Phylogeny and disparate selection signatures suggest two genetically independent domestication events in pea (Pisum L.)

Timo Hellwig1,2,3,* 1, Shahal Abbo1 1 and Ron Ophir2 2
1The Levi Eshkol School of Agriculture, The Hebrew University of Jerusalem, Jerusalem, Rehovot, Israel, 2Volcani Center, Agricultural Research Organization, Rishon LeZion, Israel, and 3Institute of Plant Genetics, Heinrich-Heine-University, Düsseldorf, Germany

Received 13 November 2021; accepted 15 January 2022; published online 21 January 2022.
*For correspondence (e-mail timo.hellwig@hhu.de).

SUMMARY

Domestication is considered a model of adaptation that can be used to draw conclusions about the modus operandi of selection in natural systems. Investigating domestication may give insights into how plants react to different intensities of human manipulation, which has direct implication for the continuing efforts of crop improvement. Therefore, scientists of various disciplines study domestication-related questions to understand the biological and cultural bases of the domestication process. We employed restriction site-associated DNA sequencing (RAD-seq) of 494 Pisum sativum (pea) samples from all wild and domesticated groups to analyze the genetic structure of the collection. Patterns of ancient admixture were investigated by analysis of admixture graphs. We used two complementary approaches, one diversity based and one based on differentiation, to detect the selection signatures putatively associated with domestication. An analysis of the subpopulation structure of wild Pisum sativum revealed five distinct groups with a notable geographic pattern. Pisum abyssinicum clustered unequivocally within the P. sativum complex, without any indication of hybrid origin. We detected 32 genomic regions putatively subjected to selection: 29 in P. sativum ssp. sativum and three in P. abyssinicum. The two domesticated groups did not share regions under selection and did not display similar haplotype patterns within those regions. Wild P. sativum is structured into well-diverged subgroups. Although Pisum sativum ssp. humile is not supported as a taxonomic entity, the so-called ‘southern humile’ is a genuine wild group. Introgression did not shape the variation observed within the sampled germplasm. The two domesticated pea groups display distinct genetic bases of domestication, suggesting two genetically independent domestication events.

Keywords: crop wild relative, domestication, genetic diversity, introgression, phylogeny, signatures of selection.

INTRODUCTION

Ironically, despite being a small genus and a genetic model system ever since Mendel (1866) published his work, the taxonomy of Pisum is still constantly debated. The Kew database (http://www.kew.org/data/) recognizes two distinct Pisum species, Pisum fulvum Sibth. & Smith and Pisum sativum L. Whereas P. fulvum is listed as a monophyletic species without any synonyms, P. sativum has 17 synonyms. Depending on the author, some of those synonyms are considered distinct species, subspecies or ecotypes, or are not considered a taxonomic group at all, which illustrates the uncertainties concerning the taxonomy of this genus. The first taxonomy of the Pisum genus was proposed by Boissier (1867), who recognized based on morphology the domesticated field pea, P. sativum and three wild species, P. fulvum, Pisum elatius M. Bieb. and Pisum humile Boiss. et Noe. Ben-Ze’ev and Zohary (1973) adopted this taxonomy when they studied the crossability relationships between different Pisum taxa. They confirmed the claims of botanists that P. elatius, P. humile and P. sativum should be considered as a single entity (a biological species) because crosses between them mostly result in fertile hybrids. Yet, the authors also found that P. humile accessions, which originated from central Turkey and northern Israel, share a chromosomal translocation with P. sativum that was not present in P. elatius and in
P. humile accessions from southern Israel. Although P. fulvum can be crossed with individuals from the other groups, Ben-Zé’ev and Zohary (1973) have shown that these crosses produce only few fertile progeny and exhibit several chromosomal rearrangements. The results of Ben-Zé’ev and Zohary (1973) support the classification proposed by Davis (1970), who recognized two species: namely P. fulvum and P. sativum, which includes the domesticated Pisum sativum ssp. sativum and the two wild subspecies Pisum sativum ssp. elatius and Pisum sativum ssp. humile. In the context of the nomenclature it should be noted that the epithet P. sativum ssp. humile is in fact incorrect, as this entity was previously named Pisum sativum ssp. syriacum (Berger, 1928). However, to enable readers to cross-reference the past rich literature on Pisum genetics and taxonomy, herein we will refer to this entity as P. sativum ssp. humile. Davis’ (1970) classification gained further support from DNA polymorphism studies by Palmer et al. (1985) and Hoey et al. (1996). However, the germplasm samples analysed in the latter two studies largely overlapped with those of Ben-Zé’ev and Zohary (1973), which were mostly from the Near East. For example, the domesticated Pisum abyssinicum A. Br. was missing from that collection.

Pisum abyssinicum was initially described at the species rank, even though the author mentioned that it could be a subspecies (Braun, 1841). One hundred and eighty years hence, the discussion about the taxonomic rank of P. abyssinicum is not yet resolved. Although some authors consider it as a good species because of its isolation in phylogenetic analyses (Ellis et al., 1998; Maxted & Ambrose, 2001; Zong et al., 2009), others consider it a subspecies or an ecotype (a cultivar group) of the domesticated P. sativum ssp. sativum (Makasheva, 1984; Nasiri et al., 2009; Weeden, 2018). Studies based on nuclear DNA markers produced slightly different phylogenies than studies based on morphological criteria or crossability patterns (Ellis, 2011; Jing et al., 2007, 2010; Vershinin et al., 2003).

The latter studies support the classification of Maxted and Ambrose (2001), who recognized three species, P. fulvum, P. abyssinicum and P. sativum, and divided the P. sativum complex into a domesticated subspecies, P. sativum ssp. sativum, and a single wild subspecies, P. sativum ssp. elatius. As we write, most researchers follow the taxonomy of Maxted and Ambrose (2001) and consider P. sativum ssp. humile a synonym of P. sativum ssp. elatius, because it was argued that it could not be genetically distinguished from P. sativum ssp. elatius accessions (Ellis, 2011; Vershinin et al., 2003). However, recently, using a large array of P. sativum ssp. elatius and P. sativum ssp. humile accessions, we confirmed the genetic distinctiveness of the ‘elatius’ and ‘humile’ groups (Hellwig, Abbo, & Ophir, 2021). Further, the divergence of two ecologically unique P. sativum ssp. humile stocks, first identified by Ben-Zé’ev and Zohary (1973), and proposed as the two varieties P. sativum ssp. humile var. syriacum (‘northern humile’) and P. sativum ssp. humile var. humile (‘southern humile’; Ladizinsky & Abbo, 2015), was corroborated by Hellwig, Abbo, & Ophir (2021). However, the so-called northern and southern humile ecotypes were just as genetically distinct from each other as they were from the other wild P. sativum subgroups (Hellwig, Abbo, & Ophir, 2021).

Based on the similar chromosome linear order (as evident from meiotic pairing patterns) uniquely shared by P. sativum ssp. sativum and the tested northern humile stocks, Ben-Zé’ev and Zohary (1973) inferred that northern humile represents the ancestral wild stock of the domesticated P. sativum ssp. sativum. Yet, they also considered potential contributions of other stocks to the primary domesticated gene pool as very likely, as the chromosomal cytotypes across the P. sativum complex are rather similar. Indeed, phylogenetic trees based on chloroplast DNA variation reported by Palmer et al. (1985) suggest a close relationship between P. sativum ssp. humile and all but one P. sativum ssp. sativum accession, and hence agrees with the conclusions of Ben-Zé’ev and Zohary (1973). Reservations concerning the domestication model proposed by Ben-Zé’ev and Zohary (1973) were raised by Hoey et al. (1996) who combined morphological, allozyme and random amplified polymorphic DNA (RAPD) markers to reconstruct the Pisum phylogeny. Their results were ambiguous and depending on the method of cladogram construction, the closest relative of P. sativum ssp. sativum was either northern humile or a polytom of northern humile and P. sativum ssp. elatius. Vershinin et al. (2003) suggested that P. sativum ssp. sativum was derived from P. sativum ssp. elatius, as it had the highest number of exclusively shared transposon-based markers with P. sativum ssp. elatius. However, note that only two P. sativum ssp. humile accessions were employed by Vershinin et al. (2003).

Jing et al. (2010) supported the notion of P. sativum ssp. elatius as the wild progenitor of P. sativum ssp. sativum and proposed the following model of pea domestication: early farmers selected and extensively grew P. sativum ssp. elatius in the Fertile Crescent; early cultivars spread eastwards, which resulted in the emergence of pea landraces still grown in Central Asia and in the Himalayan region today; and westward expansion of early cultivars gave rise to the European domestic pea (P. sativum ssp. sativum) germplasm, which eventually developed into modern elite varieties. It is noteworthy that studies proposing P. sativum ssp. humile as a separate taxonomic unit and as the primary wild stock of P. sativum ssp. sativum have broadly employed germplasm arrays similar to those of Ben-Zé’ev and Zohary (1973). However, studies proposing P. sativum ssp. elatius as the wild progenitor of P. sativum ssp. sativum and considering P. sativum ssp. humile as a minor taxonomic unit were undertaken with a meagre
share of *P. sativum* ssp. *humile* individuals. This underlines the need for a study based on a balanced germplasm array to investigate the taxonomic status of this entity and to elucidate the unclear mode of pea domestication.

In the context of pea taxonomy and domestication, attention should be paid to the status of the so-called ‘southern *humile*’. As noted above, although many pea geneticists consider *P. sativum* ssp. *humile* a synonym of *P. sativum* ssp. *elatius*, the affinity of southern *humile* to disturbed (manmade) habitats is often overlooked. Ben-Ze’ev and Zohary (1973) described the occurrence of southern *humile* on the edges of cultivation and as a weed within cereal fields. In such secondary habitats, it was documented across a wide range of different Israeli environments (Abbo, Lev-Yadun, et al., 2013). A prime example of the ecological preferences of southern *humile* can be observed in an abandoned olive grove near Beit Guvrin (Israel) where southern *humile* grows in sympathy with *P. fulvum* (observation confirmed in April 2021). Although *P. fulvum* invades the olive grove from the adjacent range-land and can be found on the hillside as well as in the olive grove, southern *humile* individuals are confined to the abandoned olive plantation and edges of the adjacent cereal fields (Abbo, Lev-Yadun, et al., 2013). This distribution pattern of southern *humile* may be interpreted as an indication of feral origin, but no investigation has ever confirmed or rejected this notion.

Another issue mentioned above is the fact that some authors consider *P. abyssinicum* as a subspecies or ecotype of *P. sativum* ssp. *sativum*, thereby assuming it shares the same domestication history. The Abyssinian pea is often regarded as a problematic taxon. Govorov (1937) first mentioned the possibility that *P. abyssinicum* may be of hybrid origin, derived from a cross between *P. sativum* ssp. *sativum* and *P. fulvum*. Kloz (1971) formulated the hypothesis that *P. abyssinicum* may not have been derived from *P. sativum* ssp. *sativum* but rather originated from a cross between *P. sativum* ssp. *humile* and *P. fulvum*, based on serological reactions of *Pisum* sp. seed proteins. This claim attracted more attention after Ellis et al. (1998) suggested that *P. abyssinicum* is a distinct taxonomic unit that was domesticated independently of *P. sativum* ssp. *sativum*. Note that we use the term ‘independent domestication’ in a strict genetic sense throughout this text, i.e. the adoption of distinct wild stocks without substantial introgression between the early domesticated lineages. We differentiate this genetic definition from the cultural usage of terms like ‘independent domestication’ (or ‘multiple domestications’) *sensu* Willcox (2005), where authors assume that the concept of cultivation occurred independently in different regions within independent cultural contexts, which we consider an unlikely scenario for the Near East (Abbo et al., 2010; Gopher et al., 2021). Moreover, we distinguish between (i) the domestication episode, i.e. a short period of time (below archaeological and $^{14}$C dating resolution) when plants were transferred from wild populations into human-made habitats, accompanied by just a few crucial genetic changes that led to phenotypes that enabled the economically feasible cultivation of these early domesticates, and (ii) later in time, subsequent crop evolution that led to significant genetic adaptations to the artificial habitat of arable fields and human agronomic and culinary preferences (e.g. Abbo et al., 2012; Abbo et al., 2014; Gopher et al., 2021). As *P. abyssinicum* clustered between *P. sativum* ssp. *elatius* and *P. fulvum* in neighbor-joining dendrograms and *P. abyssinicum* shared most markers with these groups, Vershinin et al. (2003) proposed, in line with Ellis et al. (1998), that *P. abyssinicum* originated from ancient hybridization between *P. sativum* ssp. *elatius* and *P. fulvum*. This hypothesis was supported by Jing et al. (2010) who also clustered *P. abyssinicum* between these two wild taxa. Based on the variation of a histone H1 gene, Zaytseva et al. (2012) suggested that *P. abyssinicum* is distinct from *P. sativum* ssp. *sativum* and concluded that *P. abyssinicum* was independently domesticated, probably from a wild *P. sativum* subspecies. Weeden (2018) concluded from the sequence variation of 54 genes that *P. fulvum* did not contribute to the *P. abyssinicum* gene pool, thereby casting doubts on the interpretations of Vershinin et al. (2003) and Jing et al. (2010). Weeden (2018) did not reject the claim that *P. abyssinicum* was derived from *P. sativum* ssp. *sativum*, but mentioned the possibility that *P. abyssinicum* may have originated from a cross between *P. sativum* ssp. *sativum* and *P. sativum* ssp. *elatius*.

Herein we approach several of the disputed questions regarding the phylogeny and domestication history of pea by analysing genome-wide genotyping by sequencing data from a large germplasm array of 494 pea accessions, including *P. abyssinicum*, *P. fulvum*, *P. sativum* ssp. *elatius*, *P. sativum* ssp. *humile* and *P. sativum* ssp. *sativum*. More specifically, we address the following questions: (i) is there any genetic evidence for a divergence of *P. sativum* ssp. *elatius* and *P. sativum* ssp. *humile*; (ii) if so, is *P. sativum* ssp. *humile* var. *humile* (southern *humile*) of feral origin or is it a genuine wild stock; (iii) which wild group is the most likely stock of the domesticated pea, *P. sativum* ssp. *sativum*; (iv) should *P. abyssinicum* be considered as part of the *P. sativum* complex; and (v) if so, was *P. abyssinicum* derived from a hybrid origin and was it domesticated independently of the central stock of *P. sativum* ssp. *sativum*?

RESULTS

Genotyping by sequencing

In total, we called 4 553 130 variants and retained 136 532 and 91 384 single-nucleotide polymorphisms (SNPs) after quality filtering and linkage disequilibrium (LD) pruning, respectively. The SNPs were distributed throughout the
entire genome. Although the SNP density varied over the linkage groups, we did not observe large regions without any SNPs in the entire SNP set as well as in the LD pruned set (Figure S1). Both SNP sets exhibited low missing genotypes overall, with a median and mean of 11% and 14%, respectively (Figure S2). Observed heterozygosity was extremely low with only 1% at the third quartile (Figure S2). We observed an abundance of rare alleles with minor allele frequency (MAF) at the third quartile of 0.13 and 0.11 in the filtered and the LD pruned sets, respectively (Figure S2). The mean and median values of nucleotide diversity were 0.08 and 0.14, respectively, in the entire variant set. After LD pruning, similar values were observed (Figure S2).

Population structure

Wild samples were clearly separated into *P. fulvum* and *P. sativum* (Figures 1 and 2). Within the *P. sativum* complex, we observed five distinct groups (WS 1–5) in sparse non-negative matrix factorization (sNMF) analysis with *K* = 8. We chose *K* = 8 because the results with *K* = 8 were most congruent with the other analyses (phylogenetic network, neighbor-joining tree and principal coordinate analysis (PCoA)). With increasing *K* values, more subgroups of *P. fulvum* emerged (Figure 1b). This structure was also supported by the other methods. The first split within *P. sativum* in the neighbor-joining tree separated WS3 from all other *P. sativum* subpopulations. WS3, which contained all southern *humile* samples, was also a distinct subpopulation in the principal coordinate analysis, phylogenetic network and in sNMF, where it appeared as an independent group already with *K* = 3. Low levels of admixture were observed in WS3. WS1 and WS2 had a close genetic relationship in all analyses employed. In sNMF, several WS1 samples showed comparably high subpopulation fractions from WS2. In the phylogenetic network, WS1 was divided into two subgroups by WS2. All samples of WS2 were northern *humile* from the Tel Abu Nida location in northern Israel. One of the WS1 subgroups contained only northern *humile* samples, with one exception (PeAb03). The other WS1 subgroup contained mostly *P. sativum* ssp. *elatius* samples, but also included three northern *humile* samples and one *P. fulvum* sample. The separation of WS1 and WS2 was not unambiguous. With their close genetic relationship, these groups could also be merged into a single group or separated into three subgroups according to the location of their samples in the phylogenetic network and neighbor-joining tree. Yet, as WS2 was separated from WS1 by rather long branches in the phylogenetic network and neighbor-joining tree, which is in line with the sNMF assignment, we decided to employ the clustering of sNMF for subgroup notation.

WS4 and WS5 were closely related and consisted entirely of *P. sativum* ssp. *elatius* samples. Only a few samples of these two subpopulations exhibited admixture, which was small in all cases. Ten samples (six *P. sativum* ssp. *elatius* and four northern *humile*) were highly admixed and were not assigned to any subpopulation. Three samples were closest to WS1 and WS2, whereas six samples built a small but poorly resolved group that was closest to southern *humile*. The admixed sample PeAb15 from Greece clustered between *P. fulvum* and the *P. sativum* complex.

When domesticated samples were added, the separation of the *P. sativum* complex and *P. fulvum* was still apparent and represented the uppermost level of the hierarchic structure (Figure 3). Clusters WS3, WS4 and WS6 remained distinct subgroups with only minor admixture. *Pisum abysinicum* formed a distinct clade in all analyses (Figure 3): it clustered close to southern *humile* (WS3), and only in the neighbor-joining tree did it form a clade with WS1, WS2 and *P. sativum* ssp. *sativum* (Figure 3b). *Pisum sativum* ssp. *sativum* samples formed a clade with WS1 and WS2. The results of sNMF showed that WS2 still formed a subpopulation with *K* = 8, whereas WS1 samples were either highly admixed with a large subpopulation fraction from WS2 and *P. sativum* ssp. *sativum* or had more than 50% subpopulation fractions from *P. sativum* ssp. *sativum*. Only with *K* > 12 did WS1 appear as a distinct subpopulation in sNMF. In the neighbor-net phylogenetic network and the neighbor-joining tree, WS1 samples were separated into two clades. The clade closest to *P. sativum* ssp. *sativum* was sampled from southern Turkey and the other clade contained samples from south and west Turkey as well as samples from Georgia, and the northern and western shore of the Black Sea (Figure S3).

Six samples (PFO01, Paooe00, PeAb06, PeAb68, PeAb28, PhZm13) were clustered with samples that they could not be related to. This may have happened through misclassification, e.g. PeAb60 and PeAb68 from the John Innes Center are likely to be incorrectly classified as wild, or through a mix-up during DNA extraction or the sequencing procedure, e.g. PFO01, a *P. fulvum* sample that exhibited an SNP pattern of *P. sativum*. These six samples were excluded from subsequent analyses.

Admixture analysis

*Pisum fulvum* and *P. sativum* created separated clades, and the position of the *P. sativum* ssp. *elatius* subgroups is consistent. In the neighbor-net phylogenetic network, WS5 and WS4 may be interpreted as diverged either from the *P. sativum* ssp. *elatius* and northern *humile* lineage or from southern *humile*. Although admixture was observed in WS5 (Figure 3b; sNMF), this could not fully resolve the inconsistency. We, therefore, further investigated the extent of the admixture in the *Pisum* lineage.

The first split in the admixture graph separated Vavilovia *formosa* from the *Pisum* genus (Figure 4). Within *Pisum*, the admixture graph splits into two branches, one leading to *P. fulvum* and the other leading to *P. abysinicum* and later to *P. sativum* ssp. *sativum*. None of these taxa...
Figure 1. Population structure analysis including only wild *Pisum* (pea) samples. (a) Neighbor-net phylogenetic network. (b) Neighbor-joining tree (top) and subpopulation fraction estimated by sparse non-negative matrix factorization (sNMF) with $K$ values ranging from 3 to 9 (bottom). Samples were colored according to subpopulation fractions of sNMF with $K = 8$. Samples that did not have a sub-population fraction of $>0.5$ of any genetic cluster were defined as admixed (black).
received any admixture. The results suggest a complex history of admixture in the wild P. sativum groups, which were all derived from hybrid origin. Northern humile had 89% ancestry from its most common recent ancestor with P. sativum ssp. sativum and another 11% from a group that split earlier from the direct ancestor of P. fulvum. The vast majority of P. sativum ssp. elatius ancestry (92%) came from an early ancestor of P. sativum ssp. sativum, whereas the remaining 8% came from an ancestor that originated from the P. fulvum branch. Yet, this latter ancestor strongly diverged from the direct ancestor of P. fulvum (51 drift weight). That same ancestor was involved in the hybridization event that led to southern humile (28%), which had 72% ancestry of the direct ancestor of P. abyssinicum. All values of \( f_2(\text{P. abyssinicum}; \text{B}, \text{C}) \) were slightly positive, ranging from 0.063 to 0.087. All of them were significantly different from 0, with a minimum Z score of 35.39 (Table 1).

### Diversity statistics

**Pisum abyssinicum** exhibited by far the lowest genetic diversity, with observed heterozygosity (\( H_o \)), expected heterozygosity (\( H_e \)) and nucleotide diversity (\( \pi \)) values of 0.0562, 0.0837 and 0.0850, respectively (Table 2). **Pisum sativum** ssp. *sativum* had high diversity at the lower range of its wild relatives (\( H_o = 0.0182, H_e = 0.2401, \pi = 0.2431 \)). **Pisum fulvum** (\( H_o = 0.0184, H_e = 0.1971, \pi = 0.1976 \)) had slightly lower values relative to the wild P. sativum groups and its subpopulations that, among themselves, showed a range of diversity at a comparable level except for WS2, which had the highest values of all groups with \( H_o, H_e \) and \( \pi \) values of 0.0824, 0.4469 and 0.4683, respectively.

### Genetic signatures of selection

The 5% threshold of **pcadapt** was passed by 987 SNPs in P. sativum ssp. *sativum*, whereas 851 and 720 of those exceeded the 2.5% and 1.0% thresholds, respectively (Figure 5). Thirty-nine outlier SNPs were found in the 95% \( X^2 \) threshold. Eight of those were also outliers in the 97.5% \( X^2 \) threshold and two were outliers in the 99% \( X^2 \) threshold. In P. abyssinicum, we detected 437 outlier SNPs with **pcadapt**, and 414 and 386 of those also passed the 2.5% and 1.0% thresholds, respectively. With **baypass**, 180 outliers passed the 95% threshold and 136 passed the 97.5% threshold in P. abyssinicum, whereas three SNPs exceeded the 99% threshold of the pseudo observed data set (POD).
Figure 3. Population structure analysis including all pea samples. (a) Neighbor-net phylogenetic network. (b) Neighbor-joining tree (top) and subpopulation fraction estimated by sparse non-negative matrix factorization (sNMF) with K values ranging from 3 to 9 (bottom). Samples that did not have a subpopulation fraction of >0.5 of any genetic cluster were defined as admixed. Wild samples were colored as described in Figure 1.

© 2022 The Authors. *The Plant Journal* published by Society for Experimental Biology and John Wiley & Sons Ltd., *The Plant Journal*, (2022), 110, 419–439
SNPs passing the 97.5% XTX threshold and the 2.5% PCA-applied (Figure 5; Table 3). Eight windows included outlier when the most relaxed threshold (95% and 5%) was considered candidates according to the most stringent threshold (99% and 1%). Three windows containing at least one SNP that passed the 95% XTX threshold and the 5% threshold (99% and 1%). Three windows containing at least one PCADAPT and one BAYPASS outlier that contained at least one PCADAPT and one BAYPASS outlier.

Population 1 Population 2 Population 3 $f_3$ SE Z score

| $P. abyssinicum$ | $P. fulvum$ | $P. sativum$ ssp. elatius | 0.074 | 0.0018 | 40.71 |
| $P. abyssinicum$ | $P. fulvum$ | Northern $humile$ | 0.074 | 0.0019 | 39.76 |
| $P. abyssinicum$ | $P. fulvum$ | Southern $humile$ | 0.075 | 0.0020 | 38.26 |
| $P. abyssinicum$ | $P. fulvum$ | $P. sativum$ ssp. sativum | 0.071 | 0.0019 | 37.54 |
| $P. abyssinicum$ | Southern $humile$ | $P. sativum$ ssp. elatius | 0.075 | 0.0018 | 40.52 |
| $P. abyssinicum$ | Southern $humile$ | $P. fulvum$ | 0.075 | 0.0020 | 38.26 |
| $P. abyssinicum$ | Southern $humile$ | Northern $humile$ | 0.067 | 0.0018 | 37.57 |
| $P. abyssinicum$ | Southern $humile$ | $P. sativum$ ssp. sativum | 0.063 | 0.0018 | 35.39 |
| $P. abyssinicum$ | $P. sativum$ ssp. elatius | $P. fulvum$ | 0.074 | 0.0018 | 40.71 |
| $P. abyssinicum$ | $P. sativum$ ssp. elatius | Northern $humile$ | 0.079 | 0.0018 | 43.58 |
| $P. abyssinicum$ | $P. sativum$ ssp. elatius | Southern $humile$ | 0.075 | 0.0018 | 40.52 |
| $P. abyssinicum$ | $P. sativum$ ssp. elatius | $P. sativum$ ssp. sativum | 0.079 | 0.0019 | 41.99 |
| $P. abyssinicum$ | Northern $humile$ | $P. sativum$ ssp. elatius | 0.079 | 0.0018 | 43.58 |
| $P. abyssinicum$ | Northern $humile$ | $P. fulvum$ | 0.074 | 0.0019 | 39.76 |
| $P. abyssinicum$ | Northern $humile$ | Southern $humile$ | 0.067 | 0.0018 | 37.57 |
| $P. abyssinicum$ | Northern $humile$ | $P. sativum$ ssp. sativum | 0.087 | 0.0020 | 42.65 |
| $P. abyssinicum$ | $P. sativum$ ssp. sativum | $P. sativum$ ssp. elatius | 0.079 | 0.0019 | 41.99 |
| $P. abyssinicum$ | $P. sativum$ ssp. sativum | $P. fulvum$ | 0.071 | 0.0019 | 37.54 |
| $P. abyssinicum$ | $P. sativum$ ssp. sativum | Northern $humile$ | 0.087 | 0.0020 | 42.65 |
| $P. abyssinicum$ | $P. sativum$ ssp. sativum | Southern $humile$ | 0.063 | 0.0018 | 35.39 |

In $P. sativum$ ssp. sativum, 39 5-kb windows were found that contained at least one PCADAPT and one BAYPASS outlier when the most relaxed threshold (95% and 5%) was applied (Figure 5; Table 3). Eight windows included outlier SNPs passing the 97.5% X’TX threshold and the 2.5% PCA-DAPT threshold, and two of those windows were also considered candidates according to the most stringent threshold (99% and 1%). Three windows containing at least one SNP that passed the 95% X’TX threshold and the 5% PCADAPT threshold were located less than 5 kb apart from each other in $P. abyssinicum$ (Figure 5; Table 3). Two of those windows also passed the 97.5% (2.5%) thresholds, but none passed the strictest threshold.

The genetic relationship between $P. sativum$ groups in neighbor-joining trees of the 6-Mbp areas surrounding candidate windows varied depending on the window, yet the domesticated groups did not cluster together in any of those areas (Figures 6a and S4–S33). It should be noted, however, that the bootstrap values were comparably low (<80%) at many nodes. Likewise, the haplotype patterns of $P. abyssinicum$ and $P. sativum$ ssp. sativum within candidate windows differed more from each other than the pattern of domesticated and certain wild samples (Figures 5b and S4–S33). We could not find any haplotype within the candidate windows that was unique to the domesticated groups.

DISCUSSION

Genetic diversity in wild pea

All methods employed to investigate the genetic structure in our germplasm collection clearly separated $P. fulvum$ from $P. sativum$ (Figures 1–3), corroborating previous reports (e.g. Ben-Ze’ev & Zohary, 1973; Boissier, 1867; Davis, 1970; Hoey et al., 1996; Jing et al., 2007, 2010; Palmer et al., 1985). Note, however, that Ellis (2011) claimed that ‘‘...biologically there is a good reason to consider Pisum as one species’’ because it has been reported that all pea groups can exchange genetic material (Mxated & Ambrose, 2001) and that even between distantly related stocks allelic introgression occurs (Jing et al., 2005, 2010; Vershinin et al., 2003). Our data call for a reconsideration of Ellis’ (2011) statement, as none of our respective results can clearly link the variant patterns shared among taxonomic groups to introgression events. Moreover, although $P. fulvum$ can be crossed with individuals from the $P. sativum$ complex, such crosses result in very low fractions of fertile progeny (Ben-Ze’ev & Zohary, 1973; Bobkov & Selikhova, 2017; Kosterin & Bogdanova, 2015). Like Ben-Ze’ev and Zohary (1973), who reported monitoring sympatric populations of $P. fulvum$ and wild $P. sativum$ in Israel over many years, we have never observed the typical pale orange to pinkish flower color that is typical to such hybrids during repeated surveys of such Israeli sympatric populations for more than 20 flowering seasons (for partial site lists, see Abbo et al., 2008; Abbo, Zezak, et al., 2013). The extent of natural hybridization between $P. fulvum$ and wild $P. sativum$ in wild populations could be a key to resolve this issue but has yet to be investigated. Our results do not support the notion of extensive introgression between $P. fulvum$ and the $P. sativum$ complex, yet we refrain from making any taxonomic reclassification. For the genetic structure within $P. fulvum$, readers may consult Hellwig, Abbo, Sherman, and Ophir (2021).
Independent domestication events in peas

Table 2 Statistics of diversity ($H_o$, observed heterozygosity; $H_e$, expected heterozygosity; $\pi$, nucleotide diversity) of each Pisum group and genetic cluster within Pisum sativum ssp. sativum

|          | $H_o$   | $H_e$   | $\pi$  |
|----------|---------|---------|--------|
| P. fulvum| 0.0184  | 0.1971  | 0.1976 |
| P. sativum ssp. elatus | 0.0239  | 0.2046  | 0.2058 |
| P. sativum ssp. humile var. syriacum (northern) | 0.0152  | 0.2524  | 0.2573 |
| P. sativum ssp. humile var. humile (southern) | 0.0511  | 0.2306  | 0.2328 |
| WS1      | 0.0121  | 0.2514  | 0.2555 |
| WS2      | 0.0824  | 0.4469  | 0.4683 |
| WS3      | 0.0511  | 0.2306  | 0.2328 |
| WS4      | 0.0566  | 0.2643  | 0.2669 |
| WS5      | 0.0387  | 0.2893  | 0.3005 |
| P. abyssinicum | 0.0562  | 0.0837  | 0.0850 |
| P. sativum ssp. sativum | 0.0192  | 0.2401  | 0.2431 |

Within the P. sativum complex, five distinct subpopulations were observed (Figures 1 and 2). The genetic structure of 82 of our 199 wild P. sativum samples was described earlier by Hellwig, Abbo, and Ophir (2021), who named their six clusters C1–C6. Both sample sets produced overlapping results. WS5 contains the same accessions as cluster C3 of Hellwig, Abbo, and Ophir (2021), which were sampled from southern Europe. WS3 contains all samples classified as southern humile on top of those employed by Hellwig, Abbo, and Ophir (2021). WS4 is the equivalent of C5 located in northern Israel. WS2 contains northern humile accessions that were sampled in the vicinity of Tel Abu Nida, Golan Heights, Israel. This corresponds to the approximate origin of Ben-Ze’ev and Zohary (1973) P. sativum ssp. humile sample #716 (the probable source of JI1794 of the John Innes Centre gene bank) that has featured in numerous studies ever since (e.g. Kosterin & Bogdanova, 2008; Palmer et al., 1985; Weeden, 2018), and in fact is the ‘type’ for all samples referred to, one way or another, as ‘northern humile’ forms. WS2 had no equivalent in Hellwig, Abbo, and Ophir (2021), most likely because it was represented by a single accession only. Individuals that formed C6 in Hellwig, Abbo, and Ophir (2021) were highly admixed in the current analyses and therefore have no WS# equivalent. C4 and C11 in Hellwig, Abbo, and Ophir (2021) merged into a single subpopulation, WS1, which holds two subgroups in the neighbor-net phylogenetic network and neighbor-joining tree. Overall, the genetic structure herein corroborates the genetic structure presented by Hellwig, Abbo, and Ophir (2021), Smykal et al. (2017) and Trněný et al. (2018). Yet, neither Smykal et al. (2017) nor Trněný et al. (2018) resolved southern humile as a distinct genetic group. It has been argued that P. sativum ssp. humile is not genetically distinct from P. sativum ssp. elatus and, therefore, both should be considered synonyms (Ellis, 2011; Vershinin et al., 2003). Those studies, however, were based on the pea collection of the John Innes Centre, which contains only four P. sativum ssp. humile accessions and in effect employed only one of those. Indeed, a number of reports on pea genetic diversity include only a few P. sativum ssp. humile samples or do not acknowledge this group at all (Ellis, 2011; Ellis et al., 1998; Jing et al., 2010; Vershinin et al., 2003). We are not quite sure about the reason for this under-representation of P. sativum ssp. humile in such studies. One possible reason may be the lack of genuine P. sativum ssp. humile samples. Another reason could be the difficulties that arise when samples from the wild P. sativum subgroups grown in common gardens must be distinguished based on morphological characters. For example, Kosterin et al. (2020) noted that growth conditions can strongly influence the morphology of wild peas, resulting in a different phenotype relative to the appearance of the same samples when observed in their native habitats.

The results of Hellwig, Abbo, and Ophir (2021) indicate that southern humile is indeed a distinct subgroup within the P. sativum complex, further corroborated herein with a larger sample size where southern humile forms a subgroup in central and south Israel, which diverged early from other P. sativum groups in all methods employed. This is evident from the admixture graph, where southern humile is separated from the other wild P. sativum groups by very high drift weights (Figure 4). This pattern did not change when domesticated samples were included (Figures 1 and 2). Likewise, Hellwig, Abbo, and Ophir (2021) suggest that northern humile is a distinct subgroup because 17 out of their 21 northern humile samples clustered together. This pattern was only partially confirmed herein. Although WS2 consisted of only northern humile individuals except for one accession (PeKv01 from northern Israel), this subpopulation most likely clustered together due the proximity of their sampling sites. This is expected given that wild peas do not grow in extensive stands but rather in patchy small populations that are rather isolated from each other (Ladizinsky & Abbo, 2015). Such a growth pattern causes close genetic relationships.
| Candidate | Group | Chromosome | Start BP | End BP | \(X'X\) outlier SNP | PCADOPT outlier SNP | Threshold | Gene ID | Description |
|-----------|-------|------------|----------|--------|---------------------|---------------------|-----------|---------|-------------|
| AC1       | *P. abyssinicum* | chr1LG6 | 34726784 | 34736784 | chr1LG6_34731784 | chr1LG6_34736023 | 0.975     | Psat1g025240 | Ion transport protein |
| AC2       | *P. abyssinicum* | chr1LG6 | 317327454 | 317347454 | chr1LG6_317342454 | chr1LG6_317342478 | 0.975     | Psat1g166440 | Translocase of chloroplast 159/132 + membrane anchor domain |
| AC3       | *P. abyssinicum* | chr7LG7 | 40893786 | 40903786 | chr7LG7_40898786 | chr7LG7_40897930 | 0.95      | Psat7g027480 | Frigida-like |
| SC1       | *P. sativum ssp. sativum* | chr1LG6 | 169226368 | 169236368 | chr1LG6_169231368 | chr1LG6_169231368 | 0.99      | Psat1g097640 | Aminotransferase class I and II |
| SC2       | *P. sativum ssp. sativum* | chr2LG1 | 55566791 | 55576791 | chr2LG1_55571791 | chr2LG1_55571987 | 0.95      | Psat2g036960 | PRONE (plant-specific rpo nucleotide exchanger) |
| SC3       | *P. sativum ssp. sativum* | chr1LG6 | 364300478 | 364310478 | chr1LG6_364305478 | chr1LG6_364305433 | 0.95      | Psat1g214880 | Unknown gene |
| SC4       | *P. sativum ssp. sativum* | chr2LG1 | 423066751 | 423076751 | chr2LG1_423071751 | chr2LG1_423071744 | 0.95      | Psat2g168480 | Protein kinase domain |
| SC5       | *P. sativum ssp. sativum* | chr3LG5 | 179799834 | 179809834 | chr3LG5_179804834 | chr3LG5_179804656 | 0.95      | Psat3g086760 | Cystatin-independent synthase + N-terminal domain |
| SC6       | *P. sativum ssp. sativum* | chr3LG5 | 343809183 | 343819183 | chr3LG5_343814183 | chr3LG5_343814183 | 0.95      | Psat3g157120 | RNA polymerase III RPC4 |
| SC7       | *P. sativum ssp. sativum* | chr4LG4 | 47132434 | 47142434 | chr4LG4_47137434 | chr4LG4_47137434 | 0.95      | Psat4g032760 | Elongation factor G + domain IV |
| SC8       | *P. sativum ssp. sativum* | chr6LG2 | 35511849 | 35521849 | chr6LG2_35516849 | chr6LG2_35516816 | 0.95      | Psat6g040560 | Filament-like plant protein + long coiled coil |
| SC9       | *P. sativum ssp. sativum* | chr1LG6 | 10397601 | 10407601 | chr1LG6_10402601 | chr1LG6_10402601 | 0.95      | Psat1g007080 | Unknown gene |
| SC10      | *P. sativum ssp. sativum* | chr2LG1 | 11666817 | 11676817 | chr2LG1_11673617 | chr2LG1_11673617 | 0.95      | Psat2g012600 | ThIF family |
| SC11      | *P. sativum ssp. sativum* | chr3LG5 | 20567490 | 20568490 | chr3LG5_20561490 | chr3LG5_20561354 | 0.95      | Psat3g007120 | MTHA/Hcf106 family |
| SC12      | *P. sativum ssp. sativum* | chr3LG5 | 181276862 | 181286862 | chr3LG5_181281862 | chr3LG5_181279060 | 0.95      | Psat3g093400 | Glycosyl hydrolases family 38 + N-terminal domain |
| SC13      | *P. sativum ssp. sativum* | chr3LG5 | 190441039 | 190451039 | chr3LG5_190446039 | chr3LG5_190446039 | 0.95      | Psat3g133440 | Subtilase family |
| SC14      | *P. sativum ssp. sativum* | chr3LG5 | 219492597 | 219502597 | chr3LG5_219497597 | chr3LG5_219497597 | 0.95      | Psat3g180600 | Zinc finger + C3HC4 type (RING finger) |
| SC15      | *P. sativum ssp. sativum* | chr3LG5 | 383674615 | 383684615 | chr3LG5_383679615 | chr3LG5_383679627 | 0.95      | Psat3g180600 | DNA gyrase/topoisomerase |
| SC16      | *P. sativum ssp. sativum* | chr4LG4 | 4763353 | 4773353 | chr4LG4_4763835 | chr4LG4_4765070 | 0.95      | Psat4g004440 | Phosphofructokinase |
| SC17      | *P. sativum ssp. sativum* | chr4LG4 | 122283125 | 122293125 | chr4LG4_122288125 | chr4LG4_122288125 | 0.95      | Psat4g073360 | Protein of unknown function (DFU1639) |
| SC18      | *P. sativum ssp. sativum* | chr4LG4 | 139118706 | 139128706 | chr4LG4_139132706 | chr4LG4_139132712 | 0.95      | Psat4g082120 | E1-E2 ATPase |
| SC19      | *P. sativum ssp. sativum* | chr4LG4 | 320237902 | 320247902 | chr4LG4_320242902 | chr4LG4_320243607 | 0.95      | Psat4g164440 | Chlorophyll a-b binding protein |
| SC20      | *P. sativum ssp. sativum* | chr4LG4 | 323790053 | 323800053 | chr4LG4_323804053 | chr4LG4_323804098 | 0.95      | Psat4g167240 | Flavon-Containing amine oxidoreductase |
| SC21      | *P. sativum ssp. sativum* | chr5LG3 | 83068696 | 83069696 | chr5LG3_83091896 | chr5LG3_83095287 | 0.95      | Psat5g045560 | WD domain + G-beta repeat |
| SC22      | *P. sativum ssp. sativum* | chr5LG3 | 192107960 | 192117960 | chr5LG3_192112960 | chr5LG3_192108081 | 0.95      | Psat5g107920 | PPR repeat |
| SC23      | *P. sativum ssp. sativum* | chr6LG2 | 22842705 | 22852705 | chr6LG2_22847705 | chr6LG2_22846833 | 0.95      | Psat6g029200 | F-box domain |
| SC24      | *P. sativum ssp. sativum* | chr6LG2 | 39567973 | 39577973 | chr6LG2_39572973 | chr6LG2_39572923 | 0.95      | Psat6g044120 | DNA gyrase/topoisomerase IV + subunit A |

(continued)
between individuals from the same location and was observed in *P. fulvum* (Hellwig, Abbo, Sherman, & Ophir, 2021) as well as in wild *P. sativum* (Smykal, Trněný, et al., 2018). The remaining northern *humile* samples, except PeAb28 (WS3) and PeAb37 (admixed), clustered together with *P. sativum* ssp. *elatius* samples (WS1). However, within WS1 there were two subgroups, one containing mostly *P. sativum* ssp. *elatius* samples and the other comprising only northern *humile* individuals, except for one *P. sativum* ssp. *elatius* sample (Figure 1). Whereas southern *humile* can be considered a distinct subgroup within the *P. sativum* complex, the case of northern *humile* seems to be not as clear as presented by Hellwig, Abbo, and Ophir (2021). The present results as well as those of Hellwig, Abbo, and Ophir (2021) suggest that northern and southern *humile* should not be placed into the same group as southern *humile* is genetically as distinct from northern *humile* samples as it is from the other wild *P. sativum* subgroups.

**Genetic diversity in domesticated pea**

Including samples from domesticated peas did not substantially change the genetic relationship between the wild *P. sativum* groups. *Pisum sativum* ssp. *sativum* samples clustered closest to the *P. sativum* ssp. *elatius* samples of WS1 (Figure 2). Contrary to Siol et al. (2017), who found a clear divergence between spring and winter peas and samples from the Middle East and Asia, we did not observe any subgroups within *P. sativum* ssp. *sativum*, and nor did Trněný et al. (2018) detect such subgroups. This may have been the result of the random SNP set used by both Trněný et al. (2018) and us herein, whereas Siol et al. (2017) used an SNP array that was designed for *P. sativum* ssp. *sativum*, and nor did Trněný et al. (2018) detect such subgroups. This may have been the result of the random SNP set used by both Trněný et al. (2018) and us herein, whereas Siol et al. (2017) used an SNP array that was designed for *P. sativum* ssp. *sativum*. Even though *P. sativum* ssp. *sativum* formed a clade distinct from WS1, several *P. sativum* ssp. *elatius* samples exhibited very close genetic relationships with *P. sativum* ssp. *sativum*. These samples originated from a rather broad geographic background (Figure S3) and may represent the stock that is genetically closest to the common ancestor of wild and domesticated *P. sativum*. It could be argued that the broad geographic background of these samples give support to the claim of Ben-Ze’ev and Zohary (1973) that several stocks may have contributed to the primary domesticated germplasm, similar to the situation described in rice (Choi et al., 2017). However, for the time being, this remains speculative and therefore the precise origin of the closest extant progenitor(s) of *P. sativum* ssp. *sativum* remains uncertain.

*Pisum abyssinicum* formed a distinct clade in all our analyses (Figures 2 and 3), while at the same time exhibiting a close genetic relationship with *P. sativum*, especially with southern *humile* and the clade containing *P. sativum* ssp. *sativum*, WS1 and WS3. In our neighbor-net
In our admixture plot, the lineages leading to *P. abyssinicum* and southern *humile* split early from other lineages and showed a close association of *P. abyssinicum* with southern *humile*. According to this analysis (Figure 4), southern *humile* received 72% of its genetic material from its likely common ancestor with *P. abyssinicum*. The basal position and close genetic relationship of *P. abyssinicum* and southern *humile* were also observed by Kreplak et al. (2019) as well as in phylogenetic trees based on a combination of nuclear, plastid and mitochondrial markers (Kosterin et al., 2010). *Pisum abyssinicum* and wild *P. sativum* samples from south Israel exhibited ‘combination A’ markers, which are also dominant in *P. fulvum* and are therefore considered as the ancestral marker combination (Kosterin et al., 2010). This pattern was recently corroborated by analyzing the mitochondrial genome, where the southern *humile* sample studied was grouped into the most basal clade, whereas *P. abyssinicum* fell into the second most basal clade together with other samples showing the A marker combination (Bogdanova et al., 2021). However, the plastid genome revealed a different structure (Bogdanova et al., 2021). Phylogenetic analysis using both the mitochondrial and the plastid genomes cluster *P. abyssinicum* together with *P. sativum* samples, contradicting studies (based on retrotransposon markers) that found *P. abyssinicum* to be a distinct group from both *P. fulvum* and *P. sativum* (Ellis et al., 1998; Jing et al., 2010; Vershinin et al., 2003). Based on gene sequence data, Weeden (2018) argued that if *P. sativum* ssp. *elatius* is not granted a species rank, *P. abyssinicum* should not be either, but rather be considered another subspecies among the *P. sativum* complex; in accord with the results of Trněný et al. (2018). Apparently, the use of different marker types results in different locations of *P. abyssinicum* in the phylogenetic trees. The pea genome seems to evolve faster relative to other studied legumes, especially through the high activity of transposable elements (Kreplak et al., 2019). Indeed, the nuclear genome size varies substantially between and within taxa (97.7–114.9% of the *P. sativum* ssp. *sativum* genome; Baraný et al., 1996). We, therefore, suspect that the discrepancy between the phylogenetic relationships estimated by different markers results from large structural variation in the nuclear DNA, which cannot be captured by all marker types. Whole-genome sequences could enable researchers to analyze such large structural variations and may shed light on this issue. Based on the results from mitochondrial and plastid genomes (Bogdanova et al., 2021), and simple sequence repeat (SSR) and SNP markers (herein, Kreplak et al., 2019; Trněný et al., 2018; Weeden, 2018), we endorse Weeden’s (2018) conclusion and opt to consider *P. abyssinicum* as a subgroup within the *P. sativum* complex. Still, we emphasize that this issue cannot be finally resolved yet.

A hypothesis regarding the possible hybrid origin of *P. abyssinicum* was first formulated by Govorov (1937). This view was later adopted by Kloz (1971) and reintroduced based on phylogenetic considerations (Ellis et al., 1998; Jing et al., 2010; Vershinin et al., 2003; Zaytseva et al., 2012). Although the emergence of *P. abyssinicum* from a cross involving *P. fulvum* was considered unlikely by Weeden (2018), an alternative option of a hybrid origin involving *P. sativum* ssp. *sativum* and *P. sativum* ssp. *elatius* was mentioned. However, an analysis of genetic diversity based on genome-wide SNPs did not yield any evidence of a hybrid origin of *P. abyssinicum* but rather suggested that it represents a distinct taxonomic entity (Trněný et al., 2018). Our subpopulation structure analyses are in line with the findings of Trněný et al. (2018) and we did not observe *P. abyssinicum* as a recipient of any admixture (Figures 2 and 3a). Moreover, all $f_3$(*P. abyssinicum*; *B, C*) statistics were significantly larger than zero (Table 1), which speak against the hypothesis of a hybrid origin of *P. abyssinicum*.

**Domestication of pea**

From our results, the possible feral origin of southern *humile*, as hypothesized from its affinity with disturbed habitats (Abbo, Lev-Yadun, et al., 2013; Ben-Ze’ev & Zohary, 1973), seems unlikely. Theoretically, single or very few back-mutations are required for an individual to acquire a feral trait (Scossa & Fernie, 2021; Wu et al., 2021), e.g. increased dormancy or pod dehiscence, which allows a fraction of its progeny to escape cultivation. If southern *humile* populations were indeed of feral origin, one would expect to see at least some similarity in their haplotype patterns at domestication loci, relative to the
domesticated groups, which is not the case for our candidates and likewise for Psat5g244000 (Figure 6). Therefore, a feral origin of southern *humile* seems unlikely. Southern *humile* inhabits areas in the far south of the Levant, which is a region that has been heavily impacted by human activities since the early days of the Neolithic Revolution, and the genetic variation of southern *humile* is strongly affected by the land cover, particularly by agricultural activities (Hellwig, Abbo, & Ophir, 2021). It is more likely that the described affinity with disturbed habitats is the result of natural adaptation to such habitats rather than a reflection of feral origin.

Several modes of domestication have been proposed for different crops. A single domestication event was reported for peanut (Bertoli et al., 2016) and *Zea mays* (maize) (Matsumuka et al., 2002). Domesticated lentils were probably derived in one domestication event (Ladizinsky, 1999). Yet other authors could not rule out that more than one wild *Lens orientalis* stock may have been involved (Liber et al., 2021). In *Glycine max* (soybean), a single domestication event is widely accepted but other models were suggested (for review, see Šedivy et al., 2017). Similar to the case in pea, where Braun (1841) already mentioned substantial morphological differences between *P. abyssinicum* and *P. sativum* ssp. *sativum*, there are two phenotypically distinct domesticated cultivar groups in *Cicer arietinum* (chickpea): *desi* and *kabuli*. Despite the large phenotypic variation between those types, chickpea has been domesticated only once and the *kabuli* type is considered to represent a polyphyletic group, the cultivars of which were repeatedly isolated in farmers’ fields through selection for brighter and larger seeds (Moreno & Cubero, 1978; Penmetsa et al., 2016; Rookwaiwal et al., 2014). Several studies lend support for two independent domestication events in *Phaseolus vulgaris* L. (common bean) (e.g. Gepts, 1988, 1990; Sonnante et al., 1994; Schmutz et al., 2014; for review, see Bellucci et al., 2014), with one stock originating in Mesoamerica and another stock originating in the Andes (Bitocchi et al., 2013). Two independent domestication events were also proposed for *Phaseolus lunatus* L. (Motta-Aldana et al., 2010; Serrano-Serrano et al., 2012) and *Vitis vinifera* (grapevine) (Sivan et al., 2021). However, Sivan et al. (2021) presented this model carefully, mentioning that a complete genetic turnover of the original Levan
tine domesticated stock cannot be ruled out solely based on their analyses. More complex models of domestication where multiple stocks from different regions were involved in a single domestication event were suggested for *Oryza sativa* (rice) (Choi et al., 2017). Moreover, geneflow from wild to domesticated stocks may have occurred and can complicate analyses of the domestication history of crops such as *Helianthus annuus* (sunflower) (Baute et al., 2015), maize (Hufford et al., 2013) and * Hordeum vulgare* (barley) (Civănă et al., 2021; Hübner et al., 2012). Based on the isolated position of *P. abyssinicum* in phylogenetic trees and clustering approaches, it was proposed that *P. abyssinicum* and *P. sativum* ssp. *sativum* were domesticated independently (Ellis et al., 1998; Kloz, 1971; Trněný et al., 2018; Zaytseva et al., 2012). Our results suggest that *P. abyssinicum* should be considered part of the *P. sativum* complex. Yet, *P. abyssinicum* diverged earlier from the lineage leading to *P. sativum* ssp. *sativum* that holds many wild pea samples (Figure 3, WS1 and WS2), congruent with a reconstruction regarding two genetically independent domestication events.

Besides the phylogenetic line of evidence, the genetic architecture underlying the domestication syndrome may help to clarify the mode of pea domestication. Domesticated crops often display similar changes in phenotypic traits that were changed by human selection (Donald & Hamblin, 1983; Hammer, 1984; Harlan et al., 1973). In a monophyletic domesticated lineage, it would be fair to assume that crucial traits (sensu Abbo et al., 2014) are controlled by the same set of genes. On the other hand, more than one genetically independent domestication event may well lead to a number of lineages in which different genes are involved in the respective phenotypic changes. Such a pattern was reported in *P. vulgaris*, where the Mesoamerican gene pool displayed different signatures of selection relative to the Andean gene pool (Schmutz et al., 2014). However, another study could only partially corroborate these results (Bitocchi et al., 2017). Similarly, only a few loci, putatively under selection in grapevine, were shared between the Levantine and Eurasian varieties (Sivan et al., 2021). European and Chinese apricot varieties displayed convergent phenotypic traits associated with domestication, yet different loci exhibited signatures of artificial selection (Groppi et al., 2021). Our results suggest a comparable scenario with the one in common bean, grapevine and *Prunus armeniaca* (apricot). Signatures of selection putatively associated with the domestication syndrome did not overlap between *P. abyssinicum* and *P. sativum* ssp. *sativum* (Figure 5). *Pisum abys.
pinicum* and *P. sativum* ssp. *sativum* samples did not cluster together, nor did they display similar haplotype patterns (Figures 6 and S4–S33) in any of the candidate windows detected, as would be expected under the assumption of a single domestication event. These results are in line with the observation of some degree of segregation in pod dehiscence in crosses between *P. abyssinicum* and *P. sativum* ssp. *sativum*, which suggests that this phenotype associated with domestication is controlled by different genes in the two cultivar groups (Holden, 2009). Unfortunately, none of the loci involved in the domestication syndrome in pea has yet been precisely localized (Smykal, Nelson, et al., 2018). Psat5g244000 has been reported to be a possible candidate for the *Dpo1* gene in *P. sativum* ssp. *sativum*, one of the
major genes involved in pod indehiscence in pea (Hradilová et al., 2017). This gene and the surrounding genomic region also displayed different haplotype patterns in *P. abyssinicum* and *P. sativum* ssp. *sativum* samples, which did not cluster together in the neighbor-joining cladogram based on this region. It should be noted that *Psat5g244000* has not yet been confirmed (Smýkal, pers. comm.). Once genes associated with the domestication syndrome in pea are cloned, our approach can be used to resolve the issue of the polyphyletic versus monophyletic origin of the domesticated pea groups.

The apparent distinct origin of *P. abyssinicum* and *P. sativum* ssp. *sativum* from the wild gene pool, as well as the different selection signals in the two cultivar groups, suggest that *P. abyssinicum* and *P. sativum* ssp. *sativum* were derived from two genetically independent domestication events. Suppose that *P. abyssinicum* was indeed domesticated independently from *P. sativum* ssp. *sativum*, this raises the question of the geographic origin of its domestication and from which wild stock it was derived. Ethiopia is known as a center of domestication and crop diversity (Harlan, 1971; Purugganan & Fuller, 2009; Vavilov, 1940; Zeven & Zhukovsky, 1975), yet, to the best of our knowledge, there are no reports of wild *Pisum* from east Africa. Therefore, it is unlikely that *P. abyssinicum* was domesticated within its present distribution range in east Africa. We propose a scenario similar to the one reported for apricots. European apricot cultivars originated from wild stocks in northern central Asia, from where they were disseminated westwards, whereas Chinese cultivars were domesticated in northern central Asia and dispersed eastwards (Groppi et al., 2021). Southern *humile* appears to be the genetically closest relative of *P. abyssinicum* (partially this study; Kosterin et al., 2010; Kreplak et al., 2019; Bogdanova et al., 2021), and hence may represent its extant closest relative. Under this scenario, *P. abyssinicum* may have been domesticated in the southern Levant and subsequently dispersed to east Africa. This is in line with reports of admixture events of Eurasian and east African humans, suggesting that humans migrated from the Near East to the Horn of Africa before, as well as after, the rise of agriculture (Hodgson et al., 2014; Llorente et al., 2015).
Figure 6. Neighbor-joining trees (only bootstrap values above 70%) (a) and haplotype patterns (b) of a selection of 6-Mbp windows containing candidate single-nucleotide polymorphisms (SNPs) under selection. The red shaded areas mark genes where the candidate SNPs are located.
It should be noted that this putative domestication event of *P. abyssinicum* must have happened after the establishment of the first Near Eastern farming communities in the northern Levant (Abbo & Gopher, 2017; Gopher et al., 2021). When these practices reached the southern Levant some 500–800 years later, it may have inspired the adoption of local wild *Pisum* stocks. Later, *P. abyssinicum* may have been abandoned in the southern Levant for an unknown reason, but retained in remote and isolated locales of the east African highlands.

**EXPERIMENTAL PROCEDURES**

**Plant material and genotyping**

The sample collection contained 494 pea stocks, which were sampled by ourselves or were obtained from gene banks. The sample set represented all major pea taxonomic groups, i.e. *P. abyssinicum* (40), *P. fulvum* (213); from Hellwig, Abbo, Sherman, & Ophir, 2021), *P. sativum* ssp. *elatius* (198) and *P. sativum* ssp. *sativum* (45). Additionally, the collection included one sample of *Vavilovia formosa* (Table S1). Most researchers consider *P. sativum* ssp. *elatius* as one diverse taxonomic entity. Based on morphological characteristics, we further subdivided this group into *P. sativum* ssp. *elatius* (103), *P. sativum* ssp. *humile*, including its two varieties, *P. sativum* ssp. *humile* var. *syriacum* (northern *humile*; 31) and *P. sativum* ssp. *humile* var. *humile* (southern *humile*; 62), following Ladizinsky and Abbo (2015). This enabled us to test whether *P. sativum* ssp. *humile* is indeed a distinct group within the wild *P. sativum* complex. The domesticated samples were selected based on growth region. We did not distinguish between landraces and modern material as our work addresses wild variability and the domestication episode. Accordingly, we did not explore evolutionary trajectories under domestication (crop evolution).

One plant from each accession was grown in a glass house in Rehovot (central Israel) during the summer of 2019. Temperatures were 17°C during the day and 12°C at night. Plants were irrigated daily to keep the soil moist. After the second true leaf appeared on all plants, a single tissue sample from either the apical meristem or a true leaf was taken from each plant for DNA extraction. DNA samples were sent to Elshire Group Ltd. (https://www.elshiregroup.co.nz), who conducted restriction site-associated DNA (RAD) sequencing using an adapted version of the protocol of Elshire et al. (2011). The adaptations were made to use 150-bp paired-end reads with combinatorial bar codes produced by the X Ten sequencing platform (Illumina, https://www.illumina.com) instead of single-end reads with single barcodes by HiSeq 2500 (Illumina). For the reduction of complexity, the restriction enzyme *ApeKI* was used. Demultiplexing was performed with *axe-semuix* (Murray & Borevitz, 2017). *GBS-PREPROCESS* (https://github.com/Lanilen/GBS-PreProcess) was used to trim adapter and reverse barcodes from the raw reads. The processed reads were mapped against the *P. sativum* ssp. *sativum* reference genome (*Krepelak et al., 2019*) and variants were afterward called with *STACKS 2.5* (Catchen et al., 2011, 2013).

Quality filtering was carried out with *VCFTools* 0.1.17 (Danecik et al., 2011). Only biallelic SNPs with MAFs of >0.01 and >40% genotype calls were retained. Also, SNPs that did not match the following criteria were removed: depth ≤ 3, mean depth across samples ≤ 5 and genotype quality ≤ 25. Pruning for LD was achieved using *PLINK* 1.90b5.4 (Purcell et al., 2007). Windows of 50 kb were shifted by five variants and SNPs with $r^2 \geq 0.5$ within one window were removed.

**Genetic structure**

The genetic structure of the sample collection was examined with different approaches. Sparse non-negative matrix factorization (sNMF; Frichot & François, 2015) was conducted in the *r-statistics* environment (R Core Team, 2021) using the package *usa* (Frichot et al., 2014). The cross-entropy criterion was calculated with 10% masked genotypes. *PCA* was performed using genopodaf genetic distances (Joly et al., 2015). A neighbor-net phylogenetic network was created using *SPLITTREE 4* (Huson & Bryant, 2005). The model of nucleotide substitution employed was Hasegawa–Kishino–Yano (Hasegawa et al., 1985) with transition/transversion ratio = 1.5, empiric base frequencies and the proportion of invariable sites = 0. The shape factor of the gamma distribution was set to 0.5 to consider a wide range of unequal rates among sites, because an SNP set as input is expected to violate the assumption of equal substitution rates among sites. A neighbor-joining tree was created based on Prevosti’s distance (Prevosti et al., 1975). Node support was estimated using 100 bootstrap values.

All four approaches were used on a sample set containing only wild peas and a sample set containing all samples to evaluate the genetic structure with and without domesticated samples separately.

**Admixture analysis**

We used ADMIXTOOLS 2 (https://uqrmaie1.github.io/admixtools/index.html) to construct admixture graphs based on Patterson’s $F$ statistics (Patterson et al., 2012) using the entire sample set. We used the *find_graphs* function to find admixture graphs that fit well to the data. This function requires several input parameters that cannot be easily determined a priori. Therefore, we explored the parameter space with 0 admixture events to test which block size, volume of missing data and MAF resulted in the highest graph likelihood. The likelihood was computed as out-of-sample scores ($S$), which are computed by dividing the difference of the residuals of the estimated $f_2$ statistics and the fitted $f_2$ statistics that are expected for the given admixture graph, by the inverse covariance matrix of the estimated $f_2$ statistics (https://uqrmaie1.github.io/admixtools/index.html). The $S$ scores allow the comparison of models of different complexity (i.e. different degrees of freedom). We tested block sizes ranging from 1000 to 10000 in steps of 100, maximum fractions of allowed missing genotypes ranging from 0 to 0.4 in steps of 0.1 and MAFs of 0.01 and 0.05. As the $S$ values can vary by chance, we calculated them 100 times for each parameter combination and compared the mean of those 100 $S$ values. There was no single best block size, but the results suggested that high $S$ values can be obtained with a block size between 8000 and 10 000. An MAF of 0.01 and no missing data showed the highest $S$ values. We therefore used a block size of 8000, an MAF of 0.01 and no missing data to construct six admixture graphs using the *find_graphs* function allowing 0-5 admixture events. $S$ values were calculated for each of those six graphs. We then performed the resampling of SNP blocks (100 bootstrap values) and calculated $S$ values for each resampled set. Based on the variation of $S$, we performed pairwise comparisons between all six models and calculated $P$ values to check which of the six graphs differ significantly from each other (Table S2). The graphs that best fit the data allowing 3-5 admixture events had the highest $S$ values but were not significantly different from each other. Following the parsimony principle, we chose the graph with three admixture events as the best fitting one.
A possible hybrid origin of \textit{P. abyssinicum} was tested with Pat-terson's \textit{f}_3 test (Patterson et al., 2012). We calculated \textit{f}_3(\textit{P. abyssini-cum}; B, C), where B and C represented \textit{P. fulvum}, \textit{P. sativum} ssp. \textit{sativum} and all subgroups of \textit{P. sativum} ssp. \textit{elatius}. With this procedure, we tested every possible combination of taxo-nomic groups as parents involved in a possible hybridization event leading to \textit{P. abyssinicum}. The \textit{f}_3(\textit{P. abyssinicum}; B, C) statistics were calculated with \textsc{admixtools} 2 (https://uqurmaie1.github.io/admixtools/index.html).

**Diversity statistics**

Nei’s gene diversity \((H_g;\text{Nei}, 1978)\), observed heterozygosity \((H_o)\) and nucleotide diversity \((\alpha)\) were calculated with \textsc{adegenet} (Jombart & Ahmed, 2011) and \textsc{vcftools} 0.1.17 (Danecek et al., 2011) for each \textit{Pisum} group and \textit{P. sativum} subpopulation.

**Genome-wide scans for selection signals**

We used two approaches to identify signals of selection in the two domesticated groups. The first one was \textsc{PCAdapt}, which, based on multivariate analysis, identifies signatures of selection by calculating the correlation of loci with the ordination axes. SNPs with extreme correlation were deemed outliers (Duforet-Frebourg et al., 2016). The \textsc{R} package \textsc{PCAdapt} was used for the analysis (Luu et al., 2017). The resulting \(P\) values were transformed to \(q\) values and SNPs below false discovery rates (FDRs) of 1.0%, 2.5% and 5.0% were considered outliers. Additionally, \(X'X\) statistics were calculated using \textsc{Baypass} 2.2 (Gautier, 2015). \textsc{Baypass} estimates the differentiation of two (or more) populations while accounting for the shared evolutionary history of these populations by adjusting allele frequencies using the covariance structure among the popu-lations. As we aimed to detect loci possibly involved in the domestication syndrome, we used the domesticated pea group as one population and wild \textit{P. sativum} accessions as the other population. As a result of the claim that southern \textit{humile} may be of feral origin, this group was excluded from the wild \textit{P. sativum} accessions. We randomly sampled 2000 SNPs per chromosome (14 000 SNPs in total) that were used to estimate the covariance matrix among populations. \textsc{Baypass} was run with parameters \textit{nral} = 5000, \textit{thin} = 20 and \textit{npilot} = 30. We created a POD by simu-lating 10 000 neutral SNPs to calibrate the \textit{XTX} statistics. The entire procedure was repeated three times to ensure that the covariance matrix was not biased by the sampling and the algo-rithm converged properly. As a threshold to determine \(X'X\) outliers, we used the 99%, 97.5% and 95% quantiles of the POD. SNPs were considered candidates if they passed the equivalent threshold in both analyses (e.g. 1% FDR in \textsc{PCAdapt} and 99% POD quantile in \textsc{Baypass}) and were located <5 kb apart from each other. The entire procedure was carried out for \textit{P. abyssinicum} and \textit{P. sativum} ssp. \textit{sativum} separately.

Linkage disequilibrium (LD) decay was estimated based on pair-wise \(r^2\) in 10-Mbp windows using the formula presented by Hill and Weir (1988). We extracted genomic regions surrounding the strongest candidate SNPs with a size of 6 Mbp, which was around the size when \(r^2\) dropped below 0.1 in \textit{P. sativum} ssp. \textit{sativum} and \textit{P. abyssinicum} (Figure S34). Additionally, we extracted a 6-Mbp window containing \textit{Psat}5g244000, which has been reported to be a possible candidate for \textit{Dpo1}, one of the major genes involved in pod indehiscence in pea (Hradilová et al., 2017). The SNPs within those windows were used to construct a neighboring tree based on Prevosti’s distance (Prevosti et al., 1975) to evaluate the genetic relationships of the samples in areas putatively containing loci under selection. To this end, we additionally visualized haplotype patterns within the windows using \textsc{Haplostats} 1.3 (Marnetto & Huerta-Sánchez, 2017). We used an \(r^2\) cut-off value of 0.1 to ensure that we do not miss any regions linked to the detected signal. \textsc{Haplostats} analysis (Figures 5b and S5–S33) is not affected by the size of the windows and, hence, increasing the LD cut-off would not benefit the detection of com-mon haplotypes in candidate regions. As \textsc{Haplostats} and \textsc{LD} analy-sis required phased genotypes as input, we used \textsc{Beagle} 5.1 (Browning & Browning, 2007; Browning et al., 2018) to phase the genotypes in our data. Following the suggestions of Pook et al. (2020), we used the following parameters in \textsc{Beagle} to improve the computations: burn-in = 10, iterations = 50, imp-states = 1000, imp-segment = 10, window = 70 and ne = 100 000.

**CONFLICT OF INTEREST**

All authors declare that they have no conflicts of interest associated with this work.

**DATA AVAILABILITY STATEMENT**

Supplementary files and sequence data are deposited in the Dryad Repository (https://doi.org/10.5061/dryad.x0k6djhkq).

**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online ver-sion of this article.

**Figure S1.** Distribution of SNP density over linkage groups.

**Figure S2.** Violin plots of distribution of missing genotype calls, observed heterozygosity, minor allele frequency and site nucleo-tide diversity of the entire SNP set and the LD pruned SNP set.

**Figure S3.** Map depicting sample locations of wild \textit{Pisum sativum} samples that are genetically closest to \textit{Pisum sativum} ssp. \textit{sativum}.

**Figure S4.** Neighbor-joining tree and haplotype patterns of the 6-Mbp window surrounding SC3.

**Figure S5.** Neighbor-joining tree and haplotype patterns of the 6-Mbp window surrounding SC4.

**Figure S6.** Neighbor-joining tree and haplotype patterns of the 6-Mbp window surrounding SC5.

**Figure S7.** Neighbor-joining tree and haplotype patterns of the 6-Mbp window surrounding SC6.
REFERENCES

Abbo, S. & Gopher, A. (2017) Near Eastern plant domestication: a history of thought. Trends in Plant Science, 22, 491–511. https://doi.org/10.1016/j.tips.2017.03.010

Abbo, S., Lev-Yadun, S. & Gopher, A. (2010) Agricultural origins: centers and noncenters; a near Eastern reappraisal. Critical Reviews in Plant Sciences, 29, 317–328. https://doi.org/10.1080/0735268X.2011.645428

Abbo, S., Lev-Yadun, S. & Gopher, A. (2012) Plant domestication and crop evolution in the near East: on events and processes. Critical Reviews in Plant Sciences, 31, 241-257. https://doi.org/10.1080/07352683.2011.645428

Abbo, S., Lev-Yadun, S., Heun, M. & Gopher, A. (2013) On the ‘lost’ crops of the neolithic Near East. Journal of Experimental Botany, 64, 815-822. https://doi.org/10.1093/jxbers/373

Abbo, S., Pinhasi Van-Oss, R., Gopher, A., Saranga, Y., Ohner, I. & Peleg, Z. (2014) Plant domestication versus crop evolution: a conceptual framework for cereals and grain legumes. Trends in Plant Science, 19, 351-360. https://doi.org/10.1016/j.tips.2013.12.002

Abbo, S., Zezak, I., Schwartz, E., Lev-Yadun, S. & Gopher, A. (2008) Experimental harvesting of wild peas in Israel: implications for the origins of Near East farming. Journal of Archaeological Science, 35, 922–929. https://doi.org/10.1016/j.jas.2007.06.016

Abbo, S., Zezak, I., Zehavi, Y., Schwartz, E., Lev-Yadun, S. & Gopher, A. (2013) Six seasons of wild pea harvest in Israel: bearing on Near Eastern plant domestication. Journal of Archaeological Science, 40(4), 2095–2100. https://doi.org/10.1016/j.jas.2012.12.024

Baranyi, M., Greihuber, J. & Švicecký, W.K. (1996) Genome size in wild Pisum species. Theoretical and Applied Genetics, 93, 717–721. https://doi.org/10.1007/BF00224067

Baute, G.J., Kane, N.C., Grassa, C.J., Lai, Z. & Rieseberg, L.H. (2015) Genome scans reveal candidate domestication and improvement genes in cultivated sunflower, as well as post-domestication introgression with wild relatives. New Phytopathol, 206, 630-638. https://doi.org/10.1111/nph.13285

Bellucci, E., Bitocchi, E., Rau, D., Rodríguez, M., Biagetti, E., Giardini, A. et al. (2014) Genomics of origin, domestication and evolution of Phaseolus vulgaris. In: Tuberosa, R., Graner, A. & Frison, E. (Eds.) Genomics of Plant Genetic Resources: Volume 1. Managing, sequencing and mining genetic resources. Dordrecht: Springer Netherlands, pp. 483-507. https://doi.org/10.1007/978-94-007-7672-5_20

Ben-Ze’ev, N. & Zohary, D. (1973) Species relation in the genus Pisum L. Israel Journal of Botany, 22, 73-91.

Berger, A. (1928) Classification of Pisum. In: Vegetables of New York. Part 1 — Peas. New York, USA: New York State Agricultural Experiment Station Report.

Bertollo, D.J., Cannon, S.B., Froenicke, L., Huang, G., Farmer, A.D., Cannon, E.K.S. et al. (2016) The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. Nature Genetics, 48, 438-446. https://doi.org/10.1038/ng.3517

Bitocchi, E., Bellucci, E., Giardini, A., Rau, D., Rodríguez, M., Biagetti, E. et al. (2013) Molecular analysis of the parallel domestication of the common bean (Phaseolus vulgaris) in South America and the Andes. The New Phytologist, 197, 300-313. https://doi.org/10.1111/nph.12437.x

Bitocchi, E., Rau, D., Benazzo, A., Bellucci, E., Goretti, D., Biagetti, E. et al. (2017) High level of nonsynonymous changes in common bean suggests that selection under domestication increased functional diversity at target traits. Frontiers in Plant Science. https://doi.org/10.3389/fpls.2016.02095

Bobkov, S.V. & Selikhova, T.N. (2017) Obtaining interspecific hybrids for introgressive pea breeding. Russian Journal of Genetics: Applied Research, 7, 145–152.

Bohdanova, V.S., Shatskaya, N.V., Mglinet, A.V., Kosterin, O.E. & Vasyliev, G.V. (2021) Discordant evolution of organellar genomes in peas (Pisum L.). Molecular Phylogenetics and Evolution, 160, 107136. https://doi.org/10.1016/j.ympev.2021.107136

Boissier, E. (1867). Flora Orientalis sive enumeratio plantarum. In: Graecia et Aegypto and Indiae fines hucusque observatarum. Critical Reviews in Plant Sciences, 31, 241-257.

Graecia et Aegypto and Indiae fines hucusque observatarum. Critical Reviews in Plant Sciences, 31, 241-257. https://doi.org/10.1016/j.jas.2012.12.024

Berger, A. (1928) Classification of Pisum. In: Vegetables of New York. Part 1 — Peas. New York, USA: New York State Agricultural Experiment Station Report.

Bertollo, D.J., Cannon, S.B., Froenicke, L., Huang, G., Farmer, A.D., Cannon, E.K.S. et al. (2016) The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. Nature Genetics, 48, 438-446. https://doi.org/10.1038/ng.3517

Bitocchi, E., Bellucci, E., Giardini, A., Rau, D., Rodríguez, M., Biagetti, E. et al. (2013) Molecular analysis of the parallel domestication of the common bean (Phaseolus vulgaris) in Mesoamerica and the Andes. The New Phytologist, 197, 300-313. https://doi.org/10.1111/nph.12437.x

Bitocchi, E., Rau, D., Benazzo, A., Bellucci, E., Goretti, D., Biagetti, E. et al. (2017) High level of nonsynonymous changes in common bean suggests that selection under domestication increased functional diversity at target traits. Frontiers in Plant Science. https://doi.org/10.3389/fpls.2016.02095

Bobkov, S.V. & Selikhova, T.N. (2017) Obtaining interspecific hybrids for introgressive pea breeding. Russian Journal of Genetics: Applied Research, 7, 145–152.

Bohdanova, V.S., Shatskaya, N.V., Mglinet, A.V., Kosterin, O.E. & Vasyliev, G.V. (2021) Discordant evolution of organellar genomes in peas (Pisum L.). Molecular Phylogenetics and Evolution, 160, 107136. https://doi.org/10.1016/j.ympev.2021.107136

Boissier, E. (1867). Flora Orientalis sive enumeratio plantarum. In: Graecia et Aegypto and Indiae fines hucusque observatarum. Critical Reviews in Plant Sciences, 31, 241-257.

Graecia et Aegypto and Indiae fines hucusque observatarum. Critical Reviews in Plant Sciences, 31, 241-257. https://doi.org/10.1016/j.jas.2012.12.024
Kloz, J. (1971) Serology of the Leguminosae. In: ChemotAXonomy of the Leguminosae. London: Academic Press, pp. 306-365.

Kosterin, O.E. & Bogdanova, V.S. (2008) Relationship of wild and cultivated forms of Pisum L. as inferred from an analysis of three markers, of the plastid, mitochondrial and nuclear genomes. Genetic Resources and Crop Evolution, 55, 735-755. https://doi.org/10.1007/s10722-007-9281-y

Kosterin, O.E. & Bogdanova, V.S. (2016) Reciprocal compatibility within the genus Pisum L. as studied in F1 hybrids: 1. Crossings involving P. sativum L. subsp. sativum. Genetic Resources and Crop Evolution, 62, 691-709. https://doi.org/10.1007/s10722-014-0189-z

Kosterin, O.E., Bogdanova, V.S. & Mglinet, A.V. (2020) Wild pea (Pisum sativum L. subsp. elatus (Bieb.) Aschers. et Graebn. s.l.) at the periphery of its range: Zagros Mountains, Wavilovski Zhurnal Genetiki I Selektcii, 24, 60-68. https://doi.org/10.18695/V2J.596

Kosterin, O.E., Zaytseva, O.O., Bogdanova, V.S. & Ambrose, M.J. (2010) New data on three molecular markers from different cellular genomes in Mediterranean accessions reveal new insights into phylogeography of Pisum sativum L. subsp. elatus (Bieb.) Schmalch. Genetic Resources and Crop Evolution, 57, 733-739. https://doi.org/10.1007/s10722-009-9511-6

Kreplak, J., Madouš, M., Cápal, P., Novák, P., Labade, K., Aubert, G. et al. (2019) A reference genome for pea provides insight into legume genome evolution. Nature Genetics, 51, 1411-1422.

Ladizinsky, G. (1996) Identification of the lentil’s wild genetic stock. Genetic Resources and Crop Evolution, 46, 115-118. https://doi.org/10.1023/A:100882618871

Ladizinsky, G. & Abb, S. (2015) The Pisum Genus. In: Ladizinsky, G. & Abb, S. (Eds.) The search for wild relatives of cool season legumes. Cham: Springer International Publishing, pp. 55-69. https://doi.org/10.1007/978-3-319-14505-3_3

Liber, M., Duarte, I., Maia, A.T. & Oliveira, H.R. (2021) The history of lentil (Lens culinaris subsp. culinaris) domestication and spread as revealed by genotyping-by-sequencing of wild and Landrace accessions. Frontiers in Plant Science, 12. https://doi.org/10.3389/fpls.2021.62439

Llorente, M.G., Jones, E.R., Eriksson, A., Siska, V., Arthur, K.W., Arthur, K.W., J.W. et al. (2015) Ancient Ethiopian genome reveals extensive Eurasian admixture in Eastern Africa. Science, 350, 820-822. https://doi.org/10.1126/science.aad2879

Luu, K., Bazin, E. & Blum, M.G.B. (2017) picadapt: An R package to perform genome scans for selection based on principal component analysis. Molecular Ecology Resources, 17, 67-77. https://doi.org/10.1111/1755-099X.12592

Makaseva, R.K. (1984) The pea. Rotterdam: A.A. Balkema.

Marnetto, D. & Huerta-S. (2017) Haplostrips: revealing population structure through haplotype visualization. Methods in Ecology and Evolution, 8, 1388-1392. https://doi.org/10.1111/2041-210X.12747

Matsuoka, Y., Vigouroux, Y., Goodman, M.M., Jesus Sanchez, G., Buckler, E. & Doebely, J. (2002) A single domestication for maize shown by multiple loci microsatellite genotyping. Proceedings of the National Academy of Sciences, 99, 6089-6094. https://doi.org/10.1073/pnas.05215189

Maxted, N. & Ambrose, M. (2001) Peas (Pisum L.). In: Current plant science and biotechnology in agriculture. Plant genetic resources of legumes in ecology and biotechnology in agriculture. Plant genetic resources of legumes in the domestication process in Cicer sp: Using SNP and QAT markers. PLOS ONE, 9, e102016. https://doi.org/10.1371/journal.pone.0102016

Sivan, A., Rahimi, O., Brus, J., Rathore, A., Bariotakis, M., Smaykal, P. & Lister, A., et al. (2017) Patterns of genetic structure and linkage disequilibrium in a large collection of pea germplasm. G3 Genes Genomes Genet- ics, 7, 2461-2471. https://doi.org/10.1111/g3.134741

Sivan, A., Rahimi, O., Lavi, B., Salmon-Divon, M., Weiss, E., Drori, E. et al. (2021) Genomic evidence supports an independent history of Levantine and Eurasian grapevines. Plants, People, Planet, 3, 414-427. https://doi.org/10.1002/ppp.3.10197

Smykal, P., Hradilová, I., Trnky, O., Bariotakis, M., Smaykal, P. & Lister, A., et al. (2017) Genetic diversity and macroecology of the crop wild relatives of domesticated pea. Scientific Reports, 7, 17384. https://doi.org/10.1038/s41598-017-17623-4

Smykal, P., Nelson, M.N., Berger, J.D. & Von Wettberg, E.J.B. (2018) The impact of genetic changes during crop domestication. Agronomy, 8, 119. https://doi.org/10.3390/agronomy8070119

Smykal, P., Trnky, O., Brus, J., Rathore, A., Bariotakis, M., et al. (2017) Genetic diversity and macroecology of the crop wild relatives of domesticated pea. Scientific Reports, 7, 17384. https://doi.org/10.1038/s41598-017-17623-4

Smykal, P., Trnky, O., Brus, J., Hanaček, P., Rathore, A., Roma, R.D. et al. (2018) Genetic structure of wild pea (Pisum sativum subsp. elatus) populations in the northern part of the Fertile Crescent reflects moderate cross-polination and strong effect of geographic but not environmental distance. PLOS ONE, 13, e0194056. https://doi.org/10.1371/journal.pone.0194056

Sonante, G., Stockton, T., Nodari, R.O., Becerra Velasquez, V.L. & Gepts, P. (1994) Evolution of genetic diversity during the domestication of common-bean (Phaseolus vulgaris L.). Theoretical and Applied Genetics, 89, 629-635. https://doi.org/10.1007/BF00222459
Tayeh, N., Aluome, C., Falque, M., Jacquin, F., Klein, A., Chauveau, A. et al. (2015) Development of two major resources for pea genomics: the GenoPea 13.2K SNP Array and a high-density, high-resolution consensus genetic map. *The Plant Journal*, 84, 1257–1273. https://doi.org/10.1111/tpj.13070

Trněný, O., Brus, J., Hradilová, I., Rathore, A., Das, R., Kopecký, P. et al. (2018) Molecular evidence for two domestication events in the pea crop. *Genes*, 9, 535. https://doi.org/10.3390/genes9110535

Vavilov, N.I. (1940) The theory of the origin of cultivated plants after Darwin. *Nauka (Science)*, 2, 55–75.

Vershinin, A.V., Allnutt, T.R., Knox, T.R., Ambrose, M.J. & Ellis, T.H.N. (2003) Transposable elements reveal the impact of introgression rather than transposition, in *Pisum* diversity, evolution, and domestication. *Molecular Biology and Evolution*, 20, 2067–2075. https://doi.org/10.1093/molbev/msg220

Weeden, N.F. (2018) Domestication of pea (*Pisum sativum* L.): the case of the Abyssinian pea. *Frontiers in Plant Science*, 9. https://doi.org/10.3389/fpls.2018.00515

Willcox, G. (2005) The distribution, natural habitats and availability of wild cereals in relation to their domestication in the Near East: multiple events, multiple centres. *Vegetation History and Archaeobotany*, 14, 534–541. https://doi.org/10.1007/s00334-005-0075-x

Wu, D., Lao, S. & Fan, L. (2021) De-domestication: an extension of crop evolution. *Trends in Plant Science*, 26, 560–574. https://doi.org/10.1016/j.tplants.2021.02.003

Zaytseva, O.O., Bogdanova, V.S. & Kosterin, O.E. (2012) Phylogenetic reconstruction at the species and intraspecies levels in the genus *Pisum* (L.) (peas) using a histone H1 gene. *Gene*, 504, 192–202. https://doi.org/10.1016/j.gene.2012.05.026

Zeven, A. & Zhukovsky, P. (1975) Dictionary of cultivated plants and their centres of diversity. 2nd edition. Wageningen: Pudoc.

Zong, X., Ford, R., Redden, R.R., Guan, J. & Wang, S. (2009) Identification and analysis of genetic diversity structure within *Pisum* genus based on microsatellite markers. *Agricultural Sciences in China*, 8, 257–267. https://doi.org/10.1016/S1671-2927(08)60208-4