

UNIVERSITÀ POLITECNICA DELLE MARCHE Repository ISTITUZIONALE

Adaptation to novel environments during crop diversification

This is the peer reviewd version of the followng article:

Original

Adaptation to novel environments during crop diversification / Cortinovis, Gaia; Di Vittori, Valerio; Bellucci, Elisa; Bitocchi, Elena; Papa, Roberto. - In: CURRENT OPINION IN PLANT BIOLOGY. - ISSN 1369-5266. -56:(2020), pp. 203-217. [10.1016/j.pbi.2019.12.011]

Availability:

This version is available at: 11566/278893 since: 2024-12-05T14:08:34Z

Publisher:

Published DOI:10.1016/j.pbi.2019.12.011

Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. The use of copyrighted works requires the consent of the rights' holder (author or publisher). Works made available under a Creative Commons license or a Publisher's custom-made license can be used according to the terms and conditions contained therein. See editor's website for further information and terms and conditions. This item was downloaded from IRIS Università Politecnica delle Marche (https://iris.univpm.it). When citing, please refer to the published version.

1	Adaptation to novel environments during crop diversification
2	
3	Gaia Cortinovis, Valerio Di Vittori, Elisa Bellucci, Elena Bitocchi*, and Roberto Papa*
4	
5	Dipartimento di Scienze Agrarie, Alimentari ed Ambientali, Università Politecnica delle
6	Marche, via Brecce Bianche, 60131 Ancona, Italy
7	
8	Gaia Cortinovis, g.cortinovis@pm.univpm.it
9	Valerio Di Vittori, v.divittori@staff.univpm.it
10	Elisa Bellucci, e.bellucci@univpm.it
11	
12	*Corresponding authors
13	Roberto Papa, r.papa@univpm.it
14	Elena Bitocchi, e.bitocchi@univpm.it
15	

16 Key words: Crop diversification, crop adaptation, population genomics, adaptive17 introgression, Columbian Exchange

18 Abstract

19 In the context of the global challenge of climate change, mitigation strategies are needed to 20 adapt crops to novel environments. The main goal to address this is an understanding of 21 the genetic basis of crop adaptation to different agro-ecological conditions. The movement 22 of crops during the Colombian Exchange that started with the travels of Columbus in 1492 23 is an example of rapid adaptation to novel environments. Many diversification-related traits 24 have been characterised in multiple crop species, and association-mapping analyses have identified loci involved in these. Here, we present an overview of current knowledge 25 26 regarding the molecular basis related to the complex patterns of crop adaptation and 27 dissemination, particularly outside their centres of origin. Investigation of the genomic basis 28 of crop expansion offers a powerful contribution to the development of tools to identify and 29 exploit valuable genetic diversity and to improve and design novel resilient crop varieties.

30

31 Introduction

32 Environmental change will result in strong ecological and genetic effects on gene and allele 33 frequencies in many plant populations, as well as altering several aspects of agricultural 34 systems, such as plant physiology and phenology, water availability, soil fertility, pathogen 35 spread and host susceptibility. Many crops have evolved in response to climate change 36 under increasingly stressful conditions in which their extinction is highly possible. However, 37 there is more evidence for climate-driven range expansion than for range contraction [1]. 38 This suggests that plants can cope with climate change through adaptive mechanisms, such 39 as phenotypic plasticity and microevolution [2].

40 Genetic diversity represents the raw material on which adaptive selection acts, and 41 as such, it has a fundamental role in both evolutionary history and future evolutionary 42 pathways of a species [3]. Thus, persistent fluctuations in biotic and abiotic environmental 43 factors provide a background of changing selection pressures to which species must 44 respond, and in this way, genetic diversity is maintained within populations. Identification of 45 the molecular basis of plant adaptation is needed to drive plant breeding into the 46 development of novel varieties that can adapt to climate changes. Analysis of genetic 47 diversity through population genomics and genotype-phenotype association approaches 48 can be very useful tools to reach this aim [4], especially with the novel opportunities offered 49 by the more recent advances in genomics and DNA sequencing technologies. The success 50 of such studies critically depends on the type of plant material adopted. Moreover, if the 51 search for the signature of selection is the objective, the populations must have an ancient 52 and strong link with their growing environments, which must have variable agro-ecological 53 conditions.

54 Populations of wild plants and wild crop relatives can easily meet these prerequisites. 55 There are several examples in the literature that have focused on wild germplasm to detect 56 adaptive genetic control, along with studies on model species such as Arabidopsis thaliana [5•], with other examples available for crop species. Fustier et al. [6] investigated adaptation 57 58 in 11 populations of teosinte, the wild progenitor of maize, along two elevation gradients in 59 Mexico that showed continuous environmental changes over a short geographic scale. They 60 evaluated 1,664 individuals for 18 phenotypic traits and genotyped them for 38 microsatellite 61 markers and 171 outlier single nucleotide polymorphisms (SNPs). These significantly 62 differentiated between lowland and highland populations and/or correlated with 63 environmental variables. They showed that >50% of the traits were differentiated due to 64 local selection. A recent landscape genomics study of Rodriguez et al. [7•] reported on an

65 analysis of correlations between molecular markers and ecological variables at a continental 66 scale. They analysed a sample of 310 wild common bean georeferenced accessions that they genetically characterised at 131 SNPs. Geographic and environmental data were 67 68 combined with genetic diversity data to separate the effects of geography from those of 69 ecology, and they reported a total of 26 loci (19.9%) that were putatively under selection for 70 adaptation. Among these, different loci were shown to have compatible functions with 71 adaptation features, such as chilling susceptibility, cold acclimation, and mechanisms 72 related to drought stress [7•]. Recently, Mier y Teran et al. [8] characterised 112 wild 73 common bean accessions that were representative of the geographic distribution of the 74 Mesoamerican gene pool. This was applied at the molecular level (11,447 SNP markers) 75 and the phenotypic level (root trait evaluation, comparison of control and drought stress), 76 and considered environmental variables from the geographic coordinates of the origin of 77 each accession. They defined genomic regions that were associated with productivity and 78 drought adaptation in the wild germplasm.

79 Within the cultivated gene pools, the above-mentioned prerequisites for such studies 80 are satisfied only by populations of landraces (or, if available, by experimental populations, 81 as composite crosses specifically developed over multiple generations of experimental 82 evolution). Landraces offer unique opportunities for integration of association mapping and 83 signatures of selection analyses. Indeed, landraces are the product of an evolutionary interaction with the agro-ecosystem, and consequently, their genetic composition is 84 85 determined by both stochastic and human-mediated or natural selection over decades of evolution, which means that they have maintained a considerable amount of genetic 86 87 variability. Moreover, when multiple landrace populations grown in contrasting 88 agroecological environments are compared, it is possible to tag the signatures of divergent 89 selection [9-12]. This makes it possible to investigate the genes that are responsible for the 90 'genomic architecture' of the local adaptation of plants. After domestication, food crops 91 spread widely between different geographic and cultural areas at different levels and to 92 different extents, and this process ultimately contributes to the diversification of local 93 agricultural subsistence.

Among cereals, barley and maize are examples of crops that have achieved adaptive success worldwide (**Figure 1**). Barley is one of the primary plants that originated and was domesticated in the 'Fertile Crescent' about 11,000 years ago, and was later disseminated worldwide over a wide range of agro-climatic conditions [13]. Some of these conditions were particularly extreme, such as in Tibet, Nepal, Ethiopia and the Andes, where farmers

99 cultivated barley on mountain slopes at altitudes higher than those for any other cereals [14]. 100 Maize also has one of the broadest worldwide dissemination ranges. It was domesticated 101 once in the Balsas region in the valley of Mexico about 9,000 years ago, and it subsequently 102 spread to geographically and ecologically diverse environments, from Canada to Chile [15]. 103 Similarly, among legumes, the common bean can be considered as a crop that is now 104 successfully widespread [16]. The post-domestication phase of crops outside their centres 105 of origin (i.e., at regional, continental, worldwide levels) towards a wide range of agro-106 ecosystems has led to phenotypic and genetic divergence between domesticated forms. 107 This process can be considered a fascinating model for the study of the adaptive evolution 108 of crops, and it offers the possibility to discover new interesting genetic variants that have 109 potential use in a climate alarm context, like that which we are currently in.

110

111 **Diversification traits**

112 Meyer and Purugganan [17•] reported on several observed traits in crops that accompanied 113 their domestication and diversification, and their improvement phase. It is not particularly 114 easy to clearly distinguish between genes that underlie domestication and those that control 115 diversification traits. This is the case even if the genetic basis of adaptation might be more 116 related to diversification traits that are related to post-domestication stages, such as for 117 pigments, variations in size and chemical composition of edible parts, changes to the mating system (promoting allogamy or autogamy), resistance or tolerance to abiotic and biotic 118 119 stresses, reduced vernalisation and photoperiod sensitivity, and changes to flowering time, 120 the life cycle and dwarfism [17•]. It is important to consider that these traits can vary among 121 crop species, considering also that they relate to crops that have adapted to specific agro-122 ecological conditions and cultures. In this regard, several examples can be found in the 123 literature where the function of genes defined as putatively under selection during 124 domestication of crops can be ascribed to diversification traits, thus traits upon which both 125 natural and human selection have acted during crop expansion.

126 In common bean (Phaseolus vulgaris), Bitocchi et al. [18] compared selection 127 analysis data obtained for the same genes in different studies of varying sizes, data types 128 and methodologies. To study the effects of domestication at the genome level, they analysed 129 nucleotide diversity at 49 gene fragments on a sample of 39 wild and domesticated 130 Mesoamerican accessions of *P. vulgaris*. By applying population genomics approaches, 131 they identified several genes that showed footprints of selection. At the same time, they 132 used the SNP data of Rodriguez et al. [7•] to perform selection tests on a wider sample, 133 which included 417 and 160 wild and domesticated accessions, respectively, of common 134 bean. Finally, data were included from two further studies that focused on investigation of 135 the domestication process in common bean [19,20•]. The final comparison of the data from 136 these four studies provided independent evidence of selection for four genes: AN-Pv33, AN-137 DNAJ, Leg223, AN-Pv69. Gene-function investigations revealed that all of these genes are 138 involved in plant resistance/ tolerance to abiotic stresses, such as heat, drought and salinity. 139 In this regard, adaptation of plants to abiotic stresses is of crucial importance, because they 140 are among the major environmental factors that affect plant productivity. By accessing the 141 Arabidopsis thaliana stress-responsive gene database (http://srgdb.bicpu.edu.in/) [21], we 142 have identified a list of genes that were detected as functionally involved in abiotic stress 143 responses in Arabidopsis. The OrthoFinder algorithm [22] and the 2.1 version of the P. 144 vulgaris reference genome (https://phytozome.jgi.doe.gov/pz/portal.html) were then used to

145 identify orthologous genes in common bean. A total of 770 common bean genes were found 146 to be orthologous to Arabidopsis genes involved in abiotic stress responses (Table S1), 126 147 of which showed signature of selection during domestication in Bellucci et al. [20] and/or 148 Schmutz et al. [19] (Figure 2; Table S1). Among these genes, a very interesting candidate 149 is a homologue of K⁺ uptake transporter6 (KUP6). The KUP6 gene has been shown to be a 150 key factor in osmotic adjustment, through the balancing of potassium homeostasis in cell 151 growth and drought stress responses in A. thaliana [23]. Its function is directly mediated by 152 abscisic acid signalling, and under water-deficit stress this involves inhibition of cell 153 expansion in both roots and guard cells, which is driven by decreased turgor. In Bellucci et 154 al. [20], KUP6 was also among the small fraction of outlier genes for which selection has 155 increased the nucleotide diversity in the domesticated pool compared to the wild pool, which 156 suggests selection due to crop expansion into the new environments with unexpected biotic 157 and abiotic stresses (i.e., diversifying selection).

158 Meyer et al. [24] reported an example of geographic and environmental divergent 159 adaptation between four populations of African rice (Oryza glaberrima Steud.). They 160 sequenced the genomes of 93 landraces that spanned from west to central sub-Saharan 161 Africa, to investigate the African rice post-domestication spread, its subsequent adaptation 162 to local environments, and the genes that were involved in these processes. They focused 163 on salinity tolerance, as one of the major traits associated with geographic adaptation of 164 African rice. The accessions were phenotyped for various salinity-associated fitness traits, 165 and they found a significative loss of salinity tolerance in the southwest inland population. 166 This adaptive phenotype was thus suggested to arise from the costs of maintaining tolerance 167 in a geographic area associated with greater precipitation and decreased soil salinity. In 168 contrast, no significant differences were seen for the northwest, northeast and southeast 169 subpopulations. Genome-wide association studies allowed them to identify 11 loci that 170 contained several genes that were putatively involved in salt-stress tolerance. Among the 171 most significant of these was an orthologue of the O. sativa HAK5 gene. HAK5 has been 172 shown to be a key component in the acquisition and transport of potassium, to improve salt 173 resistance in potassium-deficient rice plants [25], and it might have had a crucial role in O. 174 glaberrima adaptation along the western Atlantic coast.

The timing of important developmental stages (e.g., flowering time) is another main diversification trait that defines adaptation of plant populations to different environments. In maize, numerous studies have focused on identification of the genetic control of flowering time and on genetic variations at identified genes in different materials from diverse 179 environments. Buckler et al. [26•] used a nested association mapping population of 200 180 recombinant inbred lines from 25 crosses, which resulted in a total of 5,000 lines for 181 identification of genes or genomic regions associated with flowering time. These lines were 182 phenotyped in four locations over 2 years. Quantitative trait locus (QTL) mapping showed 183 that the differences in flowering time were not caused by a few genes that had large effects, 184 but rather by the cumulative effects of numerous QTLs (i.e., <100), each of which had only 185 a small impact on this trait [26•]. However, to date, although a large body of mapping 186 information on the QTLs that control maize flowering time is available [27,28], the molecular 187 basis of these QTLs remains almost totally unknown, with the exception of four genes that 188 have been demonstrated to be involved in flowering time: Dwarf8 [29]; ZmCCT [30-32]; Vgt1 189 [33]; and ZCN8 [34].

190 The Dwarf8 gene has been shown to be an orthologue of the gibberellic acid 191 insensitive (GAI) gene, which is a transcription factor that negatively regulates gibberellin responses in A. thaliana. Association analysis has identified several interesting 192 193 polymorphisms in maize *Dwarf8*. One of these is a 6-bp deletion in the C-terminal region of 194 the open reading frame, and this showed strong association with flowering time [29]. With 195 the aim to evaluate the contribution of *Dwarf8* to maize adaptation to temperate climates, 196 Camus-Kulandaivelu et al. [35] analysed a wide collection of traditional landraces (144 from 197 America, 131 from Europe) for indel polymorphisms in the Dwarf8 gene. They reported a 198 variation in the frequency of the *Dwarf8* deletion associated with altitude and latitude, which 199 demonstrated that these features have an important role in driving local maize adaptation 200 [15,35]. In particular, for American landraces, they showed that the frequency of the Dwarf8 201 deletion was higher in northern Flint maize (83%) compared to maize groups from the 202 tropical Caribbean (2%) and Mexican (4%). Instead, the Andean group that was represented 203 by populations that originated from high altitudes (on average, 2,200 m a.s.l.) showed a 204 frequency of Dwarf8 deletion of 58%. Similarly, in Europe, Dwarf8 deletion has prevailed in 205 landraces from northern Europe.

Vgt1 is also one of the major maize flowering-time QTLs, and a miniature transposon that is located ~70 kb upstream of *ZMRap2.7* was shown to be the causative variant of *Vgt1* that contributes to maize adaptation to temperate regions [33,36,37]. Ducrocq et al. [36] carried out an association mapping study on 375 maize inbred lines, which included inbred lines representative of the American and European diversity, with a wide range of flowering times. They reported that the *Vgt1* early allele showed higher frequency in the tropical materials. Moreover, the frequency of *Vgt1* alleles among the tropical populations varied with the altitude of the collection site, while the early allele was rare at low altitudes. These data support the hypothesis that adaptive selection followed domestication of maize, with early and late materials adapted to high altitude and low altitude cultivation systems, respectively.

217 Yang et al. [31] showed that a CACTA-like transposon insertion within the *ZmCCT10* 218 promoter repressed ZmCCT10 expression, which makes maize insensitive to long days. 219 Likewise, Huang et al. [32] identified a Harbinger-like transposable element at ~57 kb 220 upstream of *ZmCCT9* that functions as a *cis*-acting repressor of *ZmCCT9*, to enhance maize 221 adaptation to higher latitudes. Comparisons of the gene sequence from teosinte and tropical 222 and temperate maize revealed that both the adaptive insertions were completely absent in 223 teosinte, and so they are likely to be *de-novo* mutations that occurred after the initial maize 224 domestication [30-32].

Recently, Guo et al. [34] reported that two natural *cis*-variants in the promoter of *ZCN8* were gradually targeted by selection during the spread of maize from its tropical origin to northern North America, which led to earlier flowering plants that were adapted to the temperate growing regions. In more detail, *ZCN8* was proposed to be homologous to *A*. *thaliana* FLOWERING LOCUS (FT), and they considered it to be the maize florigen gene [38,39].

231 Another interesting example was the study of Vigouroux et al. [40••] on pearl millet. 232 They analysed a total of 192 landraces that had been collected during two different periods 233 (i.e., 1976, 2003) throughout Niger, in the Sahel, which is one of the driest agro-ecosystems 234 in Africa. This geographic area had undergone recurrent drought during this interval of 25 235 years. Along with the analysis of the phenological and morphological changes in the two 236 samples evaluated in field experiments, they also investigated the genetic diversity across 237 these two samples. In particular, they analysed the change in allele frequency at the PHYC 238 flowering time locus [41], and showed that the allele that conferred earliness increased from 239 9.9% to 18.3% over this time frame. This study is an example of the strong adaptation of 240 plants to changing environmental conditions even over relatively short evolutionary 241 timescales. It also suggested that exploitation of genetic variability within landrace 242 populations represents a strategy in response to future climate changes. However, they 243 recommended the consideration of the mating system of the crop species, as they indicated 244 that this strategy might be successful for allogamous species, such as pearl millet, but that 245 further studies would be needed for autogamous species [40...].

246 SNPs are the markers of choice in different population genomics studies because 247 they are the most abundant bi-allelic and co-dominant markers that are characterised by 248 simple mutational patterns and by high-throughput and low-cost detection. Despite this, 249 many other examples exist in literature that are based on structural variations, which refers 250 to genomic changes in DNA segments of >1 kbp, such as insertions, deletions, inversions, 251 or copy-number variations. It is highly possible that genes responsible for acclimatizing and 252 adaptation to different agro-ecological conditions and stress resistances will be identified in 253 such genomic changes [42..]. As an example, Zhou et al. [43] reported the duplication and 254 evolutionary history of the COR15 gene that is involved in cold-stress defence, which was 255 previously detected in two copies in several species of Brassicaceae. They cloned the 256 homologous COR15 sequences of 10 species of Brassicaceae, and when they performed 257 evolutionary analyses they found significant inter-lineage differences in the evolutionary 258 rates between the original and the duplicated genes. The most interesting data were 259 perhaps observed for the analysis of the COR15 genes of the Draba species, which contrary 260 to the other lineages, is mainly present in cold-temperature, highly arid regions. Three 261 important lines of evidence were observed: (i) the estimated non-synonymous and 262 synonymous substitution ratio appeared to be higher among the duplicated genes; (ii) 263 positive selection was detected for the duplicate COR15 gene; and (iii) functional divergence 264 was shown between the two groups of the proteins. Overall, these observations indicated 265 that the functional differences in the *Draba* lineage between *COR15a*, as the original gene, 266 and COR15b, as the duplicated gene, have been driven by adaptive evolution. This allowed 267 its spread to cold locations during the Quaternary climatic oscillations, and subsequently its 268 expansion to arid alpine and arctic regions. Similarly, De Bolt [44] examined whether 269 Arabidopsis plants grown under different temperatures for several generations showed any 270 differences in copy number variations relative to the control situation of growth under normal 271 conditions. They showed that high temperatures promoted chromosomal segmental 272 duplications.

273 Recent studies have also suggested that polyploids might have greater phenotypic 274 flexibility for gene expression in response to environmental differences [45]. Ceccarelli et al. 275 [46] showed that chromosome endoreduplication in *Sorghum bicolor* is a fundamental part 276 of the adaptive response of plant genomes to salt stress. Their results showed that when 277 exposed to salt-induced treatments, only competent genotypes underwent endopolyploidy 278 of the root cortex cells, which allowed them to grow under sublethal salinity concentrations. 279 The wide variability obtained as a result of polyploidy events was thus directly correlated with the tolerance increase of *S. bicolor* to salinity, which highlighted the important role of this mechanism in adaptive responses to different abiotic conditions. Similarly, Saleh et al. [47] reported that citrus tetraploid rootstock is more tolerant to salt stress than their corresponding diploid.

284

285 Selection for adaptation

286 Local adaptation occurs when populations that grow under heterogeneous environmental 287 conditions evolve different phenotypic traits that provide a fitness advantage in their specific 288 environment [48]. Selection acts on sequence variation, which can derive from the standing 289 variation that has a long history of segregation within a crop before the advent of selection, 290 or *de-novo* mutations that originate in populations (i.e., wild forms or landraces), or from 291 hybridisation. Knowing the sources of variation on which selection for adaptation can act is 292 important for several reasons, such as, for example, to understand how rapidly populations 293 can adapt [3]. Exhaustive evidence that shows the relative role of standing variation or de-294 novo mutations after changes in the environment is still lacking. Adaptation is likely to be 295 slower if selection acts on *de-novo* mutations, compared to what would be expected when 296 it acts on standing variation, where beneficial alleles might already be available at higher 297 frequencies [49]. Moreover, on average, adaptation from standing variation appears to occur 298 through the fixing of more alleles with small effects [3,50], and can have greater potential for 299 adaptation if the rate of environmental change is fast, rather than slow, by traversing larger 300 distances in the phenotype space.

301 Along with useful standing variation and *de-novo* mutations, selection for adaptation 302 can also act on new genotypic variations due to recombination after hybridisation [51,52]. In 303 common bean, Bellucci et al. [20•] analysed RNA sequencing data from a set of 304 Mesoamerican wild and domesticated accessions, and they showed that most of the genes 305 detected as under selection during domestication showed reduced diversity in their 306 domesticated compared to their wild forms, as expected under positive selection from 307 standing variation. However, 2.8% of the outlier genes showed no diversity in the wild form, 308 and polymorphism in the domesticated form. This thus suggested that in some cases the 309 selection increased the nucleotide diversity of domesticated materials at target loci, the 310 function of which was associated with adaptation traits, such as abiotic stress responses 311 and flowering time [20•]. Interestingly, in the same species, Bitocchi et al. [18] analysed 312 nucleotide data of 49 gene fragments in a sample of Mesoamerican wild and domesticated 313 accessions, and they detected an excess of nonsynonymous mutations in the domesticated 314 forms, particularly in the coding regions, compared to the non-coding regions. These 315 mutations appeared to be recently derived mutations, and the investigations into the 316 functions of their relative genes (responses to biotic and abiotic stresses) support a scenario 317 where new functional mutations were selected for adaptation during diversification.

318 In maize, Guo et al. [34] asked whether the ZCN8 gene can affect natural variations 319 in flowering time. They performed association analysis by sequencing ZCN8 and its 320 upstream and downstream regions in segregant populations derived from a cross between 321 W22, a temperate Zea mays ssp. mays inbred line, and 8759, a Z. mays ssp. parviglumis 322 accession. They found a SNP in the promoter region of ZCN8 (i.e., SNP-1245) that 323 coincided precisely with the allelic differences in flowering time between all of the parents of 324 the teosinte-maize populations used in their study. They also sequenced the ZCN8 gene in 325 a panel of 513 maize inbred lines and 45 teosinte lines (including lines of Z. mays ssp. 326 parviglumis, the maize progenitor, and lines of its close relative species Z. mays ssp. 327 *mexicana*). These data revealed that the early flowering allele of SNP-1245 was present in 328 \sim 24% of the teosinte accessions, which suggested that this polymorphism was a standing 329 variant in the maize wild progenitor selected during the early domestication of maize. Guo 330 et al. [34] also detected a three-base-pair deletion variant (i.e., Indel-2339) about 1,000 331 bases from SNP-1245 that was associated with flowering time and showed higher 332 expression of ZCN8. Moreover, they did not find this allelic variant in the maize progenitor, 333 although it was present in Z. mays ssp. mexicana, from which gene flow resulted in its 334 introgression into maize [53]. Furthermore, low frequency of Indel-2339 (5%) was shown for 335 South America germplasm (i.e., tropical maize), while it was selected at a higher frequency 336 in northern United States accessions (30%; temperate maize). Overall, these data 337 suggested that two independent associated mutations (i.e., *cis*-regulatory variants) in the 338 promoter region of ZCN8 arose in a stepwise manner: SNP-1245 during the early domestication of maize, and subsequently Indel-2339 during maize diversification into the 339 340 Mexican highlands. The discovery that ZCN8 has more than one functional mutation that 341 segregates indicated that genes associated with crop domestication and diversification are 342 subject to recurrent mutations that might be selective targets at different times during 343 evolution.

344 Identification of adaptive introgression can be relatively easy when materials 345 collected at different times are available, such as with historical collections. A recent 346 example was seen by the study of Bitocchi et al. [11], where the effects were evaluated for 347 hybridisation of modern maize and landraces over a relatively short period of 50 years. 348 Bitocchi et al. [11] analysed and compared the genetic diversity of two samples of maize 349 landraces from central Italy that were collected at two different times: an old collection that 350 was carried out before the introduction of hybrid varieties, and a recent collection that had 351 evolved in co-existence with modern maize. Population structure analysis allowed the detection of introgression from modern maize. Coupled to the data of selection analyses (i.e., detection of outlier loci in comparisons between historical and recent maize collections), these data indicated that selection pressures for adaptation have favoured new alleles that were introduced by migration from hybrids over the last 50 years. These data showed the crucial role of migration in the evolution of landrace populations grown on farms.

358 The Columbian Exchange: adaptation of crops from American homelands into Europe

359 The introduction of New World crops into Europe after the Columbus 1492 voyage was one 360 of the most important evolutionary events related to agriculture, adaptation and biological 361 changes, and more generally, to human society. In 1972, the historian Alfred Crosby coined 362 the term 'Columbian Exchange', to designate the process of the biological diffusion triggered 363 by the colonisation of the Americas by Europe. The benefits of the New World crops have 364 resulted in their adoption in all parts of the world, which demonstrated that as the basis of 365 this process, the plants underwent significant adaptation to the various agro-ecological 366 conditions [54]. The growing knowledge about the adaptation of crops to new environments 367 through the study of their introduction and expansion into Europe (i.e., a historically well-368 defined event of recent introduction and rapid adaptation) will be of great use for future major 369 environmental and socio-economic changes, such as increases in temperature, variability 370 of rainfall, and new consumer preferences. Several crops were introduced into Europe from 371 the Americas (e.g., tomato, maize, beans, squash, potato, tobacco). This dissemination 372 process occurred during the same historical period for several species, and it was 373 characterised by diverse features (e.g., different mating systems and ploidy) that can be 374 exploited to investigate their effects on genome diversity and to highlight the genetic control 375 of adaptation. There are numerous studies in the literature for different crops that have 376 highlighted the changes that occurred in their genomes due to colonisation of new agro-377 ecosystems. Here, we present some examples for three crops that were involved in the 378 Columbian Exchange, and which have been among the most important: potato, maize and 379 common bean.

380 Following long debate during which most studies have suggested multiple 381 domestications for Solanum tuberosum L (potato), Spooner et al. [55] demonstrated the 382 monophyletic origin of cultivated potatoes through phylogenetic analysis and cladistic data. 383 These showed that landraces of potato originated in the Andes of southern Peru, and 384 subsequently became widespread throughout Chile, thus assuming the present-day 385 distributions of the original cultivars. Potato was not brought to Europe by Columbus or 386 others soon after the discovery of the New World in 1492; potato arrived later. The reason 387 for this is that potato is a cool temperate crop of the high Andes of South America and was 388 not discovered by the Spaniards until 1532 [56]. Potato cultivation in Europe spread rapidly, 389 and also reached locations with significative growth and climate differences. For potato, the 390 most important adaptation trait to European conditions – and a key event in its history – was 391 to overcome the short-day dependency for tuberisation, due to the equatorial origin of 392 potatoes [57-59...]. Indeed, when introduced into temperate zones, wild material forms 393 tubers only during the shorter autumnal day lengths. The gradual arrival of winter, which is 394 characterised by freezing temperatures, stops the correct maturation of the tubers, 395 consequently killing the plant. In the A. thaliana model system, the pathway that controls 396 flowering time is very complex, and the complexity of this regulation involves four intricate 397 networks of signalling pathways (i.e., photoperiod, vernalisation, autonomous, gibberellins) 398 [60] (Figure 3A). Among the proteins involved in this complex pathway, cycling Dof (DNA-399 binding with one finger) factors (CDFs) are a group of plant-specific transcription factors that 400 repress flowering by down-regulation of the expression of the CONSTANS (CO) gene, a 401 central regulator of the photoperiod pathway [61]. In potato, the plant maturity phenotype 402 has been reported as a major effect QTL that maps to chromosome 5, and this phenotype 403 is a measure of several important secondary traits. These include development of the 404 canopy, vegetative growth, onset of tuberisation, leaf senescence, life-cycle length and 405 pathogen resistance [62]. Kloosterman et al. [57] used ultra-dense amplified fragment length 406 polymorphism markers and two diploid segregant potato populations derived from crosses 407 between wild and domesticated genotypes. In this way, they narrowed down the locus 408 responsible for the plant maturity phenotype to a region of around 110 kb on chromosome 409 5. Screening for putative candidate genes, they identified the potato homologue of *CDF1* in 410 this QTL region (StCDF1, Solanum tuberosum CDF gene 1). They sequenced StCDF1 in 411 the progenies of the mapping populations, which allowed identification of three StCDF1 412 allelic variants: StCDF1.1, which was characteristic of short-day-dependent tuberisation 413 descendants, and two insertion variants, StCDF1.2 and StCDF1.3, that were typical of the 414 early maturing/ tuberising descendants. Kloosterman et al. [57] established that StCDF 415 conserves its repressive function on the two potato CONSTANS genes (StCO1/2) that 416 repress tuber formation during long days [63]. They also suggested that due to the loss of 417 their C-terminal end, the StCDF Andean variants (i.e., StCDF1.2, StCDF1.3) led to 418 accumulation of StCO1/2 repressors. This interaction indirectly induced expression of 419 StSP6A, the potato homologue of FLOWERING TIME (FT), which resulted in induction of 420 tuber development under long days (Figure 3B). The absence of post-translational 421 regulation of StCDF1.2 and StCDF1.3 allowed them to remain constant throughout the day, 422 which formed the basis of potato diversification at different latitudes. A recent investigation 423 explored haplotype diversity at the potato maturity locus StCDF1 using a panel of 58 424 samples [58]. These included South American wild species, South American landraces, and 425 North American cultivars derived from modern breeding programmes. Here, Hardigan et al.

426 [58] reported 55 haplotypes for StCDF1 that encoded 27 peptide variants. Four haplotype 427 groups contained conserved deletions that affected the structure of the StCDF1 peptide. 428 The DNA phylogeny of haplotypes at the StCDF1 locus revealed that almost all long-day 429 landraces/ cultivars contained alleles that encoded shortened StCDF1 proteins that were 430 derived from introgression from wild species. This suggested a key role for the extant natural 431 populations as essential sources of untapped adaptive potential. In the case of potato, 432 StCDF1 allele introgression from the wild species allowed potato cultivation in North 433 America, and, probably, also subsequently in Europe. A very interesting study that focused 434 on the origins and adaptation of European potatoes was carried out by Gutaker et al. [59..]. 435 The strength of their work was the investigation of historical samples that spanned 350 years 436 of potato evolution in Europe. Their materials included 29 historical herbarium specimens 437 that they obtained from different European museums, which included three Chilean and 26 438 European historical samples. They also analysed 43 South American modern samples, and 439 16 European modern samples. An array-based targeted re-sequencing approach was used 440 that allowed them to target the whole chloroplast genome and ~4.3 Mb of the nuclear 441 genome, including StCDF1 [57]. Analysis of these genetic data initially allowed Gutaker et 442 al. [59••] to highlight the very complex scenario related to the introduction and wide spread 443 of potatoes in Europe. These data indicated that the oldest European materials (i.e., 444 collected between 1650 and 1750) derived from an ancestor of the Andean landraces, while 445 in the subsequent 100 years there was introgression from newly introduced Chilean 446 potatoes. The scenario is more complex considering that twentieth century European 447 potatoes did not descend from their nineteenth century admixed predecessors, but are the 448 result of introgression from wild potato species, as they were used in twentieth-century 449 breeding programmes to introduce pathogen resistance [64]. It is also interesting that 450 Gutaker et al. [59..] highlighted the re-introduction of European potatoes into America, and 451 that this impacted upon the Andean and Chilean potato diversity; indeed, European ancestry 452 was detected in potatoes in the South American modern-day sample. Gutaker et al. [59••] 453 also investigated the origins of the long-day adaptive alleles in the StCDF1 gene. They 454 reported the appearance of StCDF1.2 and StCDF1.3 adaptive alleles in Europe starting from 455 1810 only, with none of these insertion variants present in the oldest European samples of 456 Andean descent (1650-1750), nor in the Andean landraces. For this reason, they excluded 457 (with high confidence) the possibility that adaptation to long-day tuberisation had arisen from 458 the Andean landraces standing variations. They showed the appearance of the adaptive 459 alleles in Europe in correspondence with admixture with the newly introduced Chilean

460 potatoes. However, there was no evidence of direct correlations between the adaptive 461 variants and the historical samples from the lowlands of Chile. Gutaker et al. [59••] thus 462 hypothesised that the adaptive insertions in the *StCDF1* gene originated *de novo* in Europe, 463 and then became rapidly fixed due to their dominant inheritance and breeding advantage. 464 However, they also stated that this hypothesis needs to be further confirmed, as their 465 sampling of historical Chilean specimens is not particularly representative, and thus it did 466 not allow clear rejection of the possibility of a Chilean origin of these adaptive insertions.

467 Another very important crop that became widespread in Europe during the Columbian 468 exchange was *P. vulgaris* (common bean). This species originated in Mesoamerica, and 469 wild forms became widespread by subsequent migration into South America; domestication 470 took place independently in two geographically distant areas, Mesoamerica and the Andes, 471 which represented the two main gene pools of the species [16]. The Mesoamerican common 472 bean appears to have arrived in Europe through Spain and Portugal in 1506, following the 473 first voyage of Columbus; then in 1528, the exploration of Peru by Pizarro opened the 474 possibility of the introduction of the Andean common bean. P. vulgaris spread into the Old 475 World over a very short time, and many common bean landraces rapidly evolved in Europe 476 as a result of its adaptation to new agro-ecological growth conditions. The dissemination of 477 common bean into and across Europe followed very complex pathways, which involved 478 different introductions from the Americas, and at the same time, direct exchanges among 479 countries within Europe, and between European and other Mediterranean countries [16]. To 480 investigate the evolutionary patterns of the common bean far from the Americas, Angioi et 481 al. [54] analysed a wide sample of P. vulgaris accessions, as 94 from the Americas, and 307 482 from Europe. They included chloroplast simple sequence repeats (SSRs), and nuclear data 483 (i.e., phaseolins, three indel-spanning markers of the PvSHATTERPROOF1, PvSHP1, 484 gene) and morphological data (i.e., coat pattern, seed size, colour and shape). In this way, 485 Angioi et al. [65] showed that both the Mesoamerican and Andean gene pools were present 486 in Europe and that the European germplasm was more prevalent as the Andean origin 487 (67%). The trend was maintained at a smaller scale (i.e., a country level), whereby the 488 Mesoamerican proportion was higher in the eastern parts of Europe, with a maximum of 489 46% in Greece, while the Andean type was most frequently found in three European macro 490 areas: the Iberian Peninsula, Italy and central-northern Europe. Interestingly, and contrary 491 to expectations, the European common bean did not show any strong reduction in genetic 492 diversity due to the introduction bottlenecks and selection for adaptation to these new agro 493 ecosystems and consumer preferences; indeed, Angioi et al. [54] and previous studies have

494 shown very low reductions in diversity in common bean from Europe. These findings 495 indicated a high level of gene flow among the different European geographic regions. 496 Furthermore, they highlighted the role of the breakdown of the spatial isolation between the 497 Mesoamerican and Andean accessions in Europe, with promotion of hybridisation, which 498 had a significant impact on the maintenance of genetic diversity. By combining these 499 chloroplast and nuclear data, they were able to identify hybridisation events, and they 500 estimated that 44.2% of the European landraces derived from at least one hybridisation 501 event between the Mesoamerican and Andean forms. Gioia et al. [66] complemented the 502 dataset of Angioi et al. [65] with nuclear SSRs, and analysed a set of 89 American and 256 503 European landraces. Gioia et al. [66] combined the data from the recombination of the gene-504 pool-specific chloroplast SSRs, phaseolin and PvSHP1 markers and the Bayesian 505 assignments and admixture analysis based on nuclear SSRs, through which they were able 506 to identify hybrids and distinguish them as 'pure' Mesoamerican and Andean genotypes. 507 Novel combinations of genes/ genomic regions thus arose in Europe after the common bean 508 introduction and during its dissemination, on which adaptive selection acted (i.e., adaptive 509 introgression). The new '-omics' technologies can help to fine-tune the molecular basis of 510 these adaptations of the common bean in Europe, an aspect that is ongoing in the 511 BEAN ADAPT project (funded through the 2nd ERA-CAPS call, ERA-NET for Coordinating) 512 Action in Plant Sciences). This project is based on a multidisciplinary approach (i.e., 513 genomics, population/ quantitative genetics, biochemistry, plant physiology), with the aim 514 being to dissect out the genetic basis and phenotypic consequences of the adaptation of P. 515 vulgaris and its sister species P. coccineus from their centres of origin in the Americas into 516 Europe and the new European agro-ecological environments.

517 Maize is probably the most important New World crop that was involved in the 518 Columbian Exchange. Rebourg et al. [67] characterised a set of 131 European maize 519 landraces according to morphological and genetic data (i.e., restriction fragment length 520 polymorphism), and classified them into genetic groups that showed clear differentiation 521 according to latitude. Six main European races were detected based on morphological and 522 genetic differences: 'German flint', which included landraces mainly grown in Germany or 523 the Alsace; 'north-eastern European flint', which included landraces mainly from France, 524 and also Spain, Portugal and several eastern European countries; 'southern European flint', 525 which was characterised by landraces from various countries which were mainly in southern 526 Europe; 'Italian orange flint', as Italian landraces, with some others from southern Spain; 527 'Czechoslovakian type'; and 'Pyrenees-Galicia flint', which was characterised by two 528 homogeneous subgroups, as the landraces from Galicia, and those from the Pyrenees and 529 other regions of France. Then Rebourg et al. [68] included genetic data of 88 American 530 landraces that were representative of the main American races in their previous dataset [56], 531 to infer the genetic relationships among American and European maize populations. They 532 showed signatures for the introduction of a bottleneck (European landraces retained overall 533 75% of the genetic diversity of those from America), and identified various types of American 534 maize that were introduced into Europe at different times or in different places, which gave rise to distinctive European races [69]. Beyond confirming the importance of Caribbean 535 536 germplasm, which was the first maize type to be introduced into Europe, they highlighted 537 the close relationship between southern Spain and Caribbean populations, whereby the data 538 revealed that introductions of North American flint populations had a key role in the 539 adaptation of maize to the European climate. In particular, the data supported the hypothesis 540 that present-day northern and eastern European flint germplasm was directly derived from 541 North American flint populations. Northern flint populations were relatively insensitive to day 542 length, and they had low temperature requirements for flowering. Earliness was a key factor 543 for adaptation to the more temperate climates. Brandenburg et al. [70•] sequenced 67 544 genomes from both continents that covered 11 major groups, as representative of all of the 545 American and European diversity. They used several population genomics and association 546 mapping approaches to trace the origins of the European maize, and to investigate its 547 demographic and selective history. One of the main outcomes of this study was the detection 548 of admixture in the European maize materials. In particular, they reported the admixed 549 origins of the Italian flints from two contributions, the European flint and the southern 550 European populations. This excluded the possibility of a third independent introduction, as 551 had previously been suggested by Rebourg et al. [68], and instead emphasised the pivotal 552 role of admixture in environmental maize adaptation. Moreover, the data of Brandenburg et 553 al. [70•] highlighted the admixed origins of the European flints from the northern European 554 flints and the tropical landraces. Interestingly, they also investigated the footprints of 555 selection for adaptation to a wide range of climatic and ecological conditions, and they 556 showed that numerous genes/ gene networks were involved in flowering time, drought and 557 cold tolerance, and in plant defence and starch properties. An example of the candidate 558 genes for adaptation that were detected by associations between latitude and allele 559 frequency was defined at GRMZM2G095955, a gene that is located in the vicinity of the 560 maize floral activator, ZCN8 [39]. They reported that in the ZCN8 region there was a 561 haplotype that was common to all temperate materials, and they showed segregation of this

562 'temperate' haplotype with a 'tropical' haplotype within the tropics, and to a lesser extent 563 within the corn belt dents. Along with the previously characterised genes, they also revealed 564 new candidates, including ZCN5 (also known as zen1 and pebp5), a gene from the same 565 family as ZCN8 that was recently reported to be associated with flowering time variations 566 [71]. They also defined genes associated with plant responses to biotic and abiotic stresses, 567 such as the ZmASR2 gene (abscisic acid-, stress-, and ripening induced protein 2), which 568 was shown to have increased expression at the transcript and protein level under water-569 deficit conditions [72], and the TPS23 gene that is involved in the control of the synthesis of 570 a volatile sesquiterpene that attracts natural enemies of herbivores upon release [73]. 571

572 Conclusions

573 Deeper understanding of the evolutionary processes and complex genetics mechanisms 574 that form the basis of adaptation of plants to different environmental conditions is a very 575 ambitious goal for evolutionary biologists, breeders and geneticists. It also has strong 576 implications for overcoming the current challenges that agriculture has to face, such as to 577 guarantee food security and quality, to adapt crops to marked variations in climate, and to 578 protect and improve the environment. In this context, the identification of the genetic 579 architecture both at genotype and population level, that contribute to adaptive changes, can 580 strongly influence breeding targets and strategies. The potential applications are nearly 581 infinite for the constitution of novel varieties in breeding programmes, but it will be crucial 582 also for biodiversity conservation, to provide help in the implementation of the appropriate 583 strategies. We have now in-hand novel tools and approaches that allow us to face this 584 challenge through exploiting the unprecedented experimental power available. These 585 include:

(i) Particularly advanced techniques that offer unique opportunities to scan a genome, not
only to obtain genotypic information, but also to analyse the molecular phenotype of the
whole genome, through analysis of the transcriptome, the metabolome, and the proteome
[20, 74-76].

590 (ii) We can count on improvements to the tools and approaches to analyse these data, which 591 have also evolved to catch the complexity of these biological processes. Population 592 genomics approaches allow the identification of candidate loci for adaptation using 593 genotypic data without any prior information about phenotypes. Along with classical 594 approaches aimed at detection of 'selective sweeps' [77], new methods and integrated 595 approaches can be applied that take into account the concept that genes do not often 596 actually operate as sole effectors, as they have roles in complex interactive systems, or 597 gene networks, that ultimately lead to a phenotype [78]. As an example of the impact that gene interactions can have on the determination of the phenotype, an A. thaliana genome-598 599 wide association analysis reported that for root length, epistatic effects can be so strong that 600 they overcome the additive genetic variance [79•]. In soybean, Fang et al. [80] carried out a 601 comprehensive genome-wide association studies that enabled identification of the 602 underlying genetic loci, loci interaction, and genetic networks across important traits.

603 (iii) Multidisciplinary approaches can be applied and integrated to decipher the complexity 604 of the genetic basis of adaptation. These can combine evidence from the signatures of 605 selection analyses with association mapping to increase the power for the detection of 606 regions that influence complex traits, while also reducing the number of false-positive signals 607 [81,82]. Moreover, recently, different approaches have been developed based on the use of 608 environmental variables that are treated as quantitative traits, and their association with 609 molecular traits can be exploited as a tool to identify the loci that underlie local adaptation 610 [12,83]. Similarly, network analyses can be used to investigate the roles of interactions 611 between genes in local adaptation [84], using information on linkage disequilibrium shared 612 between genome-wide multiple loci to perform linkage disequilibrium network analyses.

613 (iv) Landrace populations of crops are the 'perfect' model to apply all of these approaches 614 to investigate adaptation features in the plant genome. They also allow the possibility to 615 compare the effects of the same evolutionary process on the genome when this occurs as the following: independently on different populations of the same species (e.g., 616 617 domestication in common bean occurred independently in Mesoamerica and the Andes) 618 [18,19]; among different crop species within the same genus (e.g., different domesticated 619 *Phaseolus* species) [16]; and/ or among species of different genera (i.e., shattering trait in 620 cereals) [85] that are characterised by different features (e.g., diverse mating systems, 621 diverse ploidy levels). These aspects offer great opportunities to go deeply into the 622 molecular and developmental mechanisms at the basis of adaptation.

623 In this scenario, the Columbian Exchange represents a pivotal model. It offers a great 624 opportunity to exploit all of these available tools and approaches, along with the plant genetic 625 resources, to finally dissect out the genetic basis and phenotypic consequences of plant 626 adaptation to new environments. This can now come through the study of their introduction 627 from their respective centres of domestication in the Americas, and their expansion through 628 Europe as a recent and historically well-defined event of rapid adaptation. Numerous crop 629 species have been protagonists of these processes and have experienced adaptation in a 630 relatively short period of time in the same geographic range (i.e., with the same 631 environmental changes). What we need to do now is to investigate this process more deeply 632 in different crops, and to compare and integrate the information obtained. A better 633 understanding of variation in landscape structure across species and environments is also 634 necessary to understand and predict how populations will adapt [86]. Moreover, advances 635 in statistics and increased computing power already provide the possibility to develop 636 predictive approaches, as demonstrated by Exposito-Alonso et al. [87••] who were able to

build genome-wide environmental selection models to predict how evolutionary pressures
on species will work in inaccessible environments, or even under future hypothetical
climates.

640

641 Figure legends

Figure 1. Geographic distribution of barley (top), maize (middle) and common bean (bottom) landraces from their centres of domestication.

The centres of domestication are represented by white dots with black borders. The distributions of the landraces/ traditional cultivar accessions were obtained by plotting the geographic coordinates for where the seeds were collected. Data were extracted from the database of the Genesys platform (<u>https://www.genesys-pgr.org/</u>), which includes information from several genebanks.

649

Figure 2. Physical map of the 11 common bean chromosomes and genomic locations of genes putatively involved in abiotic stress responses and with selection signatures in common bean.

653 Common bean genes were identified based on orthology with those involved in abiotic stress 654 responses in Arabidopsis thaliana, according to The Arabidopsis Stress Responsive Gene 655 Database [21] and using the OrthoFinder algorithm [22]. The orthologous protein to the A. 656 thaliana KUP6 is also shown in chromosome Pv03. For the map representation, we selected 657 a subset of 126 common bean orthologues (see Table S1 for the full list) that show selection 658 signatures according to Schmutz et al. [19] and/or Bellucci et al. [20]. Genes potentially 659 associated to different stress responses based on the orthology with A. thaliana genes are 660 highlighted according to the legend. The physical distances in the scale are reported in 661 megabases (Mb).

662

663 Figure 3. Schematic representation of *CDF* gene function and interactions in the 664 photoperiod pathway.

During long days, in the *A. thailana* model system (A), the interaction between GIGANTEA
(*GI*) and FLAVIN-BINDING KELCH REPEAT F-BOX 1 protein (*FKF1*) induces degradation
of CYCLING DOF FACTOR (*CDF*), which is a repressor of CONSTANS (*CO*). *CO* promotes
flowering by initiating transcription of the FLOWERING TIME (*FT*) gene. In *S. tuberosum L.*(B), the *CDF* adaptive variant does not interact with the GI-FKF1 complex, which leads to
repression of *CO1*/2. In contrast to *A. thaliana*, *CO1*/2 act as repressors of *SP6A*, which is

- 671 the potato homologue of *FT*. Repression of *CO1/2* allows expression of *SP6A* and promotion
- 672 of potato tuberisation under long days, which forms the basis of potato diversification at
- 673 different latitudes. Arrow, promotion of gene expression; truncated arrow, repression of gene
- 674 expression; truncated dotted arrow, lack of repression due to pathway interruption.
- 675

676 Supplementary

677 Table S1. Genes putatively involved in abiotic stress responses in common bean.

Genes were identified based on the orthology relationships (OrthoFinder algorithm) [22] with *A. thaliana* genes listed in *The Arabidopsis Stress Responsive Gene Database* [21], and with the KUP6 protein sequence. The orthogroup (i.e., cluster of orthologous genes across *A. thaliana* and common bean), the type of stress response of the orthogroup genes, the common bean gene name, its genomic location and description based on Phytozome, and the presence of selection signatures [19;20] are reported for each common bean gene.

- 684 PN; putatively neutral gene.
- n.m; no homologous contigs from Bellucci et al. (2014) have been identified

* in the column K, according to the data of Bellucci et al. (2014), the PS (putative under
selection) contig that mapped on the gene showed polymorphisms for the nucleotide
sequence across domesticated accessions and was fixed monomorphic in the wild pool of
genotypes.











693 Figure 2



695 Figure 3

696 **References**

- Malther G.R., Post E., Convey P., Menzel A., Parmesan C., Beebee T.J.C., Fromentin
 J.M., Hoegh-Guldberg O., Bairlein F. (2002). Ecological responses to recent climate change. *Nature* 416:389–395. DOI: 10.1146/annurev-arplant-042817-040240.
- **2** Lorant A., Ross-Ibarra J., Tenaillon M.I (2018). Genomics of long- and short- term
 adaptation in maize and teosinte. *PeerJ Preprints* 6: e27190v1. DOI:
 10.7287/peerj.preprints.27190v1.
- **3** Barrett R.D.H. and Schluter D. (2008). Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23:38-44. DOI: 10.1016/j.tree.2007.09.008.
- 4 Mousavi-Derazmahalleh M., Bayer P.E., Hane J.K., Valliyodan B., Nguyen H.T., Nelson
 M.N., Erskine W., Varshney R.K., Papa R., Edwards D., (2019). Adapting legume crops to
 climate change using genomic approaches. *Plant, Cell and Environment* 42(1):6-9. DOI:
 10.1111/pce.13203.
- Fournier-Level A., Korte A., Cooper M.D., Nordborg M., Schmitt J., Wilczek A.M. (2011).
 A map of local adaptation in *Arabidopsis thaliana*. *Science* 334(6052):86-89. DOI: 10.1126/science.1209271.
- This study provides complementary analyses of local adaptation to climate in A. thaliana,
 with the combination of genome-wide SNPs, fitness estimates in the field and continentalscale climate data.
- 6 Fustier M.A., Martínez-Ainsworth N.E., Aguirre-Liguori J.A., Venon A., Corti H., Rousselet
 A., Dumas F., Dittberner H., Camarena M.G., Grimanelli D., Ovaskainen O., Falque M.,
 Moreau L., de Meaux J., Montes S., Eguiarte L.E., Vigouroux Y., Manicacci D., Tenaillon
 M.I. (2019). Common gardens in teosintes reveal the establishment of a syndrome of
 adaptation to altitude. *bioRxiv* 563585. DOI: 10.1101/563585.
- 720 7• Rodriguez M., Rau D., Bitocchi E., Bellucci E., Biagetti E., Carboni A., Gepts P., Nanni
 L., Papa R., Attene G. (2016). Landscape genetics, adaptive diversity, and population
 structure in *Phaseolus vulgaris*. *New Phytol*. 209:1781–1794. DOI: 10.1111/nph.13713. *This study represents one of the first examples of the application of landscape genomics to crops, to help in the understanding of adaptation on the scale of the natural landscape*.
- 8 Mier Y Teran B.J.C, Konzen E.R., Medina V., Palkovic A., Ariani A., Tsai, S.M., Gilbert
 M.E. (2018). Root and shoot variation in relation to potential intermittent drought adaptation

- 727 of Mesoamerican wild common bean (*Phaseolus vulgaris* L.). *UC Davis*. DOI: 728 10.1093/aob/mcy221.
- 9 Pusadee T., Jamjod S., Chiang Y.C., Rerkasem B., Schaal B.A. (2009). Genetic structure
 and isolation by distance in a landrace of Thai rice. *PNAS* 106(33):13880-13885. DOI:
 10.1073/pnas.0906720106.
- **10** Bellucci E., Bitocchi E., Rau D., Nanni L., Ferradini N., Giardini A., Rodriguez M., Attene
 G., Papa R. (2013). Population structure of barley landraces populations and geneflow with
 modern varieties. *PLoS ONE* 8(12): e83891. DOI: 10.1371/journal.pone.0083891.
- **11** Bitocchi E., Bellucci E., Rau D., Albertini E., Rodriguez M., Veronesi F., Attene G., Nanni
 L. (2015). European flint landraces grown *In-situ* reveal adaptive introgression from modern
 Maize. *PLoS ONE* 10(4): e0121381. DOI: 10.1371/journal.pone.0121381.
- 738 **12** Lasky J.R., Upadhyaya H.D., Ramu P., Deshpande S., Hash C.T., Bonnette J., Juenger
- 739 T.E., Hyma K., Acharya C., Mitchell S.E., Buckler E.S., Brenton Z., Kresovich S., Morris G.P.
- 740 (2015). Genome-environment associations in sorghum landraces predict adaptive traits.
- 741 *Science Advances* 1(6): e1400218. DOI: 10.1126/sciadv.1400218.
- 13 Lister D.L., Jones H., Hugo R., Oliveira H.R., Petrie C.A., Liu X., Cockram J., Kneale
 C.J., Kovaleva O., Jones M.K. (2018). Barley heads east: Genetic analyses reveal routes of
 spread through diverse Eurasian landscapes. *PLoS ONE* 13(7): e0196652. DOI:
 10.1371/journal.pone.0196652.
- 14 Tanto Hadado T., Rau D., Elena B., Papa R., (2009). Genetic diversity of barley
 (*Hordeum vulgare* L.) landraces from the central highlands of Ethiopia: comparison between
 the *Belg* and *Meher* growing seasons using morphological traits. *Genet Resour Crop Evol.*56(8):1131-1148. DOI: 10.1007/s10722-009-9437-z.
- 15 Manchanda N., Snodgrass S.J., Ross-Ibarra J., Hufford M.B. (2018). Evolution and
 adaptation in the maize genome. *The maize genome*, Compendium of Plant Genomes.
 Springer, Cham 319-332. DOI: 10.1007/978-3-319-97427-9 19.
- **16** Bitocchi E., Rau D., Bellucci E., Rodriguez M., Murgia M.L., Gioia T., Santo D., Nanni L.,
 Attene G., Papa R. (2017) Beans (*Phaseolus* ssp.) as a model for understanding crop
 evolution. *Front Plant Sci* 8,722. DOI: 10.3389/fpls.2017.00722.

17. Meyer R.S. and Purugganan M.D. (2013). Evolution of crop species: genetics of
domestication and diversification. *Nature Reviews Genetics* 14:840–852. DOI:
10.1038/nrg3605.

759 This review provides a detailed and comprehensive overview of the process of 760 domestication and the subsequent crop expansion.

18 Bitocchi E., Rau D., Benazzo A., Bellucci E., Goretti D., Biagetti E., Panziera A., Laidò
G., Rodriguez M., Gioia T., Attene G., McClean P., Lee R.K., Jackson S.A., Bertorelle G.,
Papa R. (2017) High level of nonsynonymous changes in common bean suggests that
selection under domestication increased functional diversity at target traits. *Front Plant Sci.*7:2005. DOI: 10.3389/fpls.2016.02005.

766 19 Schmutz J., McClean P.E., Mamidi S., Wu G.A., Cannon S.B., Grimwood J., Jenkins J., 767 Shu S., Song Q., Chavarro C., Torres-Torres M., Geffroy V., Moghaddam S.M., Gao D., 768 Abernathy B., Barry K., Blair M., Brick M.A., Chovatia M., Gepts P., Goodstein D.M., 769 Gonzales M., Hellsten U., Hyten D.L., Jia G., Kelly J.D., Kudrna D., Lee R., Richard M.M.S., 770 Miklas P.N., Osorno J.M., Rodrigues J., Thareau V., Urrea F., Carlos A., Wang M., Yu Y., Zhang M., Wing R.A., Cregan P.B., Rokhsar D.S., Jackson S.A. (2014). A reference genome 771 772 for common bean and genome-wide analysis of dual domestications. Nat. Genet. 46:707-773 713. DOI: 10.1038/ng.3008.

774 20• Bellucci E., Bitocchi E., Ferrarini A., Benazzo A., Biagetti E., Klie S., Minio A., Rau D., 775 Rodriguez M., Panziera A., Venturini L., Attene G., Albertini E., Jackson S.A., Nanni L., 776 Fernie A.R., Nikoloski Z., Bertorelle G., Delledonne M., Papa R. (2014) Decreased 777 nucleotide and expression diversity and modified co-expression patterns characterize 778 Plant Cell 26:1901-1912. DOI: domestication in the common bean. 779 10.1105/tpc.114.124040.

This study is focused on investigation of the effects of the domestication process in common bean using RNA-seq data. It is one of the pioneering reports on the consequences of domestication, not only at the genome level, but also at the gene-expression level. It highlights a decrease in gene expression diversity in domesticated compared to wild forms, a pattern that appears to be general for all or most domesticated species.

Passing 21 Borkotoky S., Saravanan V., Jaiswal A., Das B., Selvaraj S., Murali A., Lakshmi P.T.V.
(2013). The *Arabidopsis* stress responsive gene database. *International Journal of Plant Genomics* Vol. 2013, Article ID 949564. DOI: 10.1155/2013/949564.

- 22 Emms D.M. and Kelly S. (2015). OrthoFinder: solving fundamental biases in whole
 genome comparisons dramatically improves orthogroup inference accuracy. *Genome biology* 16(1):157. DOI: 10.1186/s13059-015-0721-2.
- 791 23 Osakabe Y., Arinaga N., Umezawa T., Katsura S., Nagamachi K., Tanaka H., Ohiraki H.,
- 792 Yamada K., Seo S., Abo M., Yoshimura E., Shinozaki K., Yamaguchi-Shinozaki K. (2013).
- 793 Osmotic stress responses and plant growth controlled by potassium transporters in
- 794 Arabidopsis. The Plant Cell 25:609–624. DOI: 10.1105/tpc.112.10570.
- **24** Meyer R.S., Choi J.Y., Sanches M., Plessis A., Flowers J.M., Amas J., Dorph K., Barretto
 A., Gross B., Fuller D.Q., Bimpong I.K., Ndjiondjop M.N., Hazzouri K.M., Gregorio G.B.,
 Purugganan M.D. (2016). Domestication history and geographical adaptation inferred from
 a SNP map of African rice. *Nature genetics* 48:1083–1088. DOI: 10.1038/ng.3633.
- Yang T., Zhang S., Hu Y., Wu F., Hu Q., Chen G., Cai J., Wu T., Moran N., Yu L., Xu G.
 (2014). The role of a potassium transporter OsHAK₅ in potassium acquisition and transport
 from roots to shoots in rice at low potassium supply levels. *Plant Physiology* 166(2):945959. DOI: 10.1104/pp.114.246520.
- Buckler E.S., Holland J.B., Bradbury P.J., Acharya C.B., Brown P.J., Browne C., Ersoz
 E., Flint-Garcia S., Garcia A., Glaubitz J.C., Goodman M.M., Harjes C., Guill K., Kroon D.E.,
 Larsson S., Lepak N.K., Li H.H., Mitchell S.E., Pressoir G., Peiffer J.A., Rosas M.O.,
 Rocheford T.R., Romay M.C., Romero S., Salvo S., Villeda H.S., da Silva H.S., Sun Q., Tian
 F., Upadyayula N., Ware D., Yates H., Yu J.M., Zhang Z.W., Kresovich S., McMullen M.D.
 (2009). The genetic architecture of maize flowering time. *Science* 325:714-718. DOI:
 10.1126/science.1174276.
- This study uses nested association mapping for QTL mapping in maize. This population is an extremely useful resource that was used to score the flowering time for nearly a million individual plants in four environments and over 2 years. Many QTLs with small additive effects on the flowering time were detected. This result is in contrast with other studies that investigated the genetic architecture of flowering time in other species, such as Arabidopsis and rice, where few loci of relatively large effects appear to control this trait.
- 27 Xu J., Liu Y., Liu J., Cao M., Wang J., Lan H., Xu Y., Lu Y., Pan G., Rong T. (2012). The
 genetic architecture of flowering time and photoperiod sensitivity in maize as revealed by
 QTL review and meta-analysis. *JIPB* 54(6):358-373. DOI: 10.1111/j.17447909.2012.01128.x.

- 820 **28** Navarro J.A.R, Willcox M., Burgueno J., Romay C., Swarts K., Trachsel S., Preciado E.,
- 821 Terron A., Vallejo Delgado H., Vidal V., Ortega A., Espinoza Banda A., Gomez Montiel N.O.,
- 822 Ortiz-Monasterio I., San Vicente F., Guadarrama Espinoza A., Arlin G., Wenzl P., Hearne
- 823 S., Buckler E. (2017). A study of allelic diversity underlying flowering-time adaptation in
- 824 maize landraces. *Nature Genetics* 49:476-480. DOI: 10.1038/ng.3784.
- 29 Thornsberry J.M., Goodman M.M., Doebley J., Kresovich S., Nielsen D., Buckler E.S. IV
 (2001). *Dwarf8* polymorphisms associate with variation in flowering time. *Nat Genet*. 28:286289.
- 30 Hung H.Y., Shannon L.M., Tian F., Bradbury P.J., Chen C., Flint-Garcia S.A., McMullen
 M.D., Ware D., Buckler E.S., Doebley J.F., Holland J.B. (2012). *ZmCCT* and the genetic
 basis of day-length adaptation underlying the post-domestication spread of maize. *PNAS*109(28): E1913-E1921. DOI: 10.1073/pnas.1203189109.
- 31 Yang Q., Li Z., Li W., Ku L., Wang C., Ye J., Li K., Yang N., Li Y., Zhong T., Li J., Chen
 Y., Yan J., Yang X., Xu M. (2013). CACTA-like transposable element in *ZmCCT* attenuated
 photoperiod sensitivity and accelerated the post-domestication spread of maize. *PNAS*111(42):16969-16974. DOI: 10.1073/pnas.1310949110.
- 32 Huang C., Sun H., Xu D., Chen Q., Liang Y., Wang X., Xu G., Tian C.W., Li D., Wu L.,
 Yang X., Jin W., Doebley J.F., Tian F. (2017). *ZmCCT9* enhances maize adaptation to
 higher latitudes. *PNAS* E334-E341. DOI: 10.1073/PNAS.1718058115.
- 33 Salvi S., Sponza G., Morgante M., Tomes D., Niu X., Fengler K.A., Meeley R., Ananiev
 E.V., Svitashev S., Bruggemann E., Li B., Hainey C.F., Radovic S., Zaina G., Rafalski J.A.,
 Tingey S.V., Miao G.H., Phillips R.L., Tuberosa R. (2007). Conserved noncoding genomic
 sequences associated with a flowering-time quantitative trait locus in maize. *PNAS*104(27):11376-11381. DOI: 10.1073/pnas.0704145104.
- 34 Guo L., Wang X., Zhao M., Huang C., Li C., Yang C.J., York A.M., Xue W., Xu G., Liang
 Y., Chen Q., Doebley J.F., Tian F. (2018). Stepwise *cis*-regulatory changes in *ZNC8*contribute to maize flowering-time adaptation. *Current biology* 28:3005-3015. DOI:
 10.1016/j.cub.2018.07.029.
- S., Dubreuil P., Gouesnard B., Manicacci D., Charcosset A. (2006). Maize adaptation to
 temperate climate: relationship between population structure and polymorphism of *Dwarf8*gene. *Genetics* 172:2449-63. DOI: 10.1534/genetics.105.048603.

- 36 Ducrocq S., Madur D., Veyrieras J.B., Camus-Kulandaivelu L., Kloiber-Matiz M., Presterl
 T., Ouzunova M., Manicacci D., Charcosset A. (2008). Key impact of *Vgt1* on flowering time
 adaptation in maize: evidence from association mapping and ecogeographical information. *Genetics* 178:2433-2437. DOI: 0.1534/genetics.107.084830.
- 37 Castelletti S., Tuberosa R., Pindo M., Salvi S. (2014). A MITE transposon insertion is
 associated with differential methylation at the maize flowering time QTL *Vgt1*. *G3: genes, genomes, genetics* 4(5):805-812. DOI: 10.1534/g3.114.010686.
- 38 Lazakis C.M., Coneva V., Colasanti J. (2011). *ZCN8* encodes a potential orthologue of *Arabidopsis* FT florigen that integrates both endogenous and photoperiod flowering signal
 in maize. *J. Exp. Bot.* 62(14):4833-4842. DOI: 10.1093/jxb/err129.
- 39 Meng X., Muszynski M.G., Danilevskaya O.N. (2011). The FT-like *ZCN8* gene functions
 as a floral activator and is involved in photoperiod sensitivity in maize. *Plant Cell* 23:942960. DOI: 10.1105/tpc.110.08140.
- 40•• Vigouroux Y., Mariac C., De Mita S., Pham J.L., Gérard B., Kapran I., Sagnard F., Deu
 M., Chantereau J., Ali A., Ndjeunga J., Luong V., Thuillet A.C., Saïdou A.A., Bezançon G.
 (2011). Selection for earlier flowering crop associated with climatic variations in the Sahel. *PLoS ONE* 6(5): e19563. DOI: 10.1371/journal.pone.0019563.
- This study represents an excellent example of the usefulness of historical collections (i.e., materials collected at different times) of landraces to identify adaptive introgression. The results show strong adaptation of pearl millet landraces to changing environmental conditions, even over relatively short evolutionary timescales. It also suggests that exploitation of genetic variability within landrace populations represents a strategy in response to future climate changes.
- 41 Saïdou A.A., Mariac C., Luong V., Pham J.L., Bezançon G., Vigouroux Y. (2009).
 Association studies identify natural variation at *PHYC* linked to flowering time and
 morphological variation in pearl millet. *Genetics* 82(3):899-910. DOI:
 10.1534/genetics.109.102756.
- 42•• Khan A.W., Garg V., Roorkiwal M., Golicz A.A., Edwards D., Varshney R.K. (2019).
 Super-Pangenome by integrating the wild side of a species for accelerated crop
 improvement. *Trends in Plant Science*. DOI: 10.1016/j.tplants.2019.10.012.
- 882 This opinion paper is an updated overview of the state of understanding about the recent 883 developments in crop pan-genomics. This study emphasises the importance of the

dispensable genome as a repertoire of adaptive variability, which is useful for the elucidation
of crop evolutionary history and the development of new improved varieties that are resistant
to stress and pathogens.

43 Zhou D., Zhou J., Meng L., Wang Q., Xie H., Guan Y., Ma Z., Zhong Y., Chen F., Liu J.
(2008). Duplication and adaptive evolution of the *COR15* genes within the highly coldtolerant *Draba* lineage (Brassicaceae). *Gene* 441:36-44. DOI: 10.1016/j.gene.2008.06.024.

44 DeBolt S. (2010). Copy number variation shapes genome diversity in *Arabidopsis* over
immediate family generational scales. *Genome Biol. Evol.* 2:441-453. DOI:
10.1093/gbe/evq033.

45 Mutti J.S., Bhullar R.K., Gill K.S. (2017). Evolution of gene expression balance among
homeologs of natural polyploids. *G3: Genes, Genomes, Genetics* 7(4):1225-1237. DOI:
10.1534/g3.116.038711.

46 Ceccarelli M., Santantonio E., Marmottini F., Amzallag G.N., Cionini P.G. (2006).
Chromosome endoreduplication as a factor of salt adaptation in *Sorghum bicolor*. *Protoplasma* 227:113-118. DOI: 10.1007/s00709-005-0144-0.

47 Saleh B., Allario T., Dambier D., Ollitrault P., Morillon R. (2008). Tetraploid citrus
rootstocks are more tolerant to salt stress than diploid. *Comptes Rendus Biologies* 331(9):
703-710. DOI: 10.1016/j.crvi.2008.06.007.

48 Le Corre V. and Kremer A. (2012). The genetic differentiation at quantitative trait loci
under local adaptation. *Molecular ecology* 21(7):1548-1566. DOI: 10.1111/j.1365294X.2012.05479.x.

49 Innan H. and Kim Y. (2004). Pattern of polymorphism after strong artificial selection in a
domestication event. *PNAS*, 101(29):10667-10672. DOI: 10.1073/pnas.0401720101.

50 Matuszewski S, Hermisson J., Kopp M. (2015). Catch me if you can: adaptation from
standing variation to a moving phenotypic optimum. *Genetics* 200(4):1255-1274. DOI:
10.1534/genetics.115.178574.

910 **51** Anderson E. and Stebbins G.L. Jr. (1954). Hybridization as an evolutionary stimulus.
911 *Evolution* 8(4):378-388. DOI: 10.2307/2405784.

52 Lewontin R.C. and Birch L.C. (1966). Hybridization as a source of variation for adaptation
to new environments. *Evolution* 20(3):223-236. DOI: 10.1111/j.1558-5646.1966.tb03369.x.

- 53 Hufford M.B., Lubinksy P., Pyhäjärvi T., Devengenzo M.T., Ellstrand N.C., Ross-Ibarra
 J. (2013). The genomic signature of crop-wild introgression in maize. *PLOS genetics* DOI:
 10.1371/journal.pgen.1003477.
- 54 Crosby A.W. (2003). The Columbian Exchange: biological and cultural consequences of1492. Westport, CT: Praeger Publishers.
- 55 Spooner D.M., McLean K., Ramsay G., Waugh R., Bryan G.J. (2005). A single
 domestication for potato based on multilocus amplified fragment length polymorphism
 genotyping. *PNAS* 102(41):14694-14699. DOI: 10.1073/pnas.0507400102.
- 56 Hawkes J.G. and Francisco-Ortega J. (1993). The early history of the potato in Europe. *Euphytica* 70(1-2):1-7.
- 57 Kloosterman B., Abelenda J.A., Carretero Gomez M.M., Oortwijn M., De Boer J.M.,
 Kowitwanich K., Horvath B.M., Van Eck H.J., Smaczniak C., Prat S., Visser R.G.F., Bachem
 C.W.B. (2013). Naturally occurring allele diversity allows potato cultivation in northern
 latitudes. *Nature* 495, 246–250. DOI: 10.1038/nature11912.
- 58 Hardigan M.A., Parker F., Laimbeer E., Newton L., Crisovan E., Hamilton J.P.,
 Vaillancourt B., Wiegert-Rininger K., Wood J.C., Douches D.S., Farré E.M., Veilleux R.E.,
 Buell R. (2017). Genome diversity of tuber-bearing *Solanum* uncovers complex evolutionary
 history and targets of domestication in the cultivated potato. *PNAS*. DOI:
 10.1073/pnas.1714380114.
- 59•• Gutaker R.M., Weib C.L., Ellis D., Anglin N.L., Knapp S., Fernandez-Alonso J.L., Prat
 S., Burbano H.A. (2019). The origins and adaptation of European potatoes reconstructed
 from historical genomes. *Nature Eco. Evol.* DOI: 10.1038/s41559-019-0921-3.
- This recent study is an excellent example of the value of degraded and/or historic botanical specimens to study crop evolution. The work was based on an investigation of historical samples that spanned 350 years of potato evolution in Europe. Their materials included historical herbarium specimens that they obtained from different European museums and modern American and European potato samples. Next-generation sequencing data were used to highlight the power of combining contemporary and historical genomes to shed light on the complex evolutionary history of crops, and their adaptation to new environments.
- 60 Reeves P.H. and Coupland G. (2000). Response of plant development to environment:
 control of flowering by daylength and temperature. *Current Opinion in Plant Biology* 3(1):3742. DOI: 10.1016/S1369-5266(99)00041-2.

- 946 61 Putterill J., Robson F., Lee K., Simon R., Coupland G. (1995). The CONSTANS gene
 947 of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger
 948 transcription factors. *Cell* 80:847-857. DOI: 10.1016/0092-8674(95)90288-0.
- 62 Visker M., Keizer L., Van Eck H., Jacobsen E., Colon L., Struik P. (2003). Can the QTL
 for late blight resistance on potato chromosome 5 be attributed to foliage maturity type? *Theor. Appl. Genet.* 106:317-325. DOI: 10.1007/s00122-002-1021-2.
- 63 Gonzales-Schain N.D., Diaz-Mendoza M., Zurczak M., Suarez-Lopez P. (2012). Potato
 CONSTANS is involved in photoperiodic tuberization in a graft-transmissible manner. *The Plant Journal*. 70(4):678-690. DOI: 10.1111/j.1365-313X.2012.04909.x.
- 64 Bradshaw J.E., Bryan G.J., Ramsay G. (2006). Genetic resources (including wild and
 cultivated *Solanum* species) and progress in their utilisation in potato breeding. *Potato Res.*49:49–65.
- 65 Angioi S.A., Rau D., Attene G., Nanni L., Bellucci E., Logozzo G., Negri V., Spagnoletti
 Zeuli P.L, Papa R. (2010). Beans in Europe: origin and structure of the European landraces
 of *Phaseolus vulgaris* L. *Theor Appl Genet*. 121:829–843. DOI: 10.1007/s00122-010-13532.
- 66 Gioia T., Logozzo G., Attene G., Bellucci E., Benedettelli S., Negri V. (2013) Evidence
 for introduction bottleneck and extensive inter-gene pool (Mesoamerica x Andes)
 hybridization in the European common bean (*Phaseolus vulgaris* L.) germplasm. *PLoS ONE*8: e75974. DOI: 10.1371/journal.pone.0075974.
- 67 Rebourg C., Gouesnard B., Charcosset A. (2001). Large scale molecular analysis of
 traditional European maize populations. Relationships with morphological variation. *Heredity* 86:574-587. DOI: 10.1046/j.1365-2540.2001.00869.x.
- 68 Rebourg C., Chastanet M., Gouesnard B., Welcker C., Dubreuil P., Charcosset A. (2003).
 Maize introduction into Europe: the history reviewed in the light of molecular data. *Theor. Appl. Genet.* 106:895-903. DOI: 10.1007/s00122-002-1140-9.
- 972 69 Tenaillon M.I and Charcosset A. (2011). A European perspective on maize history.
 973 *Comptes rendus biologies* 334(3):221-228- DOI: 10.1016/j.crvi.2010.12.015.
- 974 **70** Brandenburg J.T., Mary-Huard T., Rigaill G., Hearne S.J., Corti H., Joets J., Vitte C.,
 975 Charcosset A., Nicolas S.D., Tenallion M.I. (2017). Independent introductions and

- 976 admixtures have contributed to adaptation of European maize and its American 977 counterparts. *PLOS Genet.* 13(3): e1006666. DOI: 10.1371/journal.pgen.1006666.
- 978 This study is the most recent report of the introduction of maize into Europe. Sixty-seven 979 sequenced maize genomes that were representative of all of the American and European 980 diversity were used to carry out several population genomics and association mapping 981 approaches, to trace the origins of European maize and to investigate its demographic and 982 selective history.
- 71 Li Y.X., Li C.H., Bradbury P.J., Liu X.L., Lu F., Romay C.M., Glaubitz J.C., Wu X., Peng
 B., Shi Y., Song Y., Zhang D., Buckler E.S., Zhang Z., Li Y., Wang T. (2016). Identification
 of genetic variants associated with maize flowering time using an extremely large multigenetic background population. *The Plant Journal* 86(5):391-402. DOI: 10.1111/tpj.13174.
- 987 72 Virlouvet L., Jacquemot M.P., Gerentes D., Corti H., Bouton S., Gilard F., Valot B.,
 988 Trouverie J., Tcherkez G., Falque M., Damerval C., Rogowsky P., Perez P., Noctor G., Zivy
 989 M., Coursol S. (2011). The ZmASR1 protein influences branched-chain amino acid
 900 biosynthesis and maintains kernel yield in maize under water-limited conditions. *Plant*991 *Physioogy* 157(2):917-936. DOI: 10.1104/pp.111.176818.
- 73 Köllner T.G., Held M., Lenk C., Hiltpold I., Turlings T.C.J., Gershenzon J., Degenhardt J.
 (2008). A maize (E)-β-Caryophyllene synthase implicated in indirect defense responses
 against herbivores is not expressed in most American maize varieties. *The Plant Cell*20(2):482-494. DOI: 10.1105/tpc.107.051672.
- 74 Toubiana D., Fernie A.R., Nikoloski Z., Fait A. (2013). Network analysis: tackling complex
 data to study plant metabolism. *Trends in biotechnology* 31(1):29-36. DOI:
 10.1016/j.tibtech.2012.10.011.
- 75 Sade D., Shriki O., Cuadros-Inostroza A., Tohge T., Semel Y., Haviv Y., Willmitzer L.,
 Fernie A.R., Czosnek H., Brotman Y. (2015). Comparative metabolomics and
 transcriptomics of plant response to *Tomato yellow leaf curl virus* infection in resistant and
 susceptible tomato cultivars. *Metabolomics* 11(1):81-97. DOI: 10.1007/s11306-014-0670-x.
- 76 Beleggia R., Rau D., Laidò G., Platani C., Nigro F., Fragasso P.D.V., Scossa F., Fernie
 A.R., Nikoloski Z., Papa R. (2016). Evolutionary metabolomics reveals domesticationassociated changes in tetraploid wheat kernels. *Molecular biology and evolution* 33(7):17401753. DOI: 10.1093/molbev/msw050.

- 77 Pavlidis, P. and Alachiotis N.A (2017). Survey of methods and tools to detect recent and
 strong positive selection. *J of Biol Res-Thessaloniki* 24:7. DOI:10.1186/s40709-017-00640.
- 1010 **78** Lareau C.A. and McKinney B.A. (2015). Network theory for data-driven epistasis
 1011 networks. *Epistasis* 1253:285-300. DOI: 10.1007/978-1-4939-2155-3 15.
- 79• Lachowiec J., Shen X., Queitsch C., Carlborg Ö. (2015). A genome-wide association
 analysis reveals epistatic cancellation of additive genetic variance for root length in *Arabidopsis thaliana. PLoS genetics* 11(9): e1005541. DOI: 10.1371/journal.pgen.1005541. *This study is an excellent example of a genome-wide association analysis conducted in*Arabidopsis that provides valuable insights into the genetic mechanisms that underlie
 complex quantitative traits and the influence of epistasis on evolutionary processes.
- 1018 **80** Fang C., Ma Y., Wu S. Liu Z., Wang Z., Yang R., Hu G., Zhou Z., Yu H., Zhang M., Pan
- Y., Zhou G., Ren H., Du W., Yan H., Wang Y., Han D., Shen Y., Liu S., Liu T., Zhang J., Qin
 H., Yuan J., Yuan X., Kong F., Liu B., Li J., Zhang Z., Wang G., Zhu B., Tian Z. (2017).
 Genome-wide association studies dissect the genetic networks underlying agronomical
 traits in soybean. *Genome Biol.* 18:161. DOI:10.1186/s13059-017-1289-9.
- 1023 **81** Stinchcombe J.R. and Hoekstra H.E. (2008). Combining population genomics and 1024 quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* 1025 100(2):158-170. DOI: 10.1038/sj.hdy.6800937.
- 82 Schwarzenbacher H., Dolezal M., Flisikowski K., Seefried F., Wurmser C., Schlötterer
 C., Fries R. (2012). Combining evidence of selection with association analysis increases
 power to detect regions influencing complex traits in dairy cattle. *BMC genomics* 13(1):48.
 DOI: 10.1186/1471-2164-13-48.
- 83 Eckert A.J., Heerwaarden J.V., Wegrzyn J.L., Nelson C.D., Ross-Ibarra J., GonzàlesMartìnez S.C., Neale D.B. (2010). Patterns of population structure and environmental
 associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). *Genetics*185(3):969-982. DOI: 10.1534/genetics.110.115543.
- 84 Kemppainen P., Knight C.G., Sarma D.K., Hlaing T., Prakash A., Maung Maung Y.N.,
 Somboon P., Mahanta J., Walton C. (2015). Linkage disequilibrium network analysis (LDna)
 gives a global view of chromosomal inversions, local adaptation and geographic structure.
- 1037 *Molecular ecology resources* 15(5):1031-1045. DOI: 10.1111/1755-0998.12369.

1038 **85** Lin Z., Li X., Shannon L.M., Yeh C.T., Wang M.L., Bai G., Peng Z., Li J., Trick H.N., 1039 Clemente T.E., Doebley J., Schnable P.S., Tuinstra M.R., Tesso T.T., White F., Yu J. (2012).

- 1040 Parallel domestication of the Shattering1 genes in cereals. *Nat. Genet.* 44:720–724. DOI:
- 1041 10.1038/ng.2281.

1042 **86** Blanquart F. and Bataillon T. (2009). Epistasis and the Structure of Fitness Landscapes:

1043 Are Experimental Fitness Landscapes Compatible with Fisher's Geometric Model?. 1044 *Genetics* 203(2):847-862. DOI: 10.1073/pnas.0906720106.

- 87•• Exposito-Alonso M., Burbano H.A., Bossdorf O., Nielsen R., Weigel D. (2019). Natural
 selection on the Arabidopsis thaliana genome in present and future climates. *Nature*573(7772):126-129. DOI: 10.1038/s41586-019-1520-9.
- 1048 This study is based on an analysis of the extensive genome information available for 1049 Arabidopsis thaliana and the measures of the fitness on different rainfall conditions for 517

1050 *natural* Arabidopsis lines grown in Germany and Spain. This experiment provides proof of

1051 concept for the use of genome-wide environment selection models for evolution-aware

1052 predictions of risks for biodiversity that are associated with climate changes.

1053

1054 Acknowledgements

1055 This work was supported by grants from the ERA-NET for Coordinating Action in Plant

1056 Sciences-2nd ERA-CAPS call, BEAN_ADAPT project.