increase in leaf number in cultivated maize [91], and common genetic loci  
543 have been shown to determine these traits as well [92]. Here we found strong positive  
544 correlations between traits: male and female flowering time, grain length and width,  
545 plant height and height of the lowest or highest ear. Strong negative correlations were  
546 observed instead between grain weight and both male and female flowering time.  
547 Trait correlations were therefore partly consistent with previous observations in maize,  
548 suggesting that they were inherited from wild ancestors.

549

550 **Footprints of past adaptation are relevant to detect variants involved in present**  
551 **phenotypic variation.**

552         The overall level of differentiation in our outcrossing system ( $F_{ST} \approx 22\%$ ) fell  
553 within the range of previous estimates (23% [93] and 33% [55] for samples  
554 encompassing both teosinte subspecies). It is relatively low compared to other  
555 systems such as the selfer *Arabidopsis thaliana*, where association panels typically  
556 display maximum values of  $F_{ST}$  around 60% within 10kb-windows genome-wide [94].  
557 Nevertheless, correction for sample structure is key for statistical associations  
558 between genotypes and phenotypes along environmental gradients. This is because  
559 outliers that display lowland/highland differentiation co-vary with environmental

560 factors, which themselves may affect traits [95]. Consistently, we found that 73.7%  
561 SNPs associated with phenotypic variation at  $K=5$ , but only 13.5% of them did so  
562 when considering a genetic structure with 11 populations. Except for one, the latter  
563 set of SNPs represented a subset of the former. Because teosinte subspecies  
564 differentiation was fully accounted for at  $K=5$  (as shown by the clear distinction  
565 between *mexicana* populations and the rest of the samples, Fig 4A), the inflation of  
566 significant associations at  $K=5$  is not due to subspecies differentiation, but rather to  
567 residual stratification among populations within genetic groups. Likewise, recent  
568 studies in humans, where global differentiation is comparatively low [96] have shown  
569 that incomplete control for population structure within European samples strongly  
570 impacts association results [61, 97]. Controlling for such structure may be even more  
571 critical in domesticated plants, where genetic structure is inferred *a posteriori* from  
572 genetic data (rather than *a priori* from population information) and pedigrees are  
573 often not well described. Below, we show that considering more than one correction  
574 using minor peaks delivered by the Evanno statistic (S5 Fig) can be informative.

575         Considering a structure with 5 genetic groups, the number of SNPs associated  
576 per trait varied from 1 to 55, with no association for leaf and grain coloration (S5  
577 Table). False positives likely represent a greater proportion of associations at  $K=5$  as  
578 illustrated by a slight excess of small P-values when compared with a correction with  
579 11 populations for most traits (S10 Fig). Nevertheless, our analysis recovered credible  
580 candidate adaptive loci that were no longer associated when a finer-grained  
581 population structure was included in the model. For instance at  $K=5$ , we detected  
582 *Sugary1* (*Su1*), a gene encoding a starch debranching enzyme that was selected during  
583 maize domestication and subsequent breeding [98, 99]. We found that *Su1* was  
584 associated with variation at six traits (male and female flowering time, tassel

585 branching, height of the highest ear, grain weight and stomata density) pointing to  
586 high pleiotropy. A previous study reported association of this gene to oil content in  
587 teosintes [100]. In maize, this gene has a demonstrated role in kernel phenotypic  
588 differences between maize genetic groups [101]. *Su1* is therefore most probably a  
589 true-positive. That this gene was no longer recovered with the 11-population structure  
590 correction indicated that divergent selection acted among populations. Indeed, allelic  
591 frequency was highly contrasted among populations, with most populations fixed for  
592 one or the other allele, and a single population with intermediate allelic frequency.  
593 With the 11-population correction, very low power is thus left to detect the effect of  
594 *Su1* on phenotypes.

595         Although the confounding population structure likely influenced the genetic  
596 associations, experimental evidence indicates that an appreciable proportion of the  
597 variants recovered with both  $K=5$  and 11 populations are true-positives (S5 Table).  
598 One SNP associated with female and male flowering time, as well as with plant height  
599 and grain length (at  $K=5$  only for the two latter traits) maps within the *phytochrome*  
600 *B2* (SNP\_210; *phyB2*) gene. Phytochromes are involved in perceiving light signals  
601 and are essential for growth and development in plants. The maize gene *phyB2*  
602 regulates the photoperiod-dependent floral transition, with mutants producing early  
603 flowering phenotypes and reduced plant height [102]. Genes from the  
604 phosphatidylethanolamine-binding proteins (PEBPs) family – *Zea mays*  
605 *CENTRORADIALIS* (*ZCN*) family in maize – are also well-known to act as promotor  
606 and repressor of the floral transition in plants [103]. *ZCN8* is the main floral activator  
607 of maize [104], and both *ZCN8* and *ZCN5* strongly associate with flowering time  
608 variation [101, 105]. Consistently, we found associations of male and female  
609 flowering time with *PEBP18* (SNP\_15). It is interesting to note that SNPs at two

610 flowering time genes, *phyB2* and *PEBP18*, influenced independently as well as in  
611 combination both female and male flowering time variation (S5 Table).

612           The proportion of genic SNPs associated to phenotypic variation was not  
613 significantly higher than that of non-genic SNPs (i.e, SNPs >1kb from a gene) ( $\chi^2_{(df=1)}$   
614 = 0.043, P-value = 0.84 at  $K=5$  and  $\chi^2_{(df=1)}=1.623$ , P-value =0.020 with 11  
615 populations) stressing the importance of considering both types of variants [106]. For  
616 instance, we discovered a non-genic SNP (SNP\_149) that displayed a strong  
617 association with leaf width variation as well as a pattern of allele frequency shift with  
618 altitude among populations (Fig 6B).

619

620 **Physically-linked and independent SNPs both contribute to the establishment of**  
621 **adaptive genetic correlations.**

622           We found limited LD among our outlier SNPs (S10 Fig) corroborating  
623 previous reports (LD decay within <100bp, [58, 93]). However, the subset of  
624 phenotypically-associated SNPs displayed greater LD, a pattern likely exacerbated by  
625 three Mb-scale inversions located on chromosomes 1 (*Inv1n*), 4 (*Inv4m*) and 9 (*Inv9e*)  
626 that, taken together, were enriched for SNPs associated with environmental variables  
627 related to altitude and/or SNPs associated with phenotypic variation. Previous work  
628 [55, 56] has shown that *Inv1n* and *Inv4m* segregate within both *parviglumis* and  
629 *mexicana*, while two inversions on chromosome 9, *Inv9d* and *Inv9e*, are present only  
630 in some of the highest *mexicana* populations; such that all four inversions also follow  
631 an altitudinal pattern. Our findings confirmed that three of these inversions possessed  
632 an excess of SNPs with high  $F_{ST}$  between subspecies and between low- and high-  
633 *mexicana* populations for *Inv9e* [57]. Noteworthy *Inv9d* contains a large ear leaf  
634 width quantitative trait locus in maize [106]. Corroborating these results, we found

635 consistent association between the only SNP located within this inversion and leaf  
636 width variation in teosinte populations (S5 Table). Overall, our results further  
637 strengthen the role of chromosomal inversions in teosinte altitudinal adaptation.

638         Because inversions suppress recombination between inverted and non-inverted  
639 genotypes, their spread has likely contributed to the emergence and maintenance of  
640 locally adaptive allelic combinations in the face of gene flow, as reported in a  
641 growing number of other models (reviewed in [107]) including insects [108], fish  
642 [109], birds [110] and plants [26, 111]. But we also found three cases of multi-SNP  
643 determinism of traits (male and female flowering time and height of the highest ear,  
644 Table S5) supporting selection of genetically independent loci. Consistently with  
645 Weber et al. [100], we found that individual SNPs account for small proportions of  
646 the phenotypic variance (S7 Table). Altogether, these observations are consistent with  
647 joint selection of complex traits determined by several alleles of small effects, some  
648 of which being maintained in linkage through selection of chromosomal  
649 rearrangements.

650

## 651 **Conclusion.**

652

653         Elevation gradients provide an exceptional opportunity for investigating  
654 variation of functional traits in response to continuous environmental factors at short  
655 geographical scales. Here we documented patterns indicating that local adaptation,  
656 likely facilitated by the existence of chromosomal inversions, allows teosintes to cope  
657 with specific environmental conditions in spite of gene flow. We detected an  
658 altitudinal syndrome in teosintes composed of sets of independent traits evolving  
659 under spatially-varying selection. Because traits co-varied with environmental

660 differences along gradients, however, statistical associations between genotypes and  
661 phenotypes largely depended on control of population stratification. Yet, several of  
662 the variants we uncovered seem to underlie adaptive trait variation in teosintes.  
663 Adaptive teosinte trait variation is likely relevant for maize evolution and breeding.  
664 Whether the underlying SNPs detected in teosintes bear similar effects in maize or  
665 whether their effects differ in domesticated backgrounds will have to be further  
666 investigated.

## 667 **Material and Methods**

668

### 669 **Description of teosinte populations and sampling.**

670 We used 37 teosinte populations of *mexicana* (16) and *parviglumis* (21)  
671 subspecies from two previous collections [57, 58, 112] to design our sampling. These  
672 populations (S1 Table) are distributed along two altitudinal gradients (Fig 1). We  
673 plotted their altitudinal profiles using R ‘raster’ package [113] (S1 Fig). We further  
674 obtained 19 environmental variable layers from  
675 <http://idrisi.uaemex.mx/distribucion/superficies-climaticas-para-mexico>. These high-  
676 resolution layers comprised monthly values from 1910 to 2009 estimated via  
677 interpolation methods [107]. We extracted values of the 19 climatic variables for each  
678 population (S1 Table). Note that high throughput sequencing (HTS) data were  
679 obtained in a previous study for six populations out of the 37 (M6a, P1a, M7b, P2b,  
680 M1b and P8b; Fig 1, S1 Table) to detect candidate genomic regions for local  
681 adaptation [58]. The four highest and lowest of these populations were included in the  
682 association panel described below.

683 We defined an association panel of 11 populations on which to perform a  
684 genotype-phenotype association study (S1 Table). Our choice was guided by grain  
685 availability as well as the coverage of the whole climatic and altitudinal ranges.  
686 Hence, we computed Principal Component Analyses (PCA) for each gradient from  
687 environmental variables using the FactoMineR package in R [114] and added altitude  
688 to the PCA graphs as a supplementary variable. Our association panel comprised five  
689 populations from a first gradient (*a*) – two *mexicana* and three *parviglumis*, and six  
690 populations from a second gradient (*b*) – one *mexicana* and five *parviglumis* (Fig 1).

691 Finally, we extracted available SNP genotypes generated with the  
692 MaizeSNP50 Genotyping BeadChip for 28 populations out of our 37 populations [57]  
693 (S1 Table). From this available SNP dataset, we randomly sampled 1000 SNPs found  
694 to display no selection footprint [57], hereafter neutral SNPs. Data for neutral SNPs  
695 (Data S1) are available at: [10.6084/m9.figshare.9901472](https://figshare.com/figures/data/10.6084/m9.figshare.9901472). We used this panel of 28  
696 populations to investigate correlation with environmental variation. Note that 10 out  
697 of the 28 populations were common to our association panel, and genotypes were  
698 available for 24 to 34 individuals per population, albeit different from the ones of our  
699 association mapping panel.

#### 700 **Common garden experiments**

701 We used two common gardens for phenotypic evaluation of the association  
702 panel (11 populations). Common gardens were located at INIFAP (Instituto Nacional  
703 de Investigaciones Forestales, Agrícolas y Pecuaria) experimental field stations in the  
704 state of Guanajuato in Mexico, one in Celaya municipality at the Campo  
705 Experimental Bajío (CEBAJ) (20°31'20'' N, 100°48'44'' W) at 1750 meters of  
706 elevation, and one in San Luis de la Paz municipality at the Sitio Experimental Norte  
707 de Guanajuato (SENGUA) (21°17'55'' N, 100°30'59'' W) at 2017 meters of elevation.  
708 These locations were selected because they present intermediate altitudes (S1 Fig).  
709 The two common gardens were replicated in 2013 and 2014.

710 The original sampling contained 15 to 22 mother plants per population. Eight  
711 to 12 grains per mother plant were sown each year in individual pots. After one  
712 month, seedlings were transplanted in the field. Each of the four fields (2 locations, 2  
713 years) was separated into four blocks encompassing 10 rows and 20 columns. We  
714 evaluated one offspring of ~15 mother plants from each of the 11 teosinte populations  
715 in each block, using a semi-randomized design, i.e. each row containing one or two



716 individuals from each population, and individuals being randomized within row,  
717 leading to a total of 2,640 individual teosinte plants evaluated.

718

### 719 **SSR genotyping and genetic structuring analyses on the association panel**

720 In order to quantify the population structure and individual kinship in our  
721 association panel, we genotyped 46 SSRs (S4 Table). Primers sequences are available  
722 from the maize database project [115] and genotyping protocol were previously  
723 published [116]. Genotyping was done at the GENTYANE platform (UMR INRA  
724 1095, Clermont-Ferrand, France). Allele calling was performed on electropherograms  
725 with the GeneMapper® Software Applied Biosystems®. Allele binning was carried  
726 out using Autobin software [117], and further checked manually.

727 We employed STRUCTURE Bayesian classification software to compute a  
728 genetic structure matrix on individual genotypes. Individuals with over 40% missing  
729 data were excluded from analysis. For each number of clusters ( $K$  from 2 to 13), we  
730 performed 10 independent runs of 500,000 iterations after a burn-in period of 50,000  
731 iterations, and combined these 10 replicates using the LargeKGreedy algorithm from  
732 the CLUMPP program [118]. We plotted the resulting clusters using DISTRUCT  
733 software. We then used the Evanno method [119] to choose the optimal  $K$  value. We  
734 followed the same methodology to compute a structure matrix from the outlier SNPs.

735 We inferred a kinship matrix  $\mathbf{K}$  from the same SSRs using SPAGeDI [120].  
736 Kinship coefficients were calculated for each pair of individuals as correlation  
737 between allelic states [121]. Since teosintes are outcrossers and expected to exhibit an  
738 elevated level of heterozygosity, we estimated intra-individual kinship to fill in the  
739 diagonal. We calculated ten kinship matrices, each excluding the SSRs from one out

740 of the 10 chromosomes. Microsatellite data (Data S2) are available  
741 at: [10.6084/m9.figshare.9901472](https://doi.org/10.6084/m9.figshare.9901472)

742 In order to gain insights into population history of divergence and admixture,  
743 we used 1000 neutral SNPs (i.e. SNPs genotyped by Aguirre-Liguori and  
744 collaborators [57] and that displayed patterns consistent with neutrality among 49  
745 teosinte populations) genotyped on 10 out of the 11 populations of the association  
746 panel to run a TreeMix analysis (TreeMix version 1.13 [122]. TreeMix models  
747 genetic drift to infer populations splits from an outgroup as well as migration edges  
748 along a bifurcating tree. We oriented the SNPs using the previously published  
749 MaizeSNP50 Genotyping BeadChip data from the outgroup species *Tripsacum*  
750 *dactyloides* [55]. We tested from 0 to 10 migration edges. We fitted both a simple  
751 exponential and a non-linear least square model (threshold of 1%) to select the  
752 optimal number of migration edges as implemented in the OptM R package [123]. We  
753 further verified that the proportion of variance did not substantially increase beyond  
754 the optimal selected value.

755

## 756 **Phenotypic trait measurements**

757 We evaluated a total of 18 phenotypic traits on the association panel (S2  
758 Table). We measured six traits related to plant architecture (PL: Plant Height, HLE:  
759 Height of the Lowest Ear, HHE: Height of the Highest Ear, Til: number of Tillers,  
760 LBr: number of Lateral Branches, NoE: number of Nodes with Ears), three traits  
761 related to leaf morphologies (LeL: Leaf Length, LeW: Leaf Width, LeC: Leaf  
762 Color), three traits related to reproduction (MFT: Male Flowering Time, FFT: Female  
763 Flowering Time, TBr : Tassel Branching), five traits related to grains (Gr: number of  
764 Grains per ear, GrL: Grain Length, GrWi: Grain Width, GrWe: Grain Weight, GrC:

765 Grain Color), and one trait related to Stomata (StD: Stomata Density). These traits  
766 were chosen because we suspected they could contribute to differences among  
767 teosinte populations based on a previous report of morphological characterization on  
768 112 teosinte collections grown in five localities [124].

769 We measured the traits related to plant architecture and leaves after silk  
770 emergence. Grain traits were measured at maturity. Leaf and grain coloration were  
771 evaluated on a qualitative scale. For stomata density, we sampled three leaves per  
772 plant and conserved them in humid paper in plastic bags. Analyses were undertaken at  
773 the Institute for Evolution and Biodiversity (University of Münster) as followed: 5mm  
774 blade discs were cut out from the mid length of one of the leaves and microscopic  
775 images were taken after excitation with a 488nm laser. Nine locations ( $0.15\text{mm}^2$ ) per  
776 disc were captured with 10 images per location along the z-axis (vertically along the  
777 tissue). We automatically filtered images based on quality and estimated leaf stomata  
778 density using custom image analysis algorithms implemented in Matlab. For each  
779 sample, we calculated the median stomata density over the (up to) nine locations. To  
780 verify detection accuracy, manual counts were undertaken for 54 random samples.  
781 Automatic and manual counts were highly correlated ( $R^2=0.82$ ), indicating reliable  
782 detection (see S1 Annex StomataDetection, Dittberner and de Meaux, for a detailed  
783 description). The filtered data set of phenotypic measurements (Data S3) is available  
784 at: [10.6084/m9.figshare.9901472](https://doi.org/10.6084/m9.figshare.9901472).

785

### 786 **Statistical analyses of phenotypic variation**

787 In order to test for genetic effects on teosinte phenotypic variation, we  
788 decomposed phenotypic values of each trait considering a fixed population effect plus  
789 a random mother-plant effect (model M1):

$$Y_{ijklm} = \mu + \alpha_i + \beta_j + \theta_{ij} + \gamma_{k/ij} + \delta_l + \chi_{il} + \psi_{jl} + P_{m/l} + \varepsilon_{ijklm} \quad (M1)$$

790 where the response variable  $Y$  is the observed phenotypic value,  $\mu$  is the total mean,  $\alpha_i$   
791 is the fixed year effect ( $i = 2013, 2014$ ),  $\beta_j$  the fixed field effect ( $j =$  field station,  
792 *SENGUA, CEBAJ*),  $\theta_{ij}$  is the year by field interaction,  $\gamma_{k/ij}$  is the fixed block effect  
793 ( $k = 1, 2, 3, 4$ ) nested within the year-by-field combination,  $\delta_l$  is the fixed effect of the  
794 population of origin ( $l = 1$  to 11),  $\chi_{il}$  is the year by population interaction,  $\psi_{jl}$  is the  
795 field by population interaction,  $P_{m/l}$  is the random effect of mother plant ( $m = 1$  to 15)  
796 nested within population, and  $\varepsilon_{ijklm}$  is the individual residue. Identical notations were  
797 used in all following models. For the distribution of the effects, the same variance was  
798 estimated within all populations. Mixed models were run using ASReml v.3.0 [125]  
799 and MM4LMM v2.0.1 [[https://rdr.io/cran/MM4LMM/man/MM4LMM-](https://rdr.io/cran/MM4LMM/man/MM4LMM-package.html)  
800 [package.html](https://rdr.io/cran/MM4LMM/man/MM4LMM-package.html), update by F. Laporte] R packages, which both gave very similar results,  
801 and fixed effects were tested through Wald tests.

802 For each trait, we represented variation among populations using box-plots on  
803 mean values per mother plant adjusted for the experimental design following model  
804 M'1:

$$Y_{ijklm} = \mu + \alpha_i + \beta_j + \theta_{ij} + \gamma_{k/ij} + p_{m/l} + \varepsilon_{ijklm} \quad (M'1)$$

805 where mother plant within population is considered as fixed. We used the function  
806 *predict* to obtain least-square means (ls-means) of each mother plant, and looked at  
807 the tendencies between population's values. All fixed models were computed using  
808 *lm* package in R, and we visually checked the assumptions of residues independence  
809 and normal distribution.

810 We performed a principal component analysis (PCA) on phenotypic values  
811 corrected for the experimental design, using FactoMineR package in R [114] from the  
812 residues of model M2 computed using the *lm* package in R:

$$Y_{ijklm} = \mu + \alpha_i + \beta_j + \theta_{ij} + \gamma_{k/ij} + \varepsilon_{ijklm} \quad (\text{M2})$$

813 Finally, we tested for altitudinal effects on traits by considering the altitude of  
814 the sampled population ( $l$ ) as a covariate (ALT) and its interaction with year and field  
815 in model M3:

$$Y_{ijklm} = \mu + \alpha_i + \beta_j + \theta_{ij} + \gamma_{k/ij} + c \cdot ALT_l + a_i \cdot ALT_l + b_j \cdot ALT_l + P_{m/l} + \varepsilon_{ijklm}$$

(M3)

816 where all terms are equal to those in model M1 except that the fixed effect of the  
817 population of origin was replaced by a regression on the population altitude ( $ALT_l$ ).

818

### 819 **Detection of selection acting on phenotypic traits**

820

821 We aimed at detecting traits evolving under spatially varying selection by  
822 comparing phenotypic to neutral genotypic differentiation.  $Q_{st}$  is a statistic analogous  
823 to  $F_{ST}$  but for quantitative traits, which can be described as the proportion of  
824 phenotypic variation explained by differences among populations [19, 107].  
825 Significant differences between  $Q_{ST}$  and  $F_{ST}$  can be interpreted as evidence for  
826 spatially-varying ( $Q_{ST} > F_{ST}$ ) selection [126]. We used the R package *QstFstComp*  
827 [127] that is adequate for experimental designs with randomized half-sibs in  
828 outcrossing species. We used individuals that were both genotyped and phenotyped  
829 on the association panel to establish the distribution of the difference between  
830 statistics ( $Q_{ST} - F_{ST}$ ) under the neutral hypothesis of evolution by drift - using the half-  
831 sib dam breeding design and 1000 resamples. We next compared it to the observed  
832 difference with 95% threshold cutoff value in order to detect traits under spatially-  
833 varying selection.

834 In addition to  $Q_{ST}$ - $F_{ST}$  analyses, we employed the DRIFTSEL R package [128]  
835 to test for signal of selection of traits while accounting for drift-driven population  
836 divergence and genetic relatedness among individuals (half-sib design). DRIFTSEL is a  
837 Bayesian method that compares the probability distribution of predicted and observed  
838 mean additive genetic values. It provides the S statistic as output, which measures the  
839 posterior probability that the observed population divergence arose under divergent  
840 selection ( $S \sim 1$ ), stabilizing selection ( $S \sim 0$ ) or genetic drift (intermediate S values)  
841 [59]. It is particularly powerful for small datasets, and can distinguish between drift  
842 and selection even when  $Q_{ST}$ - $F_{ST}$  are equal [59]. We first applied RAFM to estimate the  
843  $F_{ST}$  value across populations, and the population-by-population coancestry coefficient  
844 matrix. We next fitted both the RAFM and DRIFTSEL models with 15,000 MCMC  
845 iterations, discarded the first 5,000 iterations as a transient, and thinned the remaining  
846 by 10 to provide 1000 samples from the posterior distribution. Note that DRIFTSEL was  
847 slightly modified because we had information only about the dams, but not the sires,  
848 of the phenotyped individuals. We thus modified DRIFTSEL with the conservative  
849 assumption of all sires being unrelated. Because DRIFTSEL does not require that the  
850 same individuals were both genotyped and phenotyped, we used SSRs and phenotype  
851 data of the association panel as well as the set of neutral SNPs and phenotype data on  
852 10 out of the 11 populations. For the SNP analyses, we selected out of the 1000  
853 neutral SNPs the 465 most informative SNPs based on the following criteria:  
854 frequency of the less common variant at least 10%, and proportion of missing data at  
855 most 1%. Finally, we estimated from DRIFTSEL the posterior probability of the  
856 ancestral population mean for each trait as well as deviations of each population from  
857 these values.

858 Both  $Q_{ST}$ - $F_{ST}$  and DRIFTSEL rely on the assumption that the observed  
859 phenotypic variation was determined by additive genotypic variation. We thus  
860 estimated narrow-sense heritability for each trait in each population to estimate the  
861 proportion of additive variance in performance. We calculated per population narrow-  
862 sense heritabilities as the ratio of the estimated additive genetic variance over the total  
863 phenotypic variance on our common garden measurements using the MCMCglmm R  
864 package [129] where half sib family is the single random factor, and the design (block  
865 nested within year and field) is corrected as fixed factor. For three grain-related traits,  
866 we also ran the same model but including mother plants phenotypic values calculated  
867 from the remaining grains not sown. We ran 100,000 iterations with 10,000 burn-in,  
868 inverse gamma (0.001; 0.001) as priors. We then calculated the mean and standard  
869 deviation of the 11 per population  $h^2$  estimates.

870

#### 871 **Pairwise correlations between traits.**

872 We evaluated pairwise-correlations between traits by correlating the residues  
873 obtained from model M4, that corrects the experiment design (year, field and blocks)  
874 as well as the underlying genetic structure estimated from SSRs:

$$Y_{ijklm} = \mu + \alpha_i + \beta_j + \theta_j + \gamma_{k/ij} + \sum_{n=1}^4 b_n \cdot C_{ijklm}^n + \varepsilon_{ijklm} \quad (\text{M4})$$

875 where  $b_n$  is the slope of the regression of Y on the  $n^{\text{th}}$  structure covariate  $C^n$ . Structure  
876 covariate values ( $C^n$  covariates, from STRUCTURE output) were calculated at the  
877 individual level, i.e. for each offspring of mother plant  $m$  from population  $l$ , grown in  
878 the year  $i$  field  $j$  and block  $k$ .  $C^n$  are thus declared with  $ijklm$  indices, although they are  
879 purely genetic covariates.

880

## 881 **Genotyping of outlier SNPs on 28 populations**

882 We extracted total DNA from each individual plant of the association panel as  
883 well as 20 individuals from each of the 18 remaining populations that were not  
884 included in the association panel (Table 1). Extractions were performed from 30 mg  
885 of lyophilized adult leaf material following recommendations of DNeasy 96 Plant Kit  
886 manufacturer (QIAGEN, Valencia, CA, USA). We genotyped outlier SNPs using  
887 Kompetitive Allele Specific PCR technology (KASPar, LGC Group) [130]. Data for  
888 outlier SNPs (Data S4 and Data S5) are available at: [10.6084/m9.figshare.9901472](https://doi.org/10.6084/m9.figshare.9901472).

889 Among SNPs identified as potentially involved in local adaptation, 270 were  
890 designed for KASPar assays, among which 218 delivered accurate quality data. Of the  
891 218 SNPs, 141 were detected as outliers in two previous studies using a combination  
892 of statistical methods – including  $F_{ST}$ -scans [131], Bayescan [32] and Bayenv2 [35,  
893 132], Bayescenv [133] – applied to either six of our teosinte populations [58] or to a  
894 broader set of 49 populations genotyped by the Illumina® MaizeSNP50 BeadChip  
895 [57]. The remaining outlier SNPs (77) were detected by  $F_{ST}$ -scans from six  
896 populations (S7 Fig, S5 Table), following a simplified version of the rationale in [58]  
897 by considering only differentiation statistics: SNPs were selected if they displayed  
898 both a high differentiation (5% highest  $F_{ST}$  values) between highland and lowland  
899 populations in at least one of the two gradients, and a high differentiation (5% highest  
900  $F_{ST}$  values) between highland and lowland populations either within *parviglumis* (P2b  
901 and P8b) or within *mexicana* (M7b and M1b) or both in gradient *b* (S1 Fig). We  
902 thereby avoided SNPs fixed between the two subspecies.

903



904 **Association mapping**

905 We tested the association of phenotypic measurements with outlier SNPs on a  
906 subset of individuals for which (1) phenotypic measurements were available, (2) at  
907 least 60% of outlier SNPs were adequately genotyped, and (3) kinship and cluster  
908 membership values were available from SSR genotyping. For association, we  
909 removed SNPs with minor allele frequency lower than 5%.

910 In order to detect statistical associations between outlier SNPs and phenotypic  
911 variation, we used the following mixed model derived from [128]:

$$Y_{ijklm} = \mu + \alpha_i + \beta_j + \theta_{ij} + \gamma_{k/ij} + \sum_{n=1}^4 b_n \cdot C_{ijklm}^n + \zeta_o + u_{ijklm} + \varepsilon_{ijklm} \quad (\text{M5})$$

912 where  $\zeta$  is the fixed bi-allelic SNP factor with one level for each of the three  
913 genotypes ( $o=0, 1, 2$ ; with  $o=1$  for heterozygous individuals), and  $u_{ijklm}$  is the random  
914 genetic effect of the individual. We assumed that the vector of  $u_{ijklm}$  effects followed a  
915  $N(0, \mathbf{K} \sigma^2 u)$  distribution, where  $\mathbf{K}$  is the kinship matrix computed as described above.

916 A variant of model M5 was employed to test for SNP association to traits,  
917 while correcting for structure as the effect of population membership ( $\delta_i$ ),  $\delta$  being a  
918 factor with 11 levels (populations):

$$Y_{ijklm} = \mu + \alpha_i + \beta_j + \theta_{ij} + \gamma_{k/ij} + \delta_i + \zeta_o + u_{ijklm} + \varepsilon_{ijklm} \quad (\text{M5}')$$

919 In order to avoid overcorrection of neutral genetic structure and improve  
920 power, we ran the two models independently for each chromosome using a kinship  
921 matrix  $\mathbf{K}$  estimated from all SSRs except those contained in the chromosome of the  
922 tested SNP [134]. We tested SNP effects through the Wald statistics, and applied a 10%  
923 False Discovery Rate (FDR) threshold for each phenotype separately. In order to  
924 validate the correction for genetic structure, the 38 multiallelic SSR genotypes were  
925 transformed into biallelic genotypes, filtered for  $\text{MAF} > 5\%$ , and used to run

926 associations with the complete M5 and M5' models, as well as the M5 models  
927 excluding either kinship or both structure and kinship. For each trait, we generated  
928 QQplots of P-values for each of these models.

929 Multiple SNP models were built by successively adding at each step the most  
930 significant SNP, as long as its FDR was lower than 0.10. We controlled for population  
931 structure considering 11 populations and used the kinship matrix that excluded the  
932 SSR on the same chromosome as the last tested SNP.

933

### 934 **Environmental correlation of outlier SNPs**

935 We tested associations between allelic frequency at 171 outlier SNPs and  
936 environmental variables across 28 populations, using Bayenv 2.0 [40, 111]. Because  
937 environmental variables are highly correlated, we used the first two principal  
938 component axes from the environmental PCA analysis (PCenv1 and PCenv2) to run  
939 Bayenv 2.0. This software requires a neutral covariance matrix, that we computed  
940 from the available dataset of 1000 neutral SNPs (S1 Table). We performed 100,000  
941 iterations, saving the matrix every 500 iterations. We then tested the correlation of  
942 these to the last matrix obtained, as well as to an  $F_{ST}$  matrix calculated with  
943 BEDASSLE [135], as described in [57].

944 For each outlier SNP, we compared the posterior probability of a model that  
945 included an environmental factor (PCenv1 or PCenv2) to a null model. We  
946 determined a 5% threshold for significance of environmental association by running  
947 100,000 iterations on neutral SNPs. We carried out five independent runs for each  
948 outlier SNP and evaluated their consistency from the coefficient of variation of the  
949 Bayes factors calculated among runs.

950           In order to test whether environmental distance was a better predictor of allele  
951 frequencies at candidate SNPs than geography, we used multiple regression on  
952 distance matrices (MRM, [136]) implemented in the ecodist R package [137] for each  
953 outlier SNP. We used pairwise  $F_{ST}$  values as the response distance matrix and the  
954 geographic and environmental distance matrices as explanatory matrices. We  
955 evaluated the significance of regression coefficients by 1000 permutations and  
956 iterations of the MRM. We determined the total number of environmentally and  
957 geographically associated SNPs ( $P$ -value<0.05) among outliers. We employed the  
958 same methodology for our set of 1000 neutral SNPs.

959

## 960 **Ethics Statement**

961 All the field work has been done in Mexico in collaboration with Instituto Nacional  
962 de Investigaciones Forestales, Agrícolas y Pecuarias, Celaya in Celaya.

963

## 964 **Acknowledgments**

965 We thank Jessica Melique for her help in gathering stomata data. We are very grateful  
966 to Angelica Cibrian at Langebio (Cinvestav, Irapuato, Mexico) who let us use her lab  
967 to lyophilize our samples, and store them. Valeria Souza greatly helped us with  
968 logistic support in Mexico. Insights from Delphine Legrand, Pierre de Villemereuil  
969 and Elodie Marchadier were important for analyzing correlations between traits and to  
970 estimate heritabilities. We would like to thank warmly five anonymous reviewers for  
971 their insightful comments on our work.

## Supporting information captions

**Figure S1: Altitudinal profiles along gradients *a* and *b*.** Sampled populations are plotted on parallel altitudinal profiles for gradients *a* and *b*. Darker gray lines indicate lower latitude for gradient *a* and lower longitude for gradient *b*. Sampled populations are plotted by green circles (*parviglumis*) or red triangles (*mexicana*). The altitude of the two experimental fields (CEBAJ: 1750m and SENGUA: 2017m) are marked with stars on y-axes.

**Figure S2: Principal Component Analysis of 19 climate variables for 37 teosinte populations.** A: Projection of *parviglumis* (in green) and *mexicana* (in red) populations on the first PCA plane with gradients *a* and *b* indicated by triangles and circles, respectively. The 11 populations evaluated in common gardens are surrounded by a purple outline. Populations that were previously sequenced to detect selection footprints are shown in bold (S1 Table). B: Correlation circle of the 19 climatic variables on the first PCA plane. Climatic variables indicated as T<sub>n</sub> (n from 1 to 11) and P<sub>n</sub> (n from 12 to 19) are related to temperature and precipitation, respectively. Altitude, Latitude and Longitude (in blue) were added as supplementary variables, and CEBAJ and SENGUA field locations were added as supplementary individuals.

**Figure S3: Box-plots of means adjusted by field, year and block, for all traits.** Populations are ranked by altitude. *parviglumis* populations are shown in green and *mexicana* in red. Lighter colors are used for gradient 'a' and darker colors for gradient 'b'. Units of measurement correspond to those defined in S2 Table. For male and

female flowering time, we report values for all 11 populations although very few individuals from the two most lowland populations (P1a and P2b) flowered. Covariation with altitude was significant for all traits except for the number of nodes with ears on the main tiller (S3 Table).

**Figure S4: Pairwise correlations between phenotypic traits.** Pearson coefficient sign and magnitude for significant correlations between phenotypic traits after correction for experiment design (Model M'1). X: correlations that are not significant.

**Figure S5. Evanno method calculations for population number  $\Delta K$  in the association panel genotyped for 38 SSRs.**

**Figure S6. Genetic clustering of ancestry proportions in the association panel genotyped for 38 SSRs.** Genetic clustering was computed for  $K=2$  to  $K=11$ . Vertical lines (individuals) are partitioned into coloured segments whose length represents the admixture proportions from the  $K$  clusters.

**Figure S7. Determination of the migration edge number in the TreeMix model.**

Observed Log likelihood values are plotted against the number of migration edges tested from 0 to 10, and two models are fitted to the data (A). Both the simple exponential and the non-linear least squares delivered an optimal value of 3 for the number of migration edges (change points). The model with 3 migration edges explained 98.75% of the variance, a substantial increase from the null model with no migration edge which is 95.7% (B).

**Figure S8: Significance of  $Q_{ST}$ - $F_{ST}$  difference for each trait.** The dotted blue line indicates the 95% threshold of the simulated distributions and the red line refers to the observed difference. In this analysis, we considered as spatially-varying traits those for which the observed difference fell outside the 95% threshold. Note that Plant height was borderline significant. \*: Set of traits detected by DRIFTSEL.

**Figure S9: Genomic  $F_{ST}$ -scans on 6 teosinte populations.** We computed 4 pairwise- $F_{ST}$  values from 6 populations previously sequenced (S1 Table). Those include  $F_{ST}$  between lowland and highland populations of each gradient (P1a-M6a, P2b-M7b) as well as within subspecies on gradient *b* (P2b-P8b, M1b-M7b).  $F_{ST}$  values are averaged across sliding windows of 20 SNPs with a step of five SNPs (from top to bottom, chromosome 1 to 10) and normalized by subtracting the  $F_{ST}$  mean and dividing by the standard deviation across pairwise comparisons. Only the top 1% values are represented. The 1% thresholds for each pairwise comparisons are indicated by colored horizontal lines. Horizontal black bars indicate location of inversions on chromosome 1 (*Inv1n*), chromosome 4 (*Inv4m*) and chromosome 9 (*Inv9d*). The subset of 171 outlier SNPs analyzed in the present study is indicated with black diamond marks along the X axes.

**Figure S10: QQ-plots of observed P-values and expected P-values generated from 38 SSRs.** We employed three versions of the model M5 with correction for neither structure nor kinship, with correction for genetic structure (at  $K=5$ ), with correction for genetic structure (at  $K=5$  and with 11 populations) and kinship.

**Figure S11: Manhattan plots of associations between 171 outlier SNPs and 12 phenotypic traits.** X-axis indicates the positions of outlier SNPs on chromosomes 1 to 10, black and gray colors alternating per chromosome. Plotted on the Y-axis are the negative  $\text{Log}_{10}$ -transformed  $P$  values obtained for the  $K=5$  model. Significant associations (10% FDR) are indicated considering either a structure matrix at  $K=5$  (pink dots), for 11 populations (blue dots), or for both  $K=5$  and 11 populations models (purple dots).

**Figure S12: Pairwise Linkage Disequilibrium (LD) between outlier SNPs.**

Pairwise LD between 171 SNPs was estimated using  $r^2$ , and corrected for structure at  $K=5$  and kinship computed from 38 SSRs. Blue shaded bars show the 23 SNPs found to associate with at least one phenotype under the 11 populations structure correction.

**S1 Table. Description of 37 teosinte populations and sets of populations used in the present study by data types.**

**S2 Table. List of the 18 phenotypic traits measured and estimates of narrow-sense heritabilities ( $h^2$ ).**

**S3 Table. Significance of main effects for each trait as determined by models M1 and M3.**

**S4 Table. Description of 46 SSRs and genotyping success rate.**

**S5 Table. Characteristics, association with phenotypes, effects and correlation with environment of outlier SNPs.**

**S6 Table. Number of individuals used to test associations between 171 SNPs and 18 phenotypes.**

**S7 Table. Additive and dominance effects of SNPs associated to traits after the 11-population structure correction.**



## References

- 972 1. Whitlock MC. Modern approaches to local adaptation. *The American*  
973 *Naturalist*. 2015;186.
- 974 2. Bradshaw AD. Ecological significance of genetic variation between  
975 populations. *Perspectives on plant population ecology*. 1984:213-28.
- 976 3. Bulmer MGG. Multiple niche polymorphism. *The American Naturalist*.  
977 1972;106:254-7.
- 978 4. Endler JA. Natural selection in the wild. 1986:354.
- 979 5. Gay L, Crochet PA, Bell DA, Lenormand T. Comparing clines on molecular  
980 and phenotypic traits in hybrid zones: A window on tension zone models. *Evolution*.  
981 2008;62:2789-806.
- 982 6. Lande R. Natural selection and random genetic drift in phenotypic evolution.  
983 *Evolution*. 1976;30:314-34.
- 984 7. Lenormand T. Gene flow and the limits to natural selection. *Trends in*  
985 *Ecology and Evolution*. 2002;17:183-9.
- 986 8. Whitlock MC, Gomulkiewicz R. Probability of fixation in a heterogeneous  
987 environment. *Genetics*. 2005:1-40.
- 988 9. Yeaman S, Otto SP. Establishment and maintenance of adaptive genetic  
989 divergence under migration, selection and drift. *Evolution*. 2011;67:2123-9.
- 990 10. Rundle HD, Nosil P. Ecological speciation. *Ecology Letters*. 2005;8:336-52.
- 991 11. Kawecki TJ, Ebert D. Conceptual issues in local adaptation. *Ecology Letters*.  
992 2004;7:1225-41.
- 993 12. Hereford J. A quantitative survey of local adaptation and fitness trade-offs.  
994 *The American Naturalist*. 2009;173:579-88.
- 995 13. Leimu R, Fischer M. A meta-analysis of local adaptation in plants. *PLoS*  
996 *ONE*. 2008;3.
- 997 14. Tiffin P, Ross-Ibarra J. Advances and limits of using population genetics to  
998 understand local adaptation. *Trends in Ecology and Evolution*. 2014;29:673-80.
- 999 15. Garcia-Ramos G, Kirkpatrick M. Genetic models of adaptation and gene flow  
1000 in peripheral populations. *Evolution*. 1997;51:21-8.
- 1001 16. Slatkin M. Rare alleles as indicators of gene flow. *Evolution*. 1985;39(1):53-  
1002 65.
- 1003 17. Lande R. Neutral theory of quantitative genetic variance in an island model  
1004 with local extinction and colonization. *Evolution*. 1992;46:381-9.
- 1005 18. Whitlock MC. Neutral additive genetic variance in a metapopulation. *Genetics*  
1006 *Research*. 1999;74:215-21.
- 1007 19. Spitze K. Population structure in *Daphnia obtusa*: quantitative genetic and  
1008 allozymic variation. *Genetics Society of America*. 1993;135:367-74.
- 1009 20. Wright S. The genetical structure of populations. *Annals of Eugenics*.  
1010 1951;15:323-54.
- 1011 21. Luquet E, Léna J-P, Miaud C, Plénet S. Phenotypic divergence of the common  
1012 toad (*Bufo bufo*) along an altitudinal gradient: evidence for local adaptation. *Heredity*.  
1013 2015;114:69-79.
- 1014 22. Roschanski AM, Csilléry K, Liepelt S, Oddou-Muratorio S, Ziegenhagen B,  
1015 Huard Fdr, et al. Evidence of divergent selection for drought and cold tolerance at  
1016 landscape and local scales in *Abies alba* Mill. in the French Mediterranean Alps.  
1017 *Molecular Ecology*. 2016;25:776-94.

- 1018 23. Kawakami T, Morgan TJ, Nippert JB, Ocheltree TW, Keith R, Dhakal P, et al.  
1019 Natural selection drives clinal life history patterns in the perennial sunflower species,  
1020 *Helianthus maximiliani*. *Molecular Ecology*. 2011;20:2318-28.
- 1021 24. Moyers BT, Rieseberg LH. Remarkable life history polymorphism may be  
1022 evolving under divergent selection in the silverleaf sunflower. *Molecular Ecology*.  
1023 2016;25:3817-30.
- 1024 25. Kirkpatrick M, Barton N. Chromosome inversions, local adaptation and  
1025 speciation. *Genetics*. 2006.
- 1026 26. Lowry DB, Willis JH. A widespread chromosomal inversion polymorphism  
1027 contributes to a major life-history transition, local adaptation, and reproductive  
1028 isolation. *PLoS Biology*. 2010;8.
- 1029 27. Legrand D, Larranaga N, Bertrand R, Ducatez S, Calvez O, Stevens VM, et al.  
1030 Evolution of a butterfly dispersal syndrome. 2016.
- 1031 28. Bierne N, Welch J, Loire E, Bonhomme F, David P. The coupling hypothesis:  
1032 Why genome scans may fail to map local adaptation genes. *Molecular Ecology*.  
1033 2011;20:2044-72.
- 1034 29. Lewontin RC, Krakauer J. Distribution of gene frequency as a test of the  
1035 theory of the selective neutrality of polymorphisms. *Genetics*. 1973;74:175-95.
- 1036 30. Beaumont MA, Nichols RA. Evaluating loci for use in the genetic analysis of  
1037 population structure. *Proceedings of the Royal Society B: Biological Sciences*.  
1038 1996;263:1619-26.
- 1039 31. Vitalis R, Dawson K, Boursot P. Interpretation of variation across marker loci  
1040 as evidence of selection. *Genetics*. 2001;158:1811-23.
- 1041 32. Foll M, Gaggiotti O. A genome scan method to identify selected loci  
1042 appropriate for both dominant and codominant markers: A Bayesian perspective.  
1043 *Genetics*. 2008;180:977-93.
- 1044 33. Excoffier LaH, T and Foll, Matthieu. Detecting loci under selection in a  
1045 hierarchically structured population. *Heredity*. 2009;103:285-98.
- 1046 34. Bonhomme M, Chevalet C, Servin B, Boitard S, Abdallah J, Blott S, et al.  
1047 Detecting selection in population trees: The Lewontin and Krakauer test extended.  
1048 *Genetics*. 2010;186:241-62.
- 1049 35. Günther T, Coop G. Robust identification of local adaptation from allele  
1050 frequencies. *Genetics Society of America*. 2013;195:205-20.
- 1051 36. Lotterhos KE, Whitlock MC. Evaluation of demographic history and neutral  
1052 parameterization on the performance of FST outlier tests. *Molecular Ecology*.  
1053 2014;23:2178-92.
- 1054 37. Haasl RJ, Payseur BA. Fifteen years of genomewide scans for selection:  
1055 trends, lessons and unaddressed genetic sources of complication. *Molecular Ecology*.  
1056 2016;25:5-23.
- 1057 38. Le Corre V, Kremer A. The genetic differentiation at quantitative trait loci  
1058 under local adaptation. *Molecular Ecology*. 2012;21:1548-66.
- 1059 39. Yi X, Liang Y, Huerta-Sanchez E, Jin X, Xi Ping Cuo Z, Pool JE, et al.  
1060 Sequencing of fifty human exomes reveals adaptations to high altitude. *Science*.  
1061 2010;329:75-8.
- 1062 40. Coop G, Witonsky D, Di Rienzo A, Pritchard JK. Using environmental  
1063 correlations to identify loci underlying local adaptation. *Genetics*. 2010;185:1411-23.
- 1064 41. Guillot G, Renaud S, Ledevin R, Michaux J, Claude J. A unifying model for  
1065 the analysis of phenotypic, genetic, and geographic data. 2012;61:897-911.

- 1066 42. Frichot E, Schoville SD, Bouchard G, François O. Testing for associations  
1067 between loci and environmental gradients using latent factor mixed models.  
1068 *Molecular Biology and Evolution*. 2013;30:1687-99.
- 1069 43. Gautier M. Genome-wide scan for adaptive divergence and association with  
1070 population-specific covariates. *Genetics*. 2015;201:1555-79.
- 1071 44. Joost S, Bonin A, Bruford MW, Després L, Conord C, Erhardt G, et al. A  
1072 spatial analysis method (SAM) to detect candidate loci for selection: Towards a  
1073 landscape genomics approach to adaptation. *Molecular Ecology*. 2007;16:3955-69.
- 1074 45. Poncet BN, Herrmann D, Gugerli F, Taberlet P, Holderegger R, Gielly L, et  
1075 al. Tracking genes of ecological relevance using a genome scan in two independent  
1076 regional population samples of *Arabidopsis thaliana*. *Molecular Ecology*. 2010;19:2896-907.
- 1077 46. De Mita S, Thuillet AC, Gay L, Ahmadi N, Manel S, Ronfort J, et al.  
1078 Detecting selection along environmental gradients: Analysis of eight methods and  
1079 their effectiveness for outbreeding and selfing populations. *Molecular Ecology*.  
1080 2013;22:1383-99.
- 1081 47. Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, et al.  
1082 Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future  
1083 directions. *The American Naturalist*. 2016;188:379-97.
- 1084 48. Barton N, Hermisson J, Nordborg M. Why structure matters. *eLife*. 2019;8.
- 1085 49. Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek  
1086 AM. A map of local adaptation in *Arabidopsis thaliana*. *Science*. 2011;334:86-9.
- 1087 50. Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, Sperone FG,  
1088 et al. Adaptation to climate across the *Arabidopsis thaliana* genome. *Science*.  
1089 2011;334:83-6.
- 1090 51. Ross-Ibarra J, Tenaillon M, Gaut BS. Historical divergence and gene flow in  
1091 the genus *Zea*. *Genetics*. 2009;181:1399-413.
- 1092 52. Hufford MB, Martínez-Meyer E, Gaut BS, Eguiarte LE, Tenaillon MI.  
1093 Inferences from the historical distribution of wild and domesticated maize provide  
1094 ecological and evolutionary insight. *PLoS ONE*. 2012;7.
- 1095 53. Bilinski P, Albert PS, Berg JJ, Birchler JA, Grote MN, Lorant A, et al. Parallel  
1096 altitudinal clines reveal trends in adaptive evolution of genome size in *Zea mays*.  
1097 *PLoS Genetics*. 2018;14.
- 1098 54. Diez CM, Gaut BS, Meca E, Scheinvar E, Montes-Hernandez S, Eguiarte LE,  
1099 et al. Genome size variation in wild and cultivated maize along altitudinal gradients.  
1100 *New Phytologist*. 2013;199:264-76.
- 1101 55. Pyhäjärvi T, Hufford MB, Mezouk S, Ross-Ibarra J. Complex patterns of  
1102 local adaptation in teosinte. *Genome Biology and Evolution*. 2013;5:1594-609.
- 1103 56. Fang Z, Pyhäjärvi T, Weber AL, Dawe RK, Glaubitz JC, Sánchez González  
1104 JdJ, et al. Megabase-scale inversion polymorphism in the wild ancestor of maize.  
1105 *Genetics*. 2012;191:883-94.
- 1106 57. Aguirre-Liguori JA, Tenaillon MI, Vázquez-Lobo A, Gaut BS, Jaramillo-  
1107 Correa JP, Montes-Hernandez S, et al. Connecting genomic patterns of local  
1108 adaptation and niche suitability in teosintes. *Molecular Ecology*. 2017;26:4226-40.
- 1109 58. Fustier MA, Brandenburg JT, Boitard S, Lapeyronnie J, Eguiarte LE,  
1110 Vigouroux Y, et al. Signatures of local adaptation in lowland and highland teosintes  
1111 from whole-genome sequencing of pooled samples. *Molecular Ecology*.  
1112 2017;26:2738-56.
- 1113 59. Ovaskainen O, Karhunen M, Zheng C, Arias JMC, Merilä J. A new method to  
1114 uncover signatures of divergent and stabilizing selection in quantitative traits.  
1115 *Genetics*. 2011;189:621-32.

- 1116 60. McKinney GJ, Varian A, Scardina J, Nichols KM. Genetic and morphological  
1117 divergence in three strains of brook trout *Salvelinus fontinalis* commonly stocked in  
1118 Lake Superior. PLoS ONE. [Research Support, Non-U.S. Gov't].  
1119 2014;9(12):e113809.
- 1120 61. Sohail M, Maier RM, Ganna A, Bloemendal A, Martin AR, Turchin MC, et al.  
1121 Polygenic adaptation on height is overestimated due to uncorrected stratification in  
1122 genome-wide association studies. eLife. 2019;8.
- 1123 62. Desrousseaux AD, Sandron F, Siberchicot A, Cierco-Ayrolles C, Mangin B.  
1124 Package 'LDcorSV'. 2017.
- 1125 63. Mangin B, Siberchicot A, Nicolas S, Doligez A, This P, Cierco-Ayrolles C.  
1126 Novel measures of linkage disequilibrium that correct the bias due to population  
1127 structure and relatedness. Heredity. 2012;108:285-91.
- 1128 64. Savolainen O, Lascoux M, Merilä J. Ecological genomics of local adaptation.  
1129 Nature reviews Genetics. 2013;14:807-20.
- 1130 65. Anderson JT, Willis JH, Mitchell-Olds T. Evolutionary genetics of plant  
1131 adaptation. Trends in Genetics. 2011;27:258-66.
- 1132 66. Halbritter AH, Fior S, Keller I, Billeter R, Edwards PJ, Holderegger R, et al.  
1133 Trait differentiation and adaptation of plants along elevation gradients. Journal of  
1134 Evolutionary Biology. 2018;31(6):784-800.
- 1135 67. Körner C. The use of 'altitude' in ecological research. Trends in Ecology and  
1136 Evolution. 2007;22:569-74.
- 1137 68. Friend AD, Woodward FI, Switsur VR. Field measurements of  
1138 photosynthesis, stomatal conductance, leaf nitrogen and  $\delta^{13}\text{C}$  along altitudinal  
1139 gradients in Scotland. Functional Ecology. 1989;3:117.
- 1140 69. Neuner G. Frost resistance in alpine woody plants. Frontiers in Plant Science.  
1141 2014;5.
- 1142 70. Frohnmeyer H, Staiger D. Update on ultraviolet-B light responses ultraviolet-  
1143 B radiation-mediated responses in plants. Balancing damage and protection.  
1144 2014;133:1420-8.
- 1145 71. Byars SG, Papst W, Hoffmann AA. Local adaptation and cogradient selection  
1146 in the alpine plant, *Poa hiemata*, along a narrow altitudinal gradient. Evolution.  
1147 2007;61:2925-41.
- 1148 72. Luo Y, Widmer A, Karrenberg S. The roles of genetic drift and natural  
1149 selection in quantitative trait divergence along an altitudinal gradient in *Arabidopsis*  
1150 *thaliana*. Heredity. 2015;114:220-8.
- 1151 73. Guerin GR, Wen H, Lowe AJ, Guerin GR, Wen H. Leaf morphology shift  
1152 linked to climate change. Population Ecology. 2012:882-6.
- 1153 74. Kofidis G, Bosabalidis AM, Moustakas M. Contemporary seasonal and  
1154 altitudinal variations of leaf structural features in oregano (*Origanum vulgare* L.).  
1155 Annals of Botany. 2003;92(5):635-45.
- 1156 75. Mendez-Vigo B, Pico FX, Ramiro M, Martinez-Zapater JM, Alonso-Blanco  
1157 C. Altitudinal and climatic adaptation is mediated by flowering traits and FRI, FLC,  
1158 and PHYC genes in *Arabidopsis*. Plant Physiology. 2011;157:1942-55.
- 1159 76. Oleksyn J, Modrzynski J, Tjoelker MG, Zytkowaik R, Reich PB, Karolewski  
1160 P. Growth and physiology of *Picea abies* populations from elevational transects:  
1161 common garden evidence for altitudinal ecotypes and cold adaptation. 1998:573-90.
- 1162 77. Soularue JP, Kremer A. Evolutionary responses of tree phenology to the  
1163 combined effects of assortative mating, gene flow and divergent selection. Heredity.  
1164 2014;113:485-94.

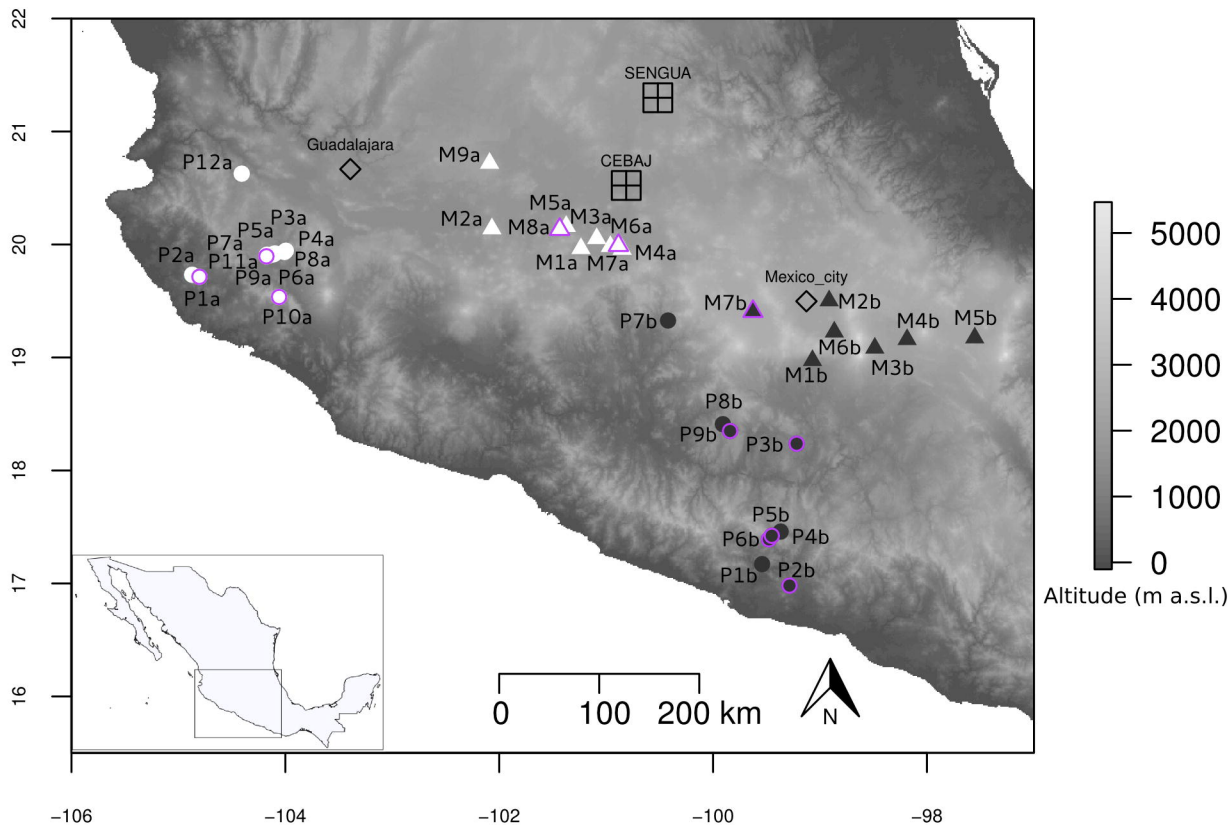
- 1165 78. Doebley JF. Maize introgression into teosinte -- a reappraisal. *Annals of the*  
1166 *Missouri Botanical Garden*. 1984;71:1100-13.
- 1167 79. Lauter N, Gustus C, Westerbergh A, Doebley J. The inheritance and evolution  
1168 of leaf pigmentation and pubescence in teosinte. *Genetics Society of America*.  
1169 2004;167:1949-59.
- 1170 80. Smith JSC, Goodman MM, Lester RN. Variation within teosinte. I. Numerical  
1171 analysis of morphological data. *Economic Botany*. 1981;35:187-203.
- 1172 81. Raven J, A. Selection pressures on stomatal evolution. *New Phytologist*.  
1173 2002;153:371-86.
- 1174 82. Dittberner H, Korte A, Mettler-Altmann T, Weber APM, Monroe G, de  
1175 Meaux J. Natural variation in stomata size contributes to the local adaptation of  
1176 water-use efficiency in *Arabidopsis thaliana*. *Molecular Ecology*. 2018:4052-65.
- 1177 83. Carlson JE, Adams CA, Holsinger KE. Intraspecific variation in stomatal  
1178 traits, leaf traits and physiology reflects adaptation along aridity gradients in a South  
1179 African shrub. *Annals of Botany*. 2016;117:195-207.
- 1180 84. Körner C, Mayr R. Stomatal behaviour in alpine plant communities between  
1181 600 and 2600 metres above sea level. In: Grace J, Ford ED, Jarvis PG, editors. *Plants*  
1182 *and their Atmospheric Environment*: Blackwell, Oxford,; 1981. p. pp 205-18.
- 1183 85. Bresson CC, Vitasse Y, Kremer A, Delzon S. To what extent is altitudinal  
1184 variation of functional traits driven by genetic adaptation in European oak and beech?  
1185 *Tree Physiology*. 2011;31:1164-74.
- 1186 86. Kooyers NJ, Greenlee AB, Colicchio JM, Oh M, Blackman BK. Replicate  
1187 altitudinal clines reveal that evolutionary flexibility underlies adaptation to drought  
1188 stress in annual *Mimulus guttatus*. *New Phytologist*. 2015;206:152-65.
- 1189 87. Körner C, Neumayer M, Menendez-Riedl SP, Smeets-Scheel A. Functional  
1190 morphology of mountain plants. *Flora*. 1989;182:353-83.
- 1191 88. Jakobsson A, Eriksson O. A comparative study of seed number, seed size,  
1192 seedling size and recruitment in grassland plants. *Oikos*. 2000;88:494-502.
- 1193 89. Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, et  
1194 al. The genetic architecture of maize flowering time. *Science*. 2009;325:714-8.
- 1195 90. Moreau L, Charcosset A, Gallais A. Use of trial clustering to study QTL x  
1196 environment effects for grain yield and related traits in maize. *Theoretical and*  
1197 *Applied Genetics*. 2004;110:92-105.
- 1198 91. Durand E, Bouchet S, Bertin P, Ressayre A, Jamin P, Charcosset A, et al.  
1199 Flowering time in maize: Linkage and epistasis at a major effect locus. *Genetics*.  
1200 2012;190:1547-62.
- 1201 92. Li D, Wang X, Zhang X, Chen Q, Xu G, Xu D, et al. The genetic architecture  
1202 of leaf number and its genetic relationship to flowering time in maize. *New*  
1203 *Phytologist*. 2015;210:256-68.
- 1204 93. Aguirre - Liguori JA, Gaut BS, Jaramillo - Correa JP, Tenaillon MI, Montes  
1205 - Hernández S, García - Oliva F, et al. Divergence with gene flow is driven by local  
1206 adaptation to temperature and soil phosphorus concentration in teosinte subspecies (  
1207 *Zea mays parviglumis* and *Zea mays mexicana*) *Molecular Ecology*. 2019:2814-30.
- 1208 94. Consortium G. 1,135 Genomes reveal the global pattern of polymorphism in  
1209 *Arabidopsis thaliana*. *Cell*. 2016;166:481-91.
- 1210 95. Novembre J, Barton NH. Tread lightly interpreting polygenic tests of  
1211 selection. *Genetics*. 2018; 208(4):1351-5.
- 1212 96. Guo J, Wu Y, Zhu Z, Zheng Z, Trzaskowski M, Zeng J, et al. Global genetic  
1213 differentiation of complex traits shaped by natural selection in humans. *Nature*  
1214 *Communications*. 2018; 9(1):1865.

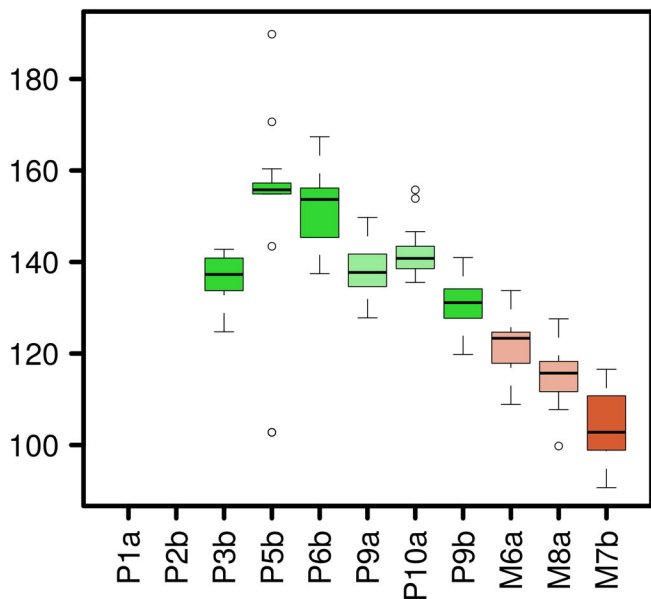
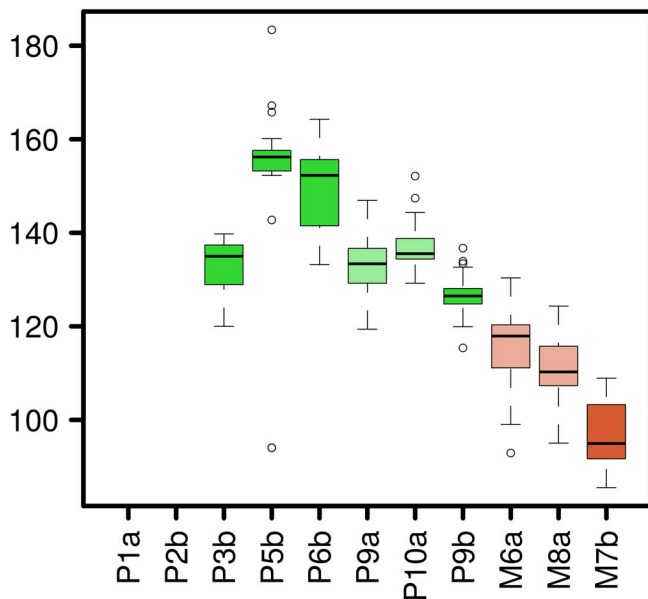
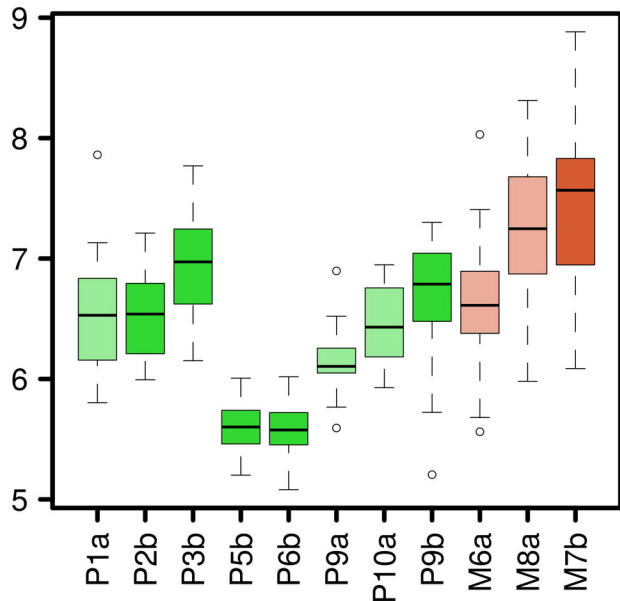
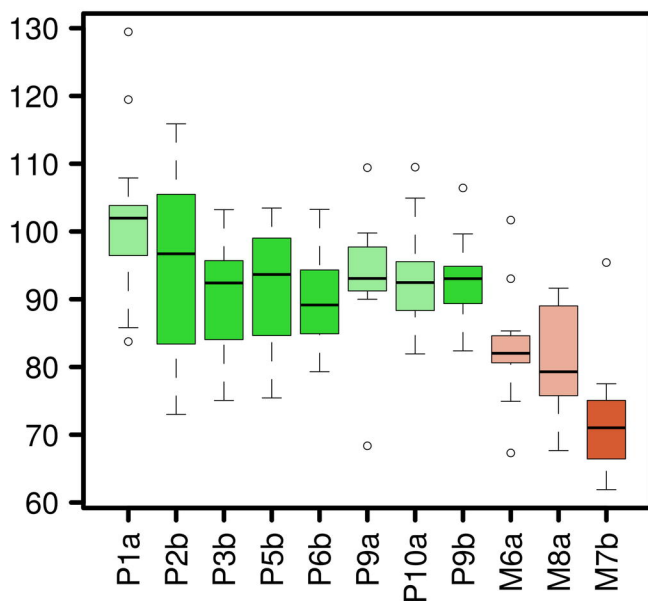
- 1215 97. Berg JJ, Harpak A, Sinnott-Armstrong N, Joergensen AM, Mostafavi H, Field  
1216 Y, et al. Reduced signal for polygenic adaptation of height in UK Biobank. *eLife*.  
1217 2019;8:1-47.
- 1218 98. Whitt SR, Wilson LM, Tenaillon MI, Gaut BS, Buckler ES. Genetic diversity  
1219 and selection in the maize starch pathway. *Proceedings of the National Academy of*  
1220 *Sciences of the United States of America*. 2002;99:12959-62.
- 1221 99. Jaenicke-Despres V, Buckler ES, Smith BD, Gilbert MTP, Cooper A, Doebley  
1222 J, et al. Early allelic selection in maize as revealed by ancient DNA. 2003;302:1206-9.
- 1223 100. Weber AL, Briggs WH, Rucker J, Baltazar BM, De Jesús Sánchez-González  
1224 J, Feng P, et al. The genetic architecture of complex traits in teosinte (*Zea mays* ssp.  
1225 *parviglumis*): New evidence from association mapping. *Genetics*. 2008;180:1221-32.
- 1226 101. Bouchet S, Servin B, Bertin P, Madur D, Combes V, Dumas F, et al.  
1227 Adaptation of maize to temperate climates: Mid-density genome-wide association  
1228 genetics and diversity patterns reveal key genomic regions, with a major contribution  
1229 of the Vgt2 (ZCN8) locus. *PLoS ONE*. 2013;8.
- 1230 102. Sheehan MJ, Kennedy LM, Costich DE, Brutnell TP. Subfunctionalization of  
1231 PhyB1 and PhyB2 in the control of seedling and mature plant traits in maize. *Plant*  
1232 *Journal*. 2007;49:338-53.
- 1233 103. Danilevskaya ON, Meng X, Hou Z, Ananiev EV, Simmons CR. A genomic  
1234 and expression compendium of the expanded PEBP gene family from maize. *Plant*  
1235 *Physiology*. 2007;146:250-64.
- 1236 104. Meng X, Muszynski MG, Danilevskaya ON. The FT-Like ZCN8 gene  
1237 functions as a floral activator and is involved in photoperiod sensitivity in maize. *The*  
1238 *Plant Cell*. 2011;23:942-60.
- 1239 105. Li YX, Li C, Bradbury PJ, Liu X, Lu F, Romay CM, et al. Identification of  
1240 genetic variants associated with maize flowering time using an extremely large multi-  
1241 genetic background population. *The Plant journal : for cell and molecular biology*.  
1242 2016;86:391-402.
- 1243 106. Yu J, Li X, Zhu C, Yeh C-T, Wu W, Takacs E, et al. Genic and non-genic  
1244 contributions to natural variation of quantitative traits in maize. *Genome research*.  
1245 2012:2436-44.
- 1246 107. Wellenreuther M, Bernatchez L. Eco-evolutionary genomics of chromosomal  
1247 inversions. *Trends in Ecology and Evolution*. 2018;33:427-40.
- 1248 108. Ayala D, Ullastres A, González J. Adaptation through chromosomal  
1249 inversions in *Anopheles*. *Frontiers in Genetics*. 2014;5:1-10.
- 1250 109. Barth JMI, Berg PR, Jonsson PR, Bonanomi S, Corell H, Hemmer-Hansen J,  
1251 et al. Genome architecture enables local adaptation of Atlantic cod despite high  
1252 connectivity. *Molecular Ecology*. 2017;26:4452-66.
- 1253 110. Lundberg M, Liedvogel M, Larson K, Sigeman H, Grahn M, Wright A, et al.  
1254 Genetic differences between willow warbler migratory phenotypes are few and cluster  
1255 in large haplotype blocks. *Evolution Letters*. 2017:155-68.
- 1256 111. Twyford AD, Friedman J. Adaptive divergence in the monkey flower *Mimulus*  
1257 *guttatus* is maintained by a chromosomal inversion. *Evolution*. 2015;69:1476-86.
- 1258 112. Díez CM, Gaut BS, Meca E, Scheinvar E, Montes - Hernandez S, Eguiarte  
1259 LE, et al. Genome size variation in wild and cultivated maize along altitudinal  
1260 gradients. *New Phytologist*. 2013;199(1):264-76.
- 1261 113. Hijmans RJ, van Etten J, Cheng J, Mattiuzzi M, Sumner M, Greenberg JA, et  
1262 al. Package ‘raster’: geographic data analysis and modeling. 2018.
- 1263 114. Husson F, Josse J, Le S, Mazet J. Package ‘FactoMineR’. An R package.  
1264 2016:96.

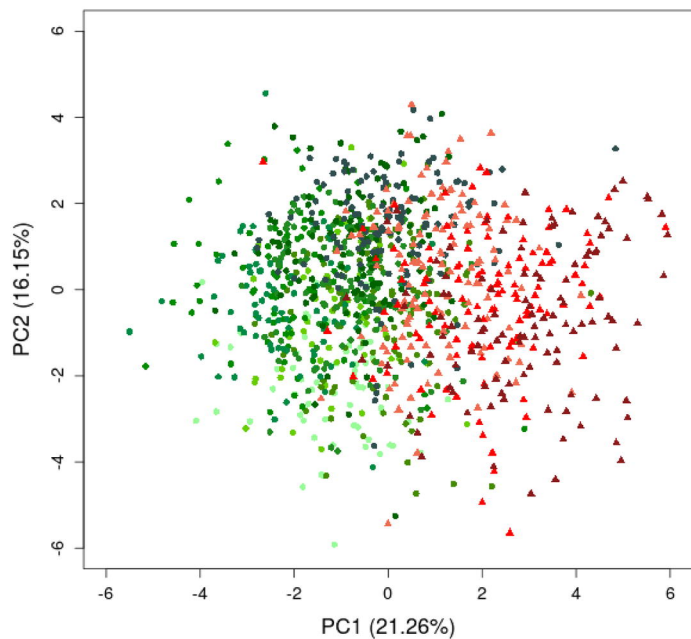
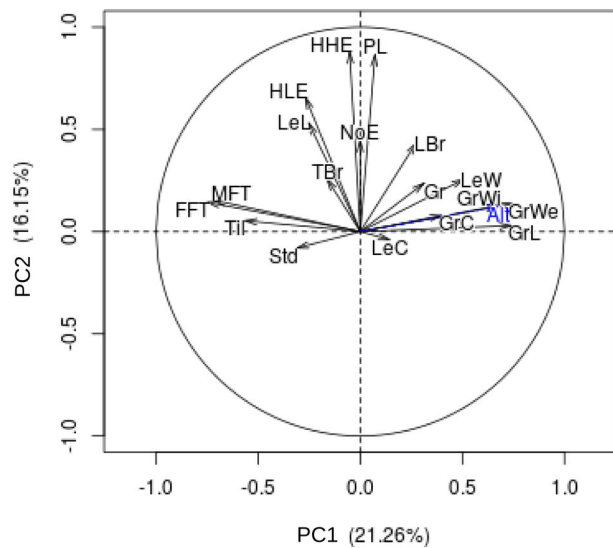
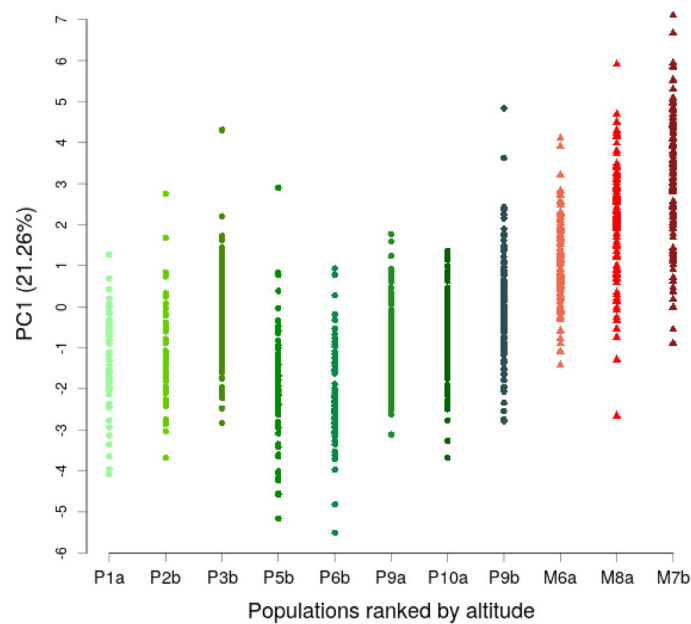
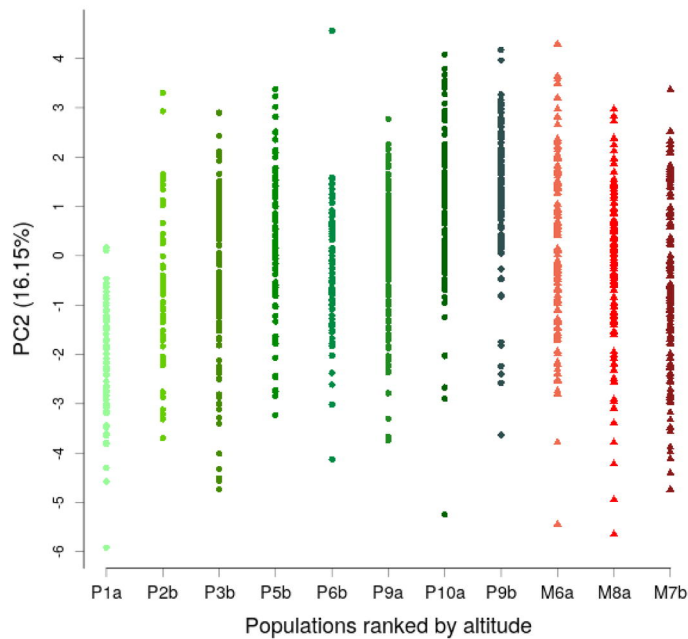
- 1265 115. Andorf CM, Cannon EK, Portwood JL, Gardiner JM, Harper LC, Schaeffer  
1266 ML, et al. MaizeGDB update: New tools, data and interface for the maize model  
1267 organism database. *Nucleic Acids Research*. 2016;44:1195-201.
- 1268 116. Camus-Kulandaivelu L, Veyrieras JB, Madur D, Combes V, Fourmann M,  
1269 Barraud S, et al. Maize adaptation to temperate climate: Relationship between  
1270 population structure and polymorphism in the Dwarf8 gene. *Genetics*.  
1271 2006;172:2449-63.
- 1272 117. Guichoux E, Lagache S, Wagner S, Chaumeil P, Léger P, Lepais O, et al.  
1273 Current trends in microsatellite genotyping. *Molecular Ecology Resources*.  
1274 2011;11:591-611.
- 1275 118. Jakobsson M, Rosenberg NA. CLUster Matching and Permutation Program  
1276 Version 1.1.2. 2007.
- 1277 119. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of  
1278 individuals using the software STRUCTURE: A simulation study. *Molecular*  
1279 *Ecology*. 2005;14:2611-20.
- 1280 120. Hardy OJ, Vekemans X. spagedi: a versatile computer program to analyse  
1281 spatial genetic structure at the individual or population levels. *Molecular Ecology*  
1282 *Notes*. 2002;2:618-20.
- 1283 121. Loiselle BA, Sork VL, Nason JD, Graham C. Spatial genetic structure of a  
1284 tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of*  
1285 *Botany*. 1995;82:1420-5.
- 1286 122. Pickrell JK, Pritchard JK. Inference of population splits and mixtures from  
1287 genome-wide allele frequency data. *PLOS Genetics*. 2012;8(11):e1002967.
- 1288 123. Fitak RRs. optM: an R package to optimize the number of migration edges  
1289 using threshold models. *Journal of Heredity*. 2019.
- 1290 124. Sanchez JdJ, Kato Yamakake TA, Aguilar Sanmiguel M, Hernandez Casillas  
1291 JM, Lopez Rodriguez A, Ruiz Corral JA. Distribución y caracterización del teocintle.  
1292 1998:165.
- 1293 125. Butler D, Cullis BR, Gilmour AR, Gogel BJ. ASReml-R reference manual.  
1294 Technical Report. 2007.
- 1295 126. Holsinger KE, Weir BS. Genetics in geographically structured populations:  
1296 defining, estimating and interpreting  $F_{ST}$ . *Nature reviews Genetics*. 2009;10:639-  
1297 50.
- 1298 127. Gilbert KJ, Whitlock MC.  $Q_{ST}$ - $F_{ST}$  comparisons with unbalanced half-sib  
1299 designs. *Molecular Ecology Resources*. 2015;15:262-7.
- 1300 128. Karhunen M, Merilä J, Leinonen T, Cano JM, Ovaskainen O. driftsel: An R  
1301 package for detecting signals of natural selection in quantitative traits. *Molecular*  
1302 *Ecology Resources*. 2013;13:746-54.
- 1303 129. Hadfield JD. MCMC methods for multi-response generalized linear mixed  
1304 models: the MCMCglmm R package. *Journal of Statistical Software*. 2010;33.
- 1305 130. Semagn K, Babu R, Hearne S, Olsen M. Single nucleotide polymorphism  
1306 genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the  
1307 technology and its application in crop improvement. *Molecular Breeding*. 2014;33:1-  
1308 14.
- 1309 131. Weir BS, Hill WG. Estimating F-statistics. *Annu Rev Genet*. 2002;36:721-50.
- 1310 132. Günther T, Coop G. A Short Manual for Bayenv2.0. 2016.
- 1311 133. De Villemereuil P, Gaggiotti OE. A new  $F_{ST}$ -based method to uncover local  
1312 adaptation using environmental variables. *Methods in Ecology and Evolution*. 2015.

- 1313 134. Rincent R, Moreau L, Monod H, Kuhn E, Melchinger AE, Malvar RA, et al.  
1314 Recovering power in association mapping panels with variable levels of linkage  
1315 disequilibrium. *Genetics*. 2014;197:375-87.
- 1316 135. Bradburd GS, Ralph PL, Coop GM. Disentangling the effects of geographic  
1317 and ecological isolation on genetic differentiation. *Evolution*. 2013;67:3258-73.
- 1318 136. Lichstein JW. Multiple regression on distance matrices: A multivariate spatial  
1319 analysis tool. *Plant Ecology*. 2007;188:117-31.
- 1320 137. Goslee SC, Urban DL. *Journal of Statistical Software* The ecodist package for  
1321 dissimilarity-based analysis of ecological data. 2007.

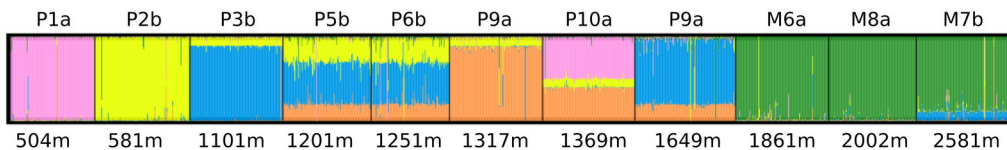




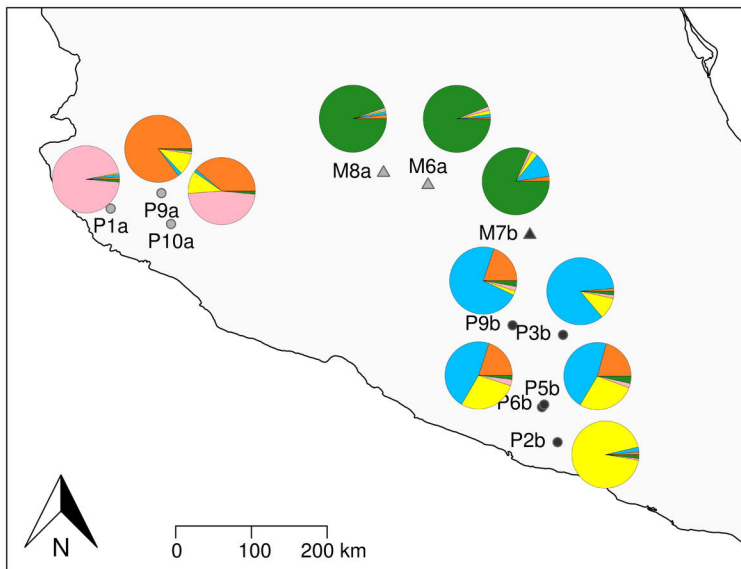
**A. Female flowering time (days)****B. Male flowering time (days)****C. Grain length (mm)****D. Stomata density (per mm<sup>2</sup>)**

**A****B****C****D**

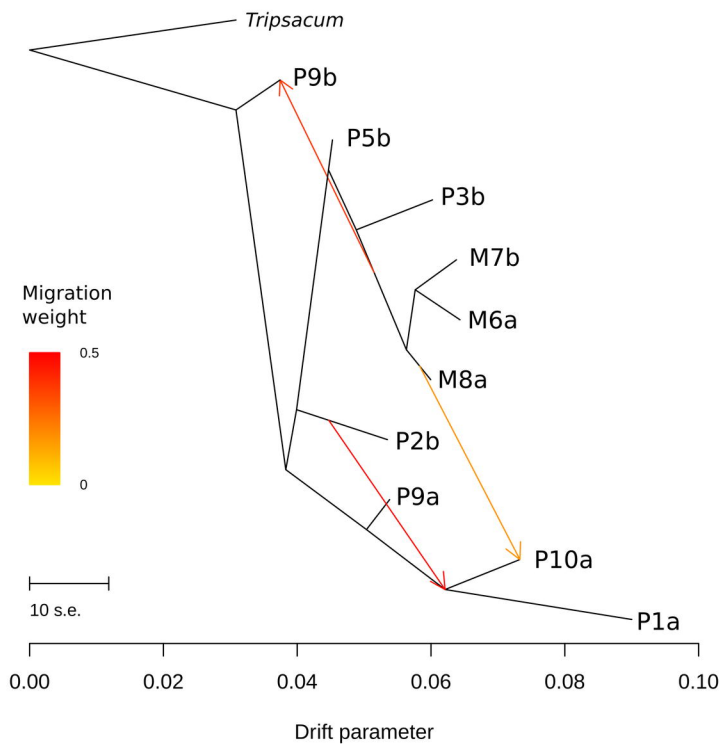
A.



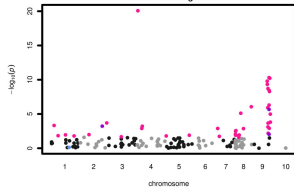
B.



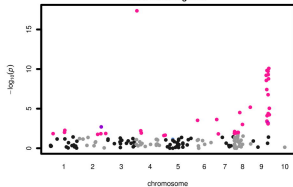
C.



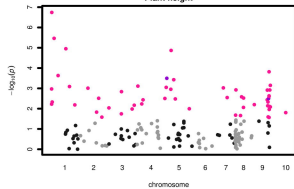
Female flowering time



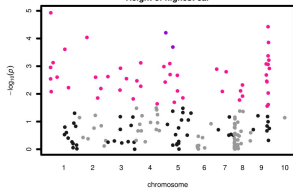
Male flowering time



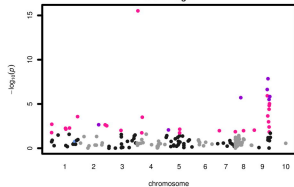
Plant height



Height of highest ear



Grain weight



Stomata density

