RESEARCH ARTICLE

Evolution of fruit and seed traits during almond naturalization

Andrés Barea-Márquez[1] | Francisco J. Ocaña-Calahorro[1] | Rodrigo Balaguer-Romano[1,2] | José María Gómez[3,4] | Eugene W. Schupp[5] | Raquel Sánchez-Pérez[6] | Jesús Guillamón[6] | Joanna Zhang[1] | Rafael Rubio de Casas[1,4] (iD)

[1]Departamento de Ecología, Universidad de Granada, Granada, Spain; [2]Mathematical and Fluid Physics Department, Faculty of Sciences, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain; [3]Estación Experimental de Zonas Áridas, EEZA-CSIC, Almería, Spain; [4]Excellence in Research Unit "Modeling Nature" (MNat), Universidad de Granada, Granada, Spain; [5]Department of Wildland Resources and the Ecology Center, Utah State University, Logan, Utah, USA and [6]Departamento de Mejora Vegetal, Centro de Edafología y Biología Aplicada del Segura-Consejo Superior de Investigaciones Científicas, Murcia, Spain

Correspondence
Andrés Barea Márquez
Email: andbarmar@ugr.es

Rafael Rubio de Casas
Email: rubiodecasas@ugr.es

Funding information
Secretaría de Estado de Investigación, Desarrollo e Innovación, Grant/Award Number: CGL2016-79950-R

Handling Editor: Renske Onstein

Abstract

1. Cultivated plant species often naturalize and enter wild communities in a process known as feralization. To successfully feralize, crops must overcome ecological barriers and may undergo selection on certain traits, diverging phenotypically and genetically from their crop ancestors. In spite of the agronomic and ecological relevance of crop feralization, the eco-evolutionary dynamics driving it remain understudied.

2. In this paper, we evaluated phenotypic and genotypic differentiation in fruit and seed traits during the naturalization of the almond tree (*Prunus dulcis* (Mill.) D.A. Webb) in SE Iberia and evaluated the potential role of natural selection in this process. To do so, we investigated the patterns of genetic divergence between cultivated and feral populations using functional (the cyanogenesis *Sk* gene) and neutral (17 SSR loci) markers and analysed morphological and biochemical traits in kernels of 342 individuals from 15 cultivated and 24 feral populations.

3. We detected very little genetic differentiation in neutral markers between cultivated and feral populations. The majority of the observed genetic variation was due to differences within each type. Conversely, the recessive allele *sk* responsible for seed toxicity was significantly more frequent in feral populations. Phenotypic differentiation between cultivated and naturalized almond populations was also significant. Feral almond kernels were smaller and lighter, had denser and more resistant shells (endocarps) and more toxic seeds. Selection analyses indicated that these genetic and phenotypic patterns might be driven by directional selection on fruit and seed traits, potentially linked to defence against predation.

Andrés Barea Márquez and Francisco J. Ocaña-Calahorro are equally contributed authors.

------

4. *Synthesis*. Our findings indicate that almond naturalization is consistent with strong directional selection on fruits and seeds, leading to smaller and more toxic seeds encased in harder endocarps. Accordingly, we propose that feralization of this crop is, at least to some degree, driven by adaptive evolution of dispersal and recruitment traits.

1 | INTRODUCTION

The domestication of crops has been ongoing since the beginning of the Holocene, determining the characteristics of cultivated plant species. The transition towards domesticated breeds involved a series of changes that resulted in phenotypes that could be easily harvested (Allaby, 2014; Larson et al., 2014; Milla et al., 2015). For example, in fruit crops, there have been changes in the relative size of edible parts, which have increased in size and become more accessible (Browicz & Zohary, 1996; Zohary & Spiegel-Roy, 1975).

Crop domestication has been documented with detail at archaeological, botanical and genetic levels. This abundance of information has shown that (a) it is not a linear process towards a domesticated 'ideotype', (b) it is not a distinct event but rather a gradient with an indefinite beginning and no clear endpoint and (c) the transition between wild and cultivated strains is not unidirectional and phenotypic changes can be reversible (Cornille et al., 2012; Dickmann et al., 1994; Larson et al., 2014; Richardson & Rejmánek, 2011). Therefore, domestication can only be fully comprehended considering its inverse process, that of de-domestication or naturalization of cultivated breeds, and the role that it has played in the evolution of cultivated varieties (i.e. cultigens; Gross et al., 2014; Wu et al., 2021).

However, the information available about crop naturalization is very limited. Scientists have studied a small number of cases, such as weeds that have emerged from the naturalization of specific cereal cultigens (Gering et al., 2019; Wu et al., 2021). There has also been interest in studying the process of gene flow between cultivated and wild strains that can produce weeds that are difficult to eradicate due to their similarity to cultivated breeds and can facilitate the propagation of genetically modified (GM) plants (Ellstrand, 2018). Nevertheless, the cases in which naturalized breeds do not pose management challenges have been almost entirely unstudied, even though they could be highly valuable for breeding purposes (Gering et al., 2019; Wu et al., 2021).

Although incomplete, the evidence obtained suggests that naturalization involves to some degree a reversion to the wild phenotype, even if it does not necessarily lead to changes in the same genic regions. Research has also shown that naturalization can be an extremely rapid process even though the genetic diversity of domestic species tends to be significantly reduced relative to that of

their direct wild ancestors (Gering et al., 2019; Zhang et al., 2017). As a consequence, selection must draw from a comparatively restricted genetic pool during feralization, providing an intriguing case study in evolutionary ecology.

The rapid transition from the cultivated to the feral (wild) state is conditioned by selection on pre-existing traits, which will ultimately determine the phenotypic and genetic characteristics of naturalized populations (Pannell et al., 2015). Since response to selection under feralization depends on the standing diversity of the crop, it can be expected to be proportional to the diversity harboured in the cultivated pool, which is highest in heterozygous and polymorphic crops such as self-incompatible trees (Miller & Gross, 2011). In addition to genetic limitations, ecological filters can also determine the outcome of feralization events. Blackburn et al. (2011) defined naturalization as part of a continuum of multiple phases, each characterized by specific selective pressures and distinct environmental factors. Accordingly, an introduced plant species can only effectively colonize a new area if there is sufficient seed dispersal, germination, and development of seedlings and juveniles (Blackburn et al., 2011; Richardson et al., 2000). In other words, successful seed dispersal and establishment constitute the first prerequisites for feralization. Therefore, traits implicated in seed dispersal and germination are likely to be under selection during naturalization.

The almond tree has been described as undergoing a process of feralization in SE Spain (Balaguer-Romano et al., 2021; Homet-Gutiérrez et al., 2015). Even though wild species closely related to the almond (*Prunus* subgen. *Amygdalus*) are not found in this region, the spontaneous recruitment of naturalized individuals near farms is a frequent phenomenon (Homet-Gutiérrez et al., 2015; Ruiz de la Torre, 2006). The establishment of these feral individuals of *P. dulcis* is the result of direct feralization from cultivated groves (i.e. they are endo-feral, *sensu* Ellstrand et al., 2010). Their presence near or even inside orchards has been traditionally accepted to some extent, as they serve as pollen donors for the mostly self-incompatible cultigens, as a source of rootstocks and as easily identifiable boundary markers in land cultivated with other crops such as olive trees (Ibancos Nuñez & Rodriguez Franco, 2010; Rubio-Cabetas, 2016; Ruiz de la Torre, 2006). In spite of this long history of tolerance of occasional semi-wild individuals, dense, self-sustaining feral populations have only been described recently (Balaguer-Romano et al.,

2021; Homet-Gutiérrez et al., 2015) and the ecological and evolutionary dynamics associated with their emergence remain poorly understood.

Taking into account the morphological characteristics of the almond, seed dispersal mechanisms are likely restricted to gravity (barochory) and synzoochory. The latter is a dispersal strategy mediated by granivores that actively transport seeds and store them (frequently underground) for later consumption (Gómez et al., 2019). Previous results by our research team confirmed the role of vertebrate vectors and synzoochory in the naturalization of almonds (Balaguer-Romano et al., 2021). However, it is important to note that in this process animal dispersers act primarily as seed predators. Thus, species dispersed by synzoochory must ensure the survival and germination of the seed, which might lead to the selection of protective features such as the development of hard shells or targeted toxicity (García et al., 2005; Struempf et al., 1999; Vander Wall & Beck, 2012).

Seed toxicity is considered an important trait for increasing seed survival in the face of potential predation (Freeland & Janzen, 1974). Seeds of *Prunus* spp. contain amygdalin, a highly toxic substance for mammals, including humans. Amygdalin is a cyanogenic diglucoside that is synthesized systemically from prunasin and accumulates in the cotyledons of seeds. Once ingested, it is hydrolysed, releasing glucose, benzaldehyde (responsible for the bitter taste) and hydrogen cyanide, a harmful substance that inhibits cellular respiration (Sánchez-Pérez et al., 2012). The almond tree (*P. dulcis*) is the only species of the genus cultivated for its 'sweet' seeds. In this species, amygdalin production appears to be a Mendelian trait, controlled by the *Sweet kernel* (*Sk*) gene. The sweet phenotype is produced by the dominant form of this locus such that homozygous dominant individuals (*SkSk*) and heterozygotes (*Sksk*) present this characteristic (Heppner, 1923, 1926; Sánchez-Pérez et al., 2019). Since the consumption of relatively small quantities of bitter almonds (as few as a couple dozen) can be lethal (Ladizinsky, 1999), almond domestication and cultivation is inextricably and exclusively linked to the non-toxic ('sweet') genotypes (Browicz & Zohary, 1996; Delplancke et al., 2013; Ladizinsky, 1999; Sánchez-Pérez et al., 2019). However, the frequency of the dominant allele in wild populations must be extremely low, to the point that it has never been found in a wild individual (a puzzling problem that was termed the 'riddle of almond domestication' by Ladizinsky, 1999) perhaps because toxic seeds are strongly favoured by predator-mediated natural selection.

Beyond seed toxicity, predation resistance or predator deterrence might involve other protective traits, such as harder endocarps or shells that increase handling costs and promote caching rather than immediate consumption and that deter animals that act solely as seed predators (García et al., 2005; Vander Wall, 2001, 2010). Simultaneously, selection might have favoured less attractive kernels, with smaller sizes and/or lower nutritional value of the seed (Gómez, 2004). However, it has also been argued that the evolution of synzoochory depends on the production of large and nutritious seeds that are desirable to granivores and stimulate caching (Vander Wall, 2001, 2010) which suggests the existence of potentially conflicting selective pressures. Both seed size and kernel (endocarp) resistance have been described as highly heritable in almond (Dicenta et al., 1993; Spiegel-Roy & Kochba, 1981). In fact, both traits might even be monogenic; endocarp resistance has been proposed to be controlled by the gene D-Q (Sánchez-Pérez et al., 2007) while recent genome-wide studies have identified a single marker with a significant association with seed weight (Pavan et al., 2021). Given the high heritability of these three traits (seed toxicity, kernel resistance and seed weight) and the high diversity of cultivated populations (Socias i Company & Felipe, 1992), a sufficiently strong selection during feralization might cause a shift in phenotypic means, even over a relatively low number of generations.

In this study, we attempted to determine whether almond feralization involves phenotypic divergence in fruit and seed traits between cultivated and naturalized populations and whether this divergence might be indicative of natural selection. For this purpose, we examined (a) whether naturalization implies genetic divergence between cultivated and feral populations of *P. dulcis* in the Iberian Peninsula; (b) the existence of quantifiable phenotypic differences between cultivated and feral populations, particularly in fruit and seed traits expected to be implicated in synzoochorous interactions, and (c) to what extent phenotypic differences can be attributed to direct selection on specific traits, namely kernel resistance, seed toxicity and weight.

2 | MATERIALS AND METHODS

2.1 | Plant samples and study area

The data in this study were obtained between 2017 and 2018 from a total of 39 populations (15 cultivated and 24 feral populations) in southern Spain (see Supplementary Methods for details). We approximated the age of feral populations by comparing photogrammetric images (Plan Nacional de Ortofotografía Aérea, 1956–2017, National Geographic Institute of Spain, IGN) and setting as baseline the earliest date when previous or coetaneous almond orchards were observed in the area and land use was compatible with the establishment of spontaneous woody vegetation (i.e. not occupied by crops or other features such as roads; Table 1; Supplementary Methods; Figure S1). We visited these populations during the period of natural dispersal (which is also the harvest season) between September and the beginning of October. In each population, we randomly sampled between 10 and 12 reproductive individuals at least 10 m apart. There were not enough reproductive trees in some feral populations and, in these cases, we took samples from all of the reproductive trees. From each tree, we randomly collected 25–30 fruits around the perimeter of the crown (or all of the available almonds in the case that the tree had <25 fruits). We also sampled 3–5 leaves from each individual and stored them in silica gel for later DNA extraction (Table 1).

TABLE 1 Populations included in this study and plant material collected in each case. The table includes the type (Cultivated/Orchard vs. Feral), year when samples were taken, coordinates and number of trees sampled. It also indicates an estimated date of establishment of each population, which was determined using aerial photographs, comparing the temporal sequence of pictures obtained from periodic photogrammetric flights between the years 1945 and 2017 (see text for details)

Location	Type	N (leaf)	N (fruits)	Year	Coordinates	Altitude (m)	Establishment
Alhama[S]	Feral	12[1]	12[2,3CU]	2018	37°01′27.0″N 3°57′45.2″W	880	After 2004
Alhama	Cult.	12[1]	12[2,3CU]	2018	37°01′24.7″N 3°58′13.5″W	870	After 1986
Barranco de Viznar[S]	Feral	12[1]	12[2,3C]	2018	37°13′07.5″N 3°33′44.4″W	945	1973–1986
Barranco de Viznar[A]	Cult.	12[1]	12[2,3C]	2018	37°13′05.5″N 3°33′48.7″W	945	1956–1976
El Burgo[S]	Feral	6	7[3U]	2017	36°46′27.8″N 4°56′16.6″W	590	1998–2003
El Burgo	Feral	2	5[3U]	2017	37°47′31.5″N 4°39′21.4″W	640	After 1980
El Burgo[S]	Feral	0	12[3U]	2017	36°48′57.4″N 4°57′5.7″ W	590	After 2005
El Burgo[A]	Cult.	0	4[3U]	2017	36°46′28.8″N 4°56′23.0″W	590	After 1956
El Burgo	Cult.	0	4[3U]	2017	37°47′31.5″N 4°39′21.4″W	640	After 1980
Campillos[S]	Feral	12	10	2017	37°01′18.1″N 4°56′29.9″W	530	After 1986
Campo Camara[S,R]	Feral	12	12	2017	37°40′10.0″N 2°48′21.6″W	800	1986–1998
Carvajales[R]	Feral	12	12	2017	37°10′44.6″N 4°40′56.9″W	480	After 1998
Cerro de San Miguel*,[A,S]	Feral	12[1]	8[2,3C]	2017	37°11′38.3″N 3°34′56.4″W	903	1973–1986
Chirivel*,[S]	Feral	5[1]	3[2,3C]	2018	37°35′06.9″N 2°15′25.2″W	1050	After 1986
Cubillas	Feral	12[1]	3	2017	37°16′37.4″N 3°39′46.5″W	650	1986–1998
Deifontes[A]	Cult.	8[1]	8[2,3CU]	2017	37°19′18.1″N 3°34′07.1″W	1000	Before 1956
Deifontes	Feral	9[1]	6[2,3CU]	2017	37°19′55.2″N 3°33′21.7″W	990	After 1980
Don Fadrique*,[R]	Feral	10	10	2017	37°54′06.8″N 2°25′41.5″W	1040	After 1956
Gor[S,R]	Feral	12	12[3U]	2017	37°23′28.2″N 3°01′07.2″W	1200	1998–2003
Gor	Cult.	0	5[3U]	2017	37°23′24.9″N 3°01′08.2″W	1200	After 1973
Grazalema*,[S]	Feral	0	12	2017	36°45′25.7″N 5°22′0.2″W	850	After 1998
Guadix	Feral	12[1]	12[2,3C]	2018	37°19′50.5″N 3°03′12.4″W	1175m	After 1986
Guadix	Cult.	12[1]	12[2,3C]	2018	37°19′47.1″N 3°03′19.8″W	1175m	After 1986
Haza del Lino	Cult.	12[1]	12[2,3CU]	2017	36°48′38.8″N 3°17′46.7″W	1200	After 1980

(Continues)

TABLE 1 (Continued)

Location	Type	N (leaf)	N (fruits)	Year	Coordinates	Altitude (m)	Establishment
Haza del Lino[S]	Feral	11[1]	9[2,3CU]	2017	36°48′38.8″N 3°17′46.7″W	1220	After 1980
Orce[R]	Feral	6	6[3U]	2017	37°42′58.3″N 2°27′39.1″W	1000	1973–1986
Orce	Cult.	4	4[3U]	2017	37°42′58.3″N 2°27′39.1″W	1000	1973–1986
Polopos[A]	Cult.	12[1]	12[2,3CU]	2017	36°47′51.8″N 3°17′33.3″W	750	After 1973
Polopos[S,R]	Feral	12[1]	11[2,3CU]	2017	36°47′51.8″N 3°17′33.3″W	750	After 1973
Santa Fe[R]	Feral	12	12[3U]	2017	37°09′44.4″N 3°43′28.0″W	610	After 1998
Santa Fe	Cult.	0	3[3U]	2017	37°09′44.4″N 3°43′28.0″W	610	After 1956
Serrato[S]	Feral	12	11	2017	36°52′55.1″N 4°57′19.9″W	650	After 1980
Serrato	Cult.	0	3	2017	36°52′55.1″N 4°57′19.9″W	650	After 1980
Sierra Maria[R]	Feral	7[1]	7[2,3C]	2017	37°42′05.6″N 2°11′03.0″W	1240	After 1973
Sierra Maria	Cult.	0	5	2017	37°42′05.6″N 2°11′03.0″W	1240	After 1973
El Torcal	Cult.	12	12[3U]	2017	37°00′15.1″N 4°33′32.2″W	590	Between 1973–1986
El Torcal	Cult.	6	6[3U]	2017	37°00′14.4″N 4°33′45.6″W	650	After 1980
El Torcal[S]	Feral	12	12[3U]	2017	37°00′12.8″N 4°33′37.4″W	590	After 1998
Vereda de la Estrella*,[S]	Feral	12[1]	12[2,3CU]	2017	37°07′41.2″N 3°21′42.4″W	1250	N/A

[1]Included in genetic studies (both SSR & SNP), [2]Used for shell resistance analyses, [3]Cyanogenesis (toxicity; C—colorimetric; U—UPLC), [A]Abandoned Orchards, [S]Feral populations on steep terrain (≥30% slope), [R]Feral populations on abandoned land or along roads or trails, *These feral populations were adjacent to almond orchards that could not be sampled.

2.2 | Genetic characterization of neutral (SSR) and functional (gene *Sk*) markers

We analysed the genetic structure of the populations using neutral markers (SSRs) and a marker that determined the presence of amygdalin in the seed (SNP of allele *Sk*, *Sweet kernel*). We extracted DNA from leaves collected from 17 populations, 16 of which were also used for the colorimetric quantification of cyanide content (5–6 individuals per population; Table 1) with the FavorPrep Plant Genomic DNA Extraction Mini kit (FAVORGEN, Taiwan). We carried out SSR genotyping in the CRAG (Center of Agrigenomic Research, CSIC-IRTA-UAB-UB) using the markers described in Sánchez-Pérez et al. (2007; Table S1). Genotyping via SNPs of the *Sk* gene was carried out in the CEBAS-CSIC following the protocol described in Sánchez-Pérez et al. (2019). About 50 ng same DNA extracted for each of the samples was used to amplify the *Sk* gene by PCR with the primers *bHLH2F-BamHI* (CACCGCGGATCCGAATGGAAGAGATCATAGCCTCAT) and *bHLH2RXhoI* (GATCCACTCGAGCTAGTTGTACCACCTTTTTATAAT)

with the Phusion High-Fidelity polymerase (Thermo Scientific) with the following conditions: 2 min at 98°C, 35 cycles of 20 s at 98°C, 20 s at 60°C and 1 min at 72°C, and one cycle of 5 min at 72°C. SNP detection was performed by the 3500 Genetic Analyzer, Applied Biosystems, by the use of the sequencing primer *bHLH2_763F* (AAGAGGGTGATACAAAAGAAGC, with a Tm of 60°C), at the CAID (University of Murcia).

2.3 | Phenotypic characterization

To study potential differences between cultivated and feral plants in the characteristics of fruits and seeds, we analysed the size of kernels (volume, in-shell [endocarp + seed] mass and seed mass), their mechanical protection (thickness, density and resistance of the endocarps) and toxicity (amygdalin/cyanide content of seeds). For the morphological measurements of the kernels, we used 2707 fruits, pooled from all of the study populations. To measure endocarp

resistance, we used 348 samples from a subset of 16 populations (6 cultivated and 10 feral, Table 1), including 4–5 individuals per population and 4–5 almonds per individual. For the quantification of amygdalin content, we used 134 seeds from 30 populations (13 cultivated and 17 feral), utilizing when possible the same seeds that had been used in previous measurements. When the same seeds were not available (e.g. because they were destroyed when cracking open the endocarp or during kernel resistance measurements), we used other almonds from the same tree if possible or the same population, ensuring in every case a consistent level of diversity (at least 2–3 individuals per population, 2–3 almonds per individual).

2.3.1 | Morphology of the fruit and seed

We dried an aliquot derived from four to five almonds from each individual (total $n = 2707$) at 50°C for 72 h. Once they were dehydrated, we took morphological measurements on the entire (in-shell) almond (a: length, b: width, c: height and M: mass). Afterward, we extracted the seeds from the endocarps and performed the same measurements on them. To estimate the volume of the almonds, we approximated their shape as an ellipsoid ($V = 4/3\pi abc$). We then estimated the density of the endocarp by dividing the difference between the in-shell mass and the unshelled seed mass by the corresponding difference in volume (i.e. $[M_{\text{in-shell}} - M_{\text{seed}}]/[V_{\text{in-shell}} - V_{\text{seed}}]$). We estimated the thickness of the endocarp by halving the difference in height between the almond with and without a shell ($[c_{\text{in-shell}} - c_{\text{seed}}]/2$). Based on the morphological results, we chose a subgroup of almonds that covered the gradient of endocarp density to estimate resistance (Table 1). In total, we subjected 348 almonds to endocarp-crushing experiments using a hydraulic press (S.A.E. IBERTEST model 1BTH-2730) to quantify cracking load.

2.3.2 | Cyanide content

To study almond toxicity, we used a modified extraction protocol and spectrophotometer assessment of cyanide using picric acid as an indicator reagent (Oshima et al., 2003; Supplementary Methods). We quantified cyanide content of 10 feral and six cultivated populations, randomly choosing three almonds from three individuals in each population (i.e. nine almonds per population, $n = 134$; Table 1). We used liquid nitrogen to grind the almond seeds into a fine, homogeneous powder that we stored at −20°C. Then, digested the almond samples with citric and tartaric acid. We quantified the amount of liberated cyanide using paper strips stained with picric acid. We washed the paper strips with ethanol and 24 h later, quantified the optical density of the fluid (proportional to the cyanide content) in a Tecan NanoQuant Infinite M200 spectrophotometer. In parallel, and as a calibration method, we determined the amygdalin content of 82 almonds (50 cultivated, 32 feral; Table 1) through ultra-performance liquid chromatography (UPLC) coupled to mass spectrometry following the protocol established by Arrazola et al. (2013) with a Waters Acquity UPLC system interfaced to a triple quadrupole mass spectrometer Xevo TQS (Waters). We carried out these analyses at the University of Granada's Center of Scientific Instrumentation (CIC-UGR).

3 | DATA ANALYSIS

3.1 | Genetic differentiation

Genetic differences among populations and between cultivation types were estimated with an analysis of the partition of molecular variance (AMOVA) in neutral (SSR) markers of all sequenced populations and orchards. We also used these data to compute an overall fixation index (Fst) and to construct a phylogenetic neighbour joining tree using Nei's genetic distances among populations (Nei, 1972). Population genetic analyses were performed with the packages 'hierfstat', 'poppr' and 'vegan' (Goudet, 2005; Kamvar et al., 2014; Oksanen et al., 2019). Additionally, we used functional genetic data (differences in the *Sk* allele) to estimate changes in cyanogenesis associated with feralization. We modelled the effect of population and cultivation type on cyanogenesis with a GLM fitted to a binomial negative distribution, including *Sk* genotype as a factor and cyanide content as the response variable. We also used the functional genetic data to determine whether there are differences in the distribution of *Sk* genotypes between feral and cultivated populations. To do this, we generated a contingency table showing genotype against cultivation type and evaluated the correlation between the two variables with a chi-square test.

3.2 | Phenotypic differences

To determine whether the phenotypic differences among almond populations were affected by cultivation type, we carried two sets of analysis of variance for each of the seed and fruit traits. First, we used all the data available to study the overall differences across all feral and cultivated almonds. Then, we performed two sets of similar analyses first using only the six paired populations included in genetic studies and then limiting the sampling to the subset of four populations that were identified as sister OTUs of their adjacent orchards in the phylogenetic tree. In these two last cases, we compared only the three highly heritable traits that were also the focus of our selection analyses (seed toxicity, endocarp resistance and seed weight) using endocarp density as a proxy for resistance, because we did not have resistance data available for all populations and both traits are strongly positively correlated ($R \approx 0.7$; p-value $= 2.2 \times 10^{-16}$). In these analyses, we used models with two predictors: almond 'type' (feral or cultivated) and the population of origin. In each case, we fit linear models that included only 'type' or both predictors, carrying out transformations to satisfy the criteria of parametric modelling when necessary. Furthermore,

we fit mixed models that considered population as a random factor, choosing the distribution function and the link function best suited for each response variable using the functions 'descdist' and 'fitdist' in the packet 'fitdistrplus' in R (Delignette-Muller & Dutang, 2015). In each case, we carried out preliminary explorations of model fit comparing the difference between the residual deviance and the residual degrees of freedom through a chi-squared test and selected the model with the lowest AIC value. When two models had comparable AIC values (six units or less) we chose the simplest, less parameterized model, considering the simplest model of all a uni-factorial linear model (the transformations applied to the data in each case and the models fitted are indicated in Table S3).

3.3 | Pst–Fst comparison

We compared the degree of neutral genetic differentiation (estimated with SSR markers) to that of phenotypic differentiation with the aim of detecting potential signals of selection in three traits: seed weight, endocarp resistance and seed amygdalin concentration (cyanogenesis) which are expected to be responsive to selection due to their high heritability (Dicenta et al., 1993; Heppner, 1923, 1926; Spiegel-Roy & Kochba, 1981). To infer whether phenotypic change is driven by selection rather than by neutral dynamics, we compared the fixation index (Fst) derived from neutral markers to Pst, an index that estimates proportion of variation in a quantitative trait caused by genetic differences, analogous to Qst. Fst is driven mostly by genetic drift and gene migration, while indexes such as Qst also incorporate the effects of selective dynamics on the phenotype. Consequently, differences between the two indices can be attributed to selection (Leinonen et al., 2013). If Qst > Fst, genetic divergence in the trait exceeds neutral expectations, which can be associated with directional selection. If Qst = Fst, neutral divergence cannot be ruled out as the cause of phenotypic divergence. Lastly, Qst < Fst can be taken as indication that phenotypes are diverging less across populations than expected under a neutral scenario, likely due to stabilizing selection (Leinonen et al., 2008; Merilä & Crnokrak, 2001). Since studying quantitative genetic variance in reproductive traits of wild (or feral) trees is very complex, we approximated Qst with the index of phenotypic divergence across populations Pst (Brommer, 2011) calculated as:

$$Pst = \frac{\sigma^2_{GB}}{c\sigma^2_{GB} + 2h^2\sigma^2_{GW}},$$

where σ_{GB} is the genetic variance between populations, σ_{GW} is the variance within populations, c is the proportion of variance due to differences caused by additive genetic effects across populations and h^2 is the narrow-sense heritability of the phenotypic characters being studied. In this index, the ratio c/h^2 measures the proportion of phenotypic differences observed between populations

that can be attributed to additive genetic variance (Brommer, 2011; Leinonen et al., 2006). In general, an effect of selection can be assumed whenever (a) Pst ≠ Fst across a wide range of values (e.g. consistently after a small value of c/h^2); (b) Pst and Fst are significantly different (non-overlapping 95% confidence intervals) and (c) the confidence interval of Pst does not include the value of Fst in the neutral scenario $c = h^2$ (i.e. when the genetic architecture of the trait is the same between populations; Brommer, 2011). In the absence of reliable point estimates of c/h^2, we calculated the Pst values of each phenotypic trait for the interval ($0 \le c/h^2 \le 2$) in increments of 0.1 in the package 'Pstat' (Blondeau da Silva & Da Silva, 2018).

4 | RESULTS

4.1 | Genetic analyses

The Analyses of Molecular Variance (AMOVA) of neutral (SSR) markers showed significant differences between cultivation types and across populations within each type. Also, the amount of genetic variance explained by cultivation type differed significantly across locations as indicated by the interaction term (Table S2). Differences between types accounted for a much smaller proportion of the variance than differences among populations within each type ($R^2 = 0.02$ vs. 0.29). Local variation in the influence of type (i.e. the type × population interaction term) had also a relatively small explanatory power ($R^2 = 0.08$). These results seem to indicate that the genetic variation depends more on stochastic variability among populations than on the degree of cultivation. The component that accounted for the most genetic variation were the differences among individuals in each population ($R^2 = 0.60$) which is congruent with high intrapopulational genetic diversity (Table S2). Genetic distances among populations also corroborate a limited amount of differentiation between cultivated and feral populations. Of the six pairs analysed, four were recovered as sister OTUs (Figure 1).

In contrast to the pattern observed in neutral markers, the functional marker associated with toxicity (SNPs of the *Sk* allele) clearly diverged with feralization. The frequency of different alleles varied significantly between types ($X^2[2, 95] = 10.671$, $p < 0.01$). These differences were due to the increased frequency of homozygous dominants *Sk/Sk* (associated with the non-toxic 'sweet' phenotype) in cultigens and an increased frequency of heterozygous and homozygous recessive individuals in feral populations (Figure 2). These differences were correlated with differences in cyanide content, which were significantly different between homozygotes (*sk/sk* vs. *Sk/Sk*) and between heterozygotes and dominant homozygotes (i.e. *Sk/sk* and *Sk/Sk*, $p < 0.0001$ in both cases), but not between heterozygotes and recessive homozygotes ($p = 0.063$, Figure 2). In other words, the changes in allele frequencies between types are concomitant with the phenotypic differentiation in toxicity.
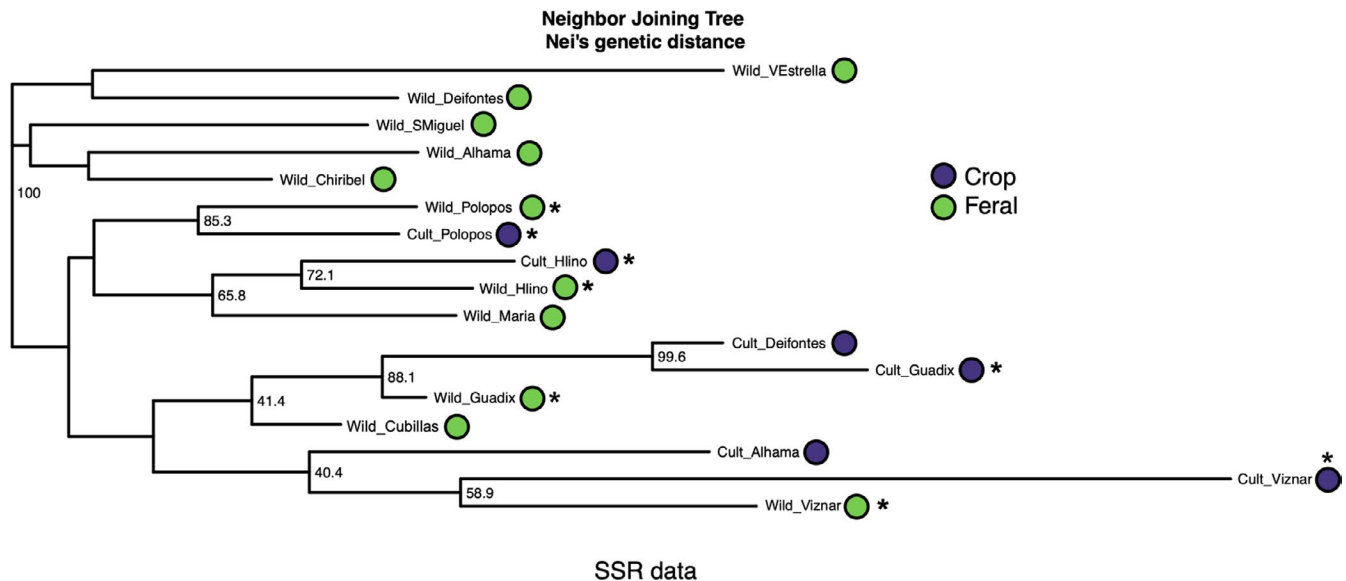
Neighbor Joining Tree
Nei's genetic distance

SSR data

FIGURE 1 Phylogenetic neighbour joining tree for the sequenced populations using Nei's genetic distance. Asterisks denote the instances in which feral and orchard populations from the same locality are paired as sister OTUS (four out of six cases)

FIGURE 2 Relationship between *Sk* genotype, almond toxicity and cultivation degree (i.e. type). (a) Mean cyanide content ± 95% CI; $N = 96$; *Sk/sk* versus *Sk/sk* or *sk/sk* significant at $p < 0.001$; *Sk/sk* vs. *sk/sk* n.s. (b) Absolute frequency of each genotype in the studied sample
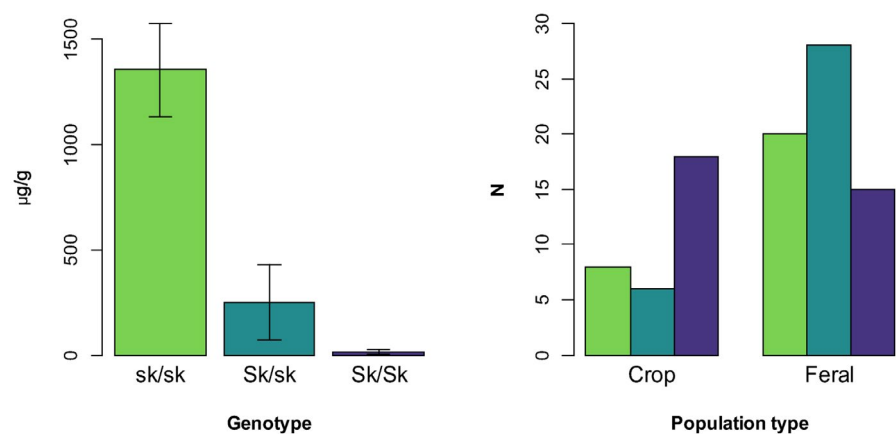
4.2 | Phenotypic differences between cultivated and feral almond populations

All feral populations were relatively young. Only in two cases (Don Fadrique and Vereda de la Estrella), we were unable to ascertain that the population was not established after 1956. Most of them (17 out of 24 feral populations) appeared to have been established in the last 40 years (after 1980) and eight were clearly established within the last 25 years (i.e. after 1998; Table 1). This short time span does not seem to have hindered phenotypic differentiation, and we found significant differences between cultivated and feral populations in total (in-shell) weight and volume, endocarp resistance and density, and in all sets of comparisons. Seed weight differed significantly between types when considering all populations and in the case of the six populations used in the genetic characterizations. However, it did not differ significantly between types when we limited the analyses to the four populations and orchards that were identified as sister groups in the phylogenetic tree (Table S3; Figure 3). Although in all cases there was also a significant effect of the population and, therefore, differences within each type should not be

disregarded, consistent trends were apparent. Feral almonds had lower weight and volume both in-shell and as naked seeds and endocarps with higher density and resistance. Toxicity differences between cultivation types were also highly significant. The average cyanide content was very significantly higher in feral almonds (Figure 3; Table S3).

4.3 | Inference of natural selection in feral almonds

Pst versus Fst comparisons indicated that the phenotypic divergence observed is consistent with the effects of directional selection. The average Fst obtained for neutral markers was 0.128 (95% CI 0.114 – 0.140). This value was clearly lower than Pst for any $c/h^2 > 0.2$ in the three traits analysed (endocarp resistance, seed weight and toxicity). Moreover, in two of the traits (endocarp resistance and seed weight) Pst = Fst only for $c/h^2 \approx 0$ (Figure 4). In other words, the additive genetic effects across populations should be <20% of the additive genetic effects within populations for Pst ≤ Fst. Furthermore, the Pst confidence intervals for the neutral hypothesis $c = h^2$ obtained
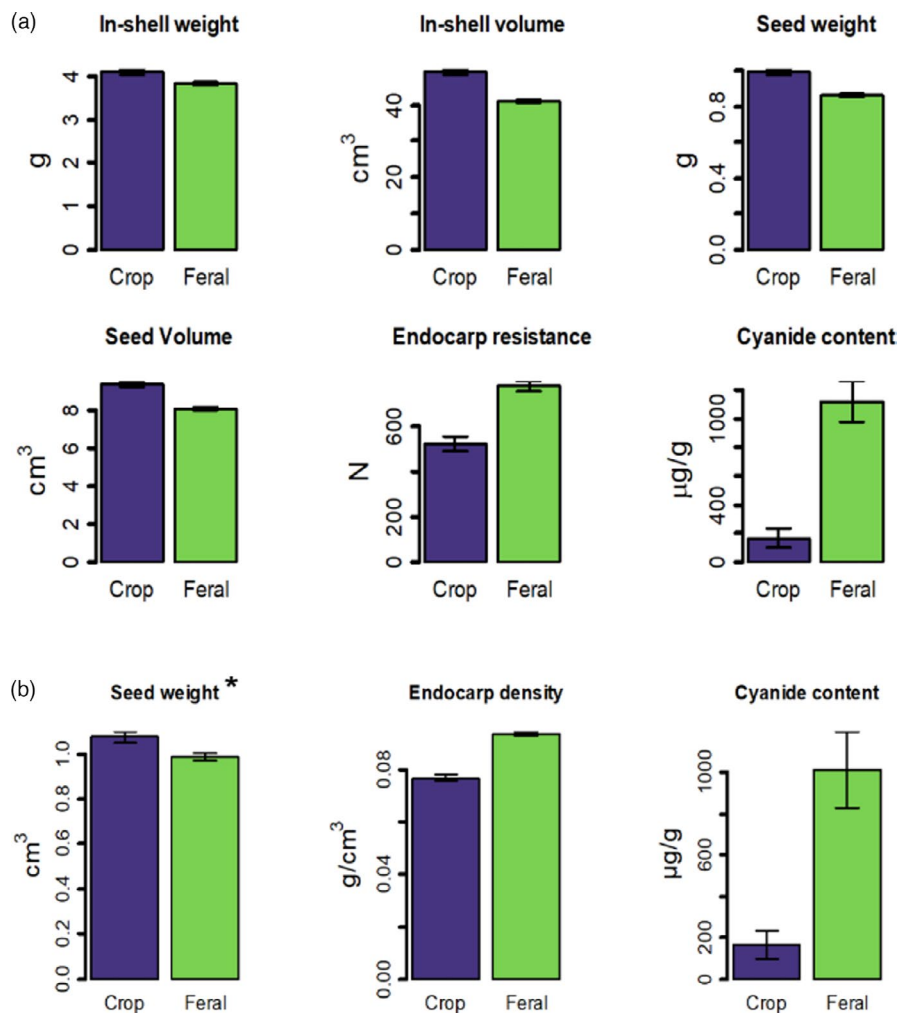
(a)

(b)

FIGURE 3 Mean trait values ±95% CI for each cultivation type. (a) Mean values of all feral and cultivated almonds available showing the overall differences between types across all populations. $N = 2707$, except for shell (endocarp) resistance and toxicity ($N = 348$ and 134, respectively). (b) Mean values for seed size, endocarp density ($N = 861$) and seed toxicity ($N = 102$) for the six feral populations and their correspondent neighbouring orchards analysed in the phylogenetic tree. These graphs are congruent with those obtained with only the four populations identified as sister groups in Figure 1, except for seed weight * which did not show significant differences between types in the subsample. In every other case, differences between crop and feral almonds were significant at $p < 0.001$

by bootstrapping were clearly higher than Fst in all three traits, especially in the case of seed weight ($Pst_{Endocarp\ Resistance} = 0.668$ [95% CI $= 0.587–0.793$]; $Pst_{Seed\ Weight} = 0.948$ [95% CI $= 0.943–0.956$]; $Pst_{Cyanide\ Content} = 0.463$ [95% CI $= 0.403–0.727$]; Figure 4).

5 | DISCUSSION

According to our genetic analyses, neutral divergence between cultivated and feral populations is relatively limited. Conversely, feralization involves significant functional differentiation. Results demonstrate clear phenotypic divergence in several fruit and seed traits. These phenotypic shifts may have taken place over the span of few (likely <4) generations. The contrast between genomic and phenotypic divergence supports the presence of strong directional selection acting on certain fruit and/or seed traits during almond naturalization.

5.1 | Genetic diversity and structure of almond populations in SE Spain

The limited genetic differentiation in neutral markers (SSR) between cultivated and feral populations stands out against the marked differentiation in the frequency of the (functional) *Sk* allele observed between types. Neutral genetic diversity was high at both inter- and intra-populational levels, which is consistent with the reproductive characteristics of the almond (self-incompatible and insect pollinated; Hamrick & Godt, 1996; Tamura et al., 2000). We found that these differences between and within populations are highly significant and account for the majority of observed genetic variance. On the other hand, differences between domesticated and feral types, albeit significant, were much lower. Conversely, the allelic frequencies of the dominant and recessive alleles of the seed toxicity marker (*Sk*) differed significantly between types. Orchards had a much lower proportion of recessive homozygotes (the putatively cyanogenic genotypes; Heppner, 1923, 1926; Sánchez-Pérez et al., 2008) and heterozygotes than their feral counterparts. The contrast between neutral and functional differentiation seems to indicate that naturalization involves a process of strong selection on at least certain seed traits (Messer et al., 2016).

5.2 | Origin of feral almond populations

Most feral almond populations were reckoned to be relatively young. According to photogrammetric data, the majority of
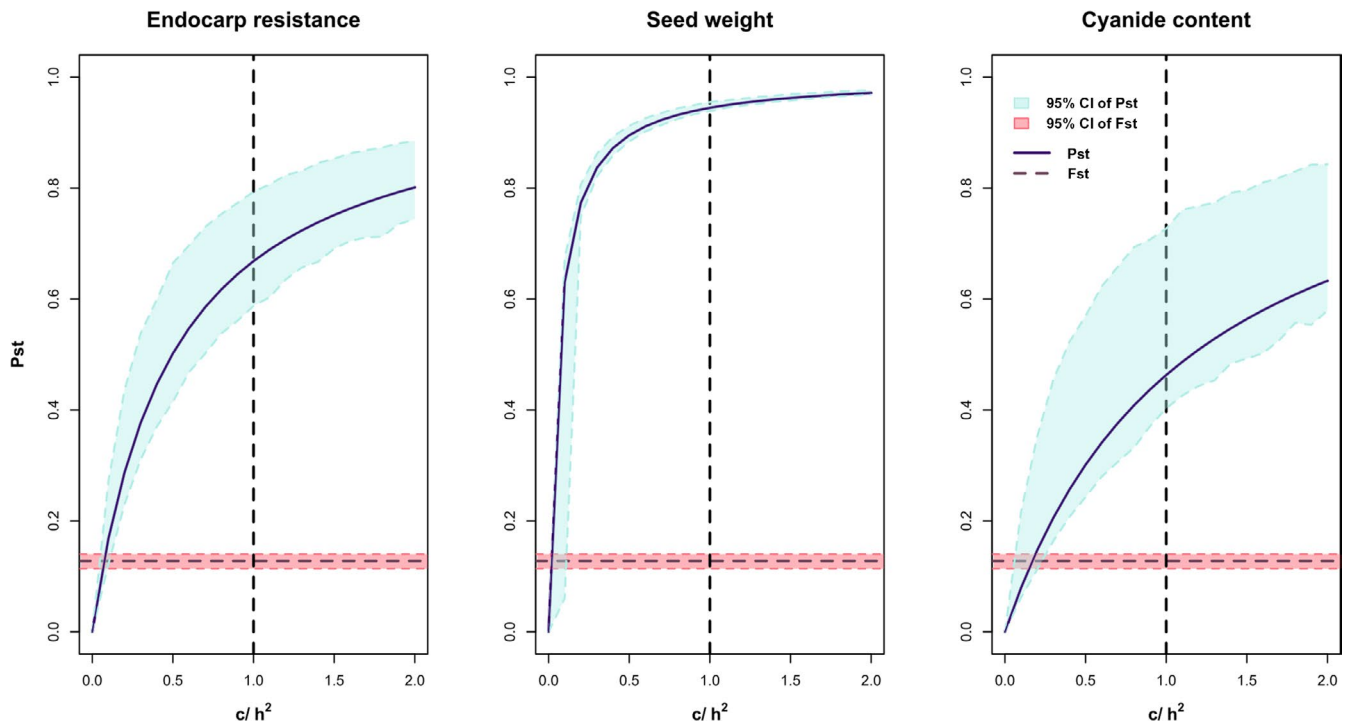
FIGURE 4 Pst versus Fst values for three highly heritable phenotypic traits. Curves represent Pst values for c/h^2 in the interval $0 \leq c/h^2 \leq 2$ at 0.1 steps. The vertical line represents the null hypothesis of equal genetic variance within and among populations, that is, $c = h^2$. Mean Fst = 0.128 is represented by the solid horizontal line. Dashed lines represent bootstrap 95% CI; (a) Endocarp resistance, $c = h^2$ 95% CI = 0.590–0.792, Fst = Pst at $c/h^2 = 0.098$; (b) Seed weight, $c = h^2$ 95% CI = 0.939–0.955, Fst = Pst at $c/h^2 = 0.022$; (c) Seed cyanogenesis, $c = h^2$ 95% CI = 0.405–0.995, Fst = Pst at $c/h^2 = 0.203$

populations (17 out of 24) did not exist 40 years ago. Consequently, divergence must have occurred over very few generations, probably <4 (assuming an average of 10 years from seed to seed, rather fast for trees; Petit & Hampe, 2006). It is unclear why feral (wild) almond populations have not been scientifically described until recently, in spite of the fact that feral trees are relatively frequent in SE Iberia and that natural populations could be a potentially important genetic resource for a valuable crop (Gering et al., 2019). A possible explanation, supported by the young age of our studied populations, is that these groves were not so common historically and represent a relatively novel feature. A series of coincidental conditions occur presently that may be facilitating the feralization of almonds. The surface covered by almond orchards has been spreading over the last decades in Spain (from ~564,000 ha in 1980 to >657,000 ha in 2018; MAPAMA, 2019). However, this increase in the total area occupied by *P. dulcis* comprises two concurrent but clearly different phenomena. New plantations tend to be located in highly productive, irrigated areas, while mountain and marginal orchards (almond has traditionally been cultivated on marginal soils of hillslopes; van Wesemael et al., 2006) have been abandoned at a steep pace, but are still standing in many cases (MAPAMA, 2019). The combination of increased gene and propagule pools and areas available to colonization might have created the perfect conditions for the spread of almonds into natural communities.

5.3 | Phenotypic divergence in the naturalization of the almond tree: morphological and biochemical traits

In spite of the lack of genetic differentiation between cultivated and feral *P. dulcis* populations and the recent origin of the later, our results demonstrated significant phenotypic differences between the two types. Some of these differences might be plastic, driven by ecological conditions and crop management practices like pruning and fertilization. Also, variation among populations of each type was significant, which might indicate that phenotypic patterns ultimately depend on local conditions, not only on whether almonds come from orchards or natural populations. Nevertheless, we observed significant and consistent divergence in traits that are highly heritable and with a very narrow genetic basis. Variation in kernel density/resistance and seed toxicity and weight are likely to be higher among than within genotypes (Pavan et al., 2021; Sánchez-Pérez et al., 2007, 2019).

Selection under cultivation has led to a wide range of almond endocarp resistance (hardness), from paper-thin to stone-hard varieties (Fornés-Comas et al., 2019). Our results are consistent with this variability; we measured cultivated almonds with cracking loads ranging from approximately 70 N to 1700 N, comparable to those observed in feral almonds. However, average resistance in the latter was significantly higher. We also observed significant differences

in the volume and mass of almonds in-shell. Almonds produced by feral trees had a lower volume although this did not correlate with a smaller in-shell weight, which was in fact slightly higher for feral almonds. Paradoxically, these same almonds tended to have lighter seeds than their cultivated counterparts. These differences were likely caused by a higher proportion of endocarp (i.e. shell). In other words, feral almond trees produced smaller kernels and seeds, but denser, more resistant endocarps. It is possible that these changes are a consequence of predator-induced selection, and that more compact, harder shells provide defence against post-dispersal predation (Fornés-Comas et al., 2019; Vander Wall & Beck, 2012; Vander Wall et al., 2019).

Predators/dispersers might also be fostering a reduction of seed size in feral almonds. However, the adaptive value of this trait and, therefore, the extent of the selective pressures acting on it may be highly variable depending on local conditions. Since they can imply a larger reward for the animal, larger, heavier seeds might be favoured and more likely to be dispersed and cached. However, survival after dispersal is not guaranteed (Gómez et al., 2019; Schupp et al., 2019) as cached seeds with larger nutritional might be consumed first (Gómez, 2004; Kuprewicz & García-Robledo, 2019; Perea et al., 2016). Moreover, there are numerous examples of smaller seeds or intermediate sized seeds being preferentially dispersed and cached (Gómez et al., 2019; Schupp et al., 2019). In addition, survival in caches can be greater for smaller seeds, for larger seeds, or be unaffected by seed size (Schupp et al., 2019). Furthermore, it is important to consider that the evolution of seed size is ultimately the net result of interactions between various selective factors, the relative importance of which can vary depending on local conditions (Kitajima & Fenner, 2000; Lázaro & Traveset, 2009; Schupp et al., 2019). In the case of the almond, one element that is likely to influence local outcomes is the genetic makeup of the local (cultivated) almond population, which might lead to differences in seed size among feral stands. According to our results, population differences within types are highly significant in every case, while differences among types were only apparent when considering a large sample. Our interpretation is that, even though selection may favour a general reduction in seed size under natural conditions, the enormous genetic variation for seed size among almond cultigens coupled with heterogeneity in local selective dynamics can result in complex patterns.

Cyanide content of feral almonds was much higher than that of the kernels collected in orchards and correlated with the increase in the frequency of the recessive allele *sk*. However, cyanide content did not behave as a simple Mendelian character and exhibited continuous values, ranging from dominant homozygotes that contain undetectable amounts to *sk/sk* almonds that have up to 5 mg/g. Heterozygotes (*Sk/sk*) in particular seem to produce highly variable amounts of cyanide. Although cultivated almonds are all sweet (Ladizinsky, 1999), many of them are heterozygotes. Moreover, recessive homozygotes can be found in cultivated orchards, used as 'wildtype' rootstocks, which occasionally flower and bear fruit, particularly in old and abandoned (or semi-abandoned) orchards (Rubio-Cabetas, 2016) which can foster the appearance of cyanogenic

genotypes in naturalized populations. These might be recessive homozygotes *sk/sk* but also, according to our results, occasionally heterozygotes.

The increase in toxicity associated with feralization might be adaptive at different stages. First, toxicity could favour dispersal and decrease post-dispersal predation, either of the seeds or seedlings (Beckman et al., 2019; Sánchez-Pérez et al., 2008). Additionally, bitter genotypes might use amygdalin as a source of nitrogen during the initial phases of growth (Gleadow & Woodrow, 2002; Sánchez-Pérez et al., 2008). With the available data, we cannot determine whether the preponderance of cyanogenic individuals in naturalized populations is due to processes that take place during the seed stage or later on.

5.4 | Feralization and phenotypic evolution in almond

Our comparisons of neutral (Fst) and quantitative (Pst) genetic differentiation support the idea that directional selection acts on at least three traits during feralization: endocarp resistance, and seed weight and toxicity. The estimated values of genetic differentiation (Pst) for these three traits were notably higher than the average value observed in neutral markers (Fst), independently of interpopulational genetic variation and heritability (as estimated by varying values of c/h^2). Although we cannot precisely quantify the variance caused by environmental differences between individuals and populations, these three traits are highly heritable (i.e. they have very high h^2). Two of them—endocarp density and toxicity—are considered monogenic (Dicenta et al., 1993; Sánchez-Pérez et al., 2007; Spiegel-Roy & Kochba, 1981) while recent results by Pavan et al. (2021) indicate that seed weight also has a very narrow genetic basis. However, given the wide genetic diversity across populations and the high genetic and phenotypic diversity of cultivated almonds (Fornés-Comas et al., 2019; Halász et al., 2019; Socias i Company & Felipe, 1992), genetic variance for all traits (i.e. the parameter c) is also expected to be significant. We detected a large proportion of heterozygotes for cyanide content in cultivated populations, and as mentioned above, kernel density/resistance and seed weight varied significantly across populations independent of type. Therefore, it seems reasonable to assume that the actual ratio of intra- and inter-population genetic variances c/h^2 is relatively high (i.e. ~1) or at least clearly different from zero. Under these circumstances, the most parsimonious explanation for the differences we observed in these traits are processes of directional selection (Brommer, 2011). However, we believe that further investigations are necessary to prove the existence of these selective dynamics. More exhaustive genetic analyses, including genomic data from multiple individuals and populations, will be needed to corroborate whether feralization entails selective sweeps on specific genetic regions.

In light of the results described here, it is also unavoidable to wonder what role dispersers are playing in this process. Synzoochory

strikes a delicate balance between mutualism and antagonism, since the animals that act as seed dispersers are, simultaneously, seed predators (Gómez et al., 2019). It is possible that the traits under selection in feral populations involve a compromise between ensuring seed dispersal and survival (Fricke & Wright, 2016; Patton et al., 1997; Vander Wall & Beck, 2012). However, at this stage, we lack evidence of the mechanisms driving this selective process.

Further research is also necessary to clarify the potential role of pre-existing isolated feral trees in the emergence of the novel, phenotypically different populations. Our sampling prioritized naturalized populations in abandoned areas near cultivated almond orchards, which were predicted to be the most probable source of pollen and seeds. Our genetic results showed that although this was indeed a plausible scenario in most cases, it did not fit the results for (at least) two of the pairs of orchard-feral populations considered. This indicates that the feral genomes might have a strong imprint from other sources. It is possible that selection during feralization may have been facilitated by admixture with pre-existing semi-wild individuals. These have been traditionally tolerated and even occasionally planted along field edges for different purposes (Ibancos Núñez & Rodríguez Franco, 2010; Ruiz de la Torre, 2006). Future studies could elucidate the relative contribution of these marginal trees to the evolution during naturalization of almonds.

6 | CONCLUSIONS

The results of this study reveal a pattern of genetic and phenotypic divergence in feral almond populations. This divergence is consistent with an evolutionary process potentially mediated by biotic dispersal agents. Our findings appear to indicate that during naturalization in the Iberian Peninsula, the fruits and seeds of *Prunus dulcis* are under strong directional selection that favours smaller and more toxic seeds encased in harder endocarps.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

AUTHORS' CONTRIBUTIONS
R.R.d.C., J.M.G. and E.W.S. conceived the study and designed the preliminary experiments; R.R.d.C., F.J.O.-C., R.B.-R. and A.B.M. finalized the experimental design; R.R.d.C., F.J.O.-C., R.B.-R., A.B.M., R.S.-P., J.G. and J.Z. collected the data; A.B.M. and R.R.d.C. analysed the data; A.B.M., R.R.d.C. and J.Z. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ORCID
Rafael Rubio de Casas https://orcid.org/0000-0003-4276-4968

REFERENCES

Allaby, R. G. (2014). Domestication syndrome in plants. In C. Smith (Ed.), Encyclopedia of global archaeology (pp. 2182–2184). Springer New York. https://doi.org/10.1007/978-1-4419-0465-2_2416

Arrazola, G., Grane, N., Martin, M. L., & Dicenta, F. (2013). Determinación de compuestos cianogénicos amigdalina y prunasina en semillas de almendras (*Prunus dulcis* L.) utilizando cromatografía liquida de alta resolución. *Revista Colombiana de Química*, *42*(3), 23–30.

Balaguer-Romano, R., Barea-Márquez, A., Ocaña-Calahorro, F., Gómez, J. M., Schupp, E. W., Zhang, J., & Rubio de Casas, R. (2021). The potential role of synzoochory in the naturalization of almond *Prunus dulcis* (Mill) D.A. Webb. *Basic and Applied Ecology*, *50*, 97–106. https://doi.org/10.1016/j.baae.2020.11.004

Beckman, N. G., Aslan, C. E., Rogers, H. S., Kogan, O., Bronstein, J. L., Bullock, J. M., Hartig, F., Zhou, Y., Zurell, D., Brodie, J. F., Bruna, E. M., Cantrell, R. S., Decker, R. R., Efiom, E., Fricke, E. C., Gurski, K., Hastings, A., Johnson, J. S., Loiselle, B. A., ... Zambrano, J. (2019). Advancing an interdisciplinary framework to study seed dispersal ecology. *AoB Plants*, *12*(plz048). https://doi.org/10.1093/aobpla/plz048

Blackburn, T. M., Pyšek, P., Bacher, S., Carlton, J. T., Duncan, R. P., Jarošík, V., Wilson, J. R. U., & Richardson, D. M. (2011). A proposed unified framework for biological invasions. Trends in Ecology & Evolution. *Trends in Ecology & Evolution*, *26*(7), 333–339. https://doi.org/10.1016/j.tree.2011.03.023

Blondeau da Silva, S., & Da Silva, A. (2018). Pstat: An R package to assess population differentiation in phenotypic traits. *The R Journal*, *10*(1), 447–454. https://doi.org/10.32614/RJ-2018-010

Brommer, J. E. (2011). Whither Pst? The approximation of Qst by Pst in evolutionary and conservation biology. *Journal of Evolutionary Biology*, *24*, 1160–1168. https://doi.org/10.1111/j.1420-9101.2011.02268.x

Browicz, K., & Zohary, D. (1996). The genus *Amygdalus* L. (Rosaceae): Species relationships, distribution and evolution under domestication. *Genetic Resources and Crop Evolution*, *43*, 229–247.

Cornille, A., Gladieux, P., Smulders, M. J. M., Roldán-Ruiz, I., Laurens, F., Le Cam, B., Nersesyan, A., Clavel, J., Olonova, M., Feugey, L., Gabrielyan, I., Zhang, X.-G., Tenaillon, M. I., & Giraud, T. (2012). New insight into the history of domesticated apple: Secondary contribution of the European wild apple to the genome of cultivated varieties. *PLoS Genetics*, *8*(5), e1002703. https://doi.org/10.1371/journal.pgen.1002703

Delignette-Muller, M.-L., & Dutang, C. (2015). Fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software*, *64*(4), 1–34.

Delplancke, M., Alvarez, N., Benoit, L., Espíndola, A., Joly, H. I., Neuenschwander, S., & Arrigo, N. (2013). Evolutionary history of almond tree domestication in the Mediterranean basin. *Molecular Ecology*, *22*(4), 1092–1104. https://doi.org/10.1111/mec.12129

Dicenta, F., García, J. E., & Carbonell, E. (1993). Heritability of fruit characters in almond. *Journal of Horticultural Science & Biotechnology*, *68*, 121–126. https://doi.org/10.1080/00221589.1993.11516335

Dickmann, D. I., Gold, M. A., & Flore, J. A. (1994). The Ideotype concept and the genetic improvement of tree crops. In J. Janick (Ed.), *Plant breeding reviews* (Vol. *12*, pp. 163–193). Wiley.

Ellstrand, N. C. (2018). Born to run? Not necessarily: Species and trait bias in persistent free-living transgenic plants. *Frontiers in Bioengineering and Biotechnology*, *6*(88). https://doi.org/10.3389/fbioe.2018.00088

Ellstrand, N. C., Heredia, S. M., Leak-Garcia, J. A., Heraty, J. M., Burger, J. C., Yao, L. I., Nohzadeh-Malakshah, S., & Ridley, C. E. (2010). Crops gone wild: Evolution of weeds and invasives from domesticated ancestors. *Evolutionary Applications*, *3*(5–6), 494–504. https://doi.org/10.1111/j.1752-4571.2010.00140.x

Fornés Comas, J., Socias i Company, R., & Alonso Segura, J. M. (2019). Shell hardness in almond: Cracking load and kernel percentage. *Scientia Horticulturae*, *245*, 7–11. https://doi.org/10.1016/j.scienta.2018.09.075

Freeland, W. J., & Janzen, D. H. (1974). Strategies in herbivory by mammals: The role of plant secondary compounds. *The American Naturalist*, *108*(961), 269–289. https://doi.org/10.1086/282907

Fricke, E. C., & Wright, S. J. (2016). The mechanical defence advantage of small seeds. *Ecology Letters*, *19*(8), 987–991. https://doi.org/10.1111/ele.12637

García, D., Obeso, J. R., & Martínez, I. (2005). Rodent seed predation promotes differential recruitment among bird-dispersed trees in temperate secondary forests. *Oecologia*, *144*(3), 435–446. https://doi.org/10.1007/s00442-005-0103-7

Gering, E., Incorvaia, D., Henriksen, R., Conner, J. K., Getty, T., & Wright, D. (2019). Getting back to nature: Feralization in animals and plants. *Trends in Ecology & Evolution*, *34*(12), 1137–1151. https://doi.org/10.1016/j.tree.2019.07.018

Gleadow, R. M., & Woodrow, I. E. (2002). Constraints on effectiveness of cyanogenic glycosides in herbivore defense. *Journal of Chemical Ecology*, *28*, 1301–1313.

Gómez, J. M. (2004). Bigger is not always better: Conflicting selective pressures on seed size in *Quercus ilex*. *Evolution*, *58*(1), 71–80. https://doi.org/10.1554/02-617

Gómez, J. M., Schupp, E. W., & Jordano, P. (2019). Synzoochory: The ecological and evolutionary relevance of a dual interaction. *Biological Reviews*, *94*, 874–902. https://doi.org/10.1111/brv.12481

Goudet, J. (2005). HIERSTAT, a package for R to compute and test hierarchical *F*-statistics. *Molecular Ecology Resources*, *5*(1), 184–186.

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

Halász, J., Kodad, O., Galiba, G. M., Skola, I., Ercisli, S., Ledbetter, C. A., & Hegedűs, A. (2019). Genetic variability is preserved among strongly differentiated and geographically diverse almond germplasm: An assessment by simple sequence repeat markers. *Tree Genetics & Genomes*, *15*(1), 12. https://doi.org/10.1007/s11295-019-1319-8

Hamrick, J. L., & Godt, J. W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London, Series B*, *351*, 1291–1298.

Heppner, M. J. (1923). The factor for bitterness in the sweet almond. *Genetics*, *8*(4), 390. https://doi.org/10.1093/genetics/8.4.390

Heppner, M. J. (1926). Further evidence on the factor for bitterness in the sweet almond. *Genetics*, *11*(6), 605. https://doi.org/10.1093/genetics/11.6.605

Homet-Gutiérrez, P., Schupp, E. W., & Gómez, J. M. (2015). Naturalization of almond trees (*Prunus dulcis*) in semi-arid regions of the Western Mediterranean. *Journal of Arid Environments*, *113*, 108–113. https://doi.org/10.1016/j.jaridenv.2014.10.005

Ibancos Núñez, C., & Rodríguez Franco, R. (2010). *Biodiversidad y conocimiento local Las variedades cultivadas autóctonas en el Entorno de Doñana*. Consejería de Agricultura y Pesca, Junta de Andalucía.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal and/or sexual reproduction. *PeerJ*, *2*, e281.

Kitajima, K., & Fenner, M. (2000). Ecology of seedling regeneration. In M. Fenner (Ed.), *Seeds: Ecology of regeneration in plant communities* (2nd ed., pp. 331–360). CABI Publishing.

Kuprewicz, E. K., & García-Robledo, C. (2019). Deciphering seed dispersal decisions: Size, not tannin content, drives seed fate and survival in a tropical forest. *Ecosphere*, *10*(1), e02551. https://doi.org/10.1002/ecs2.2551

Ladizinsky, G. (1999). On the origin of almond. *Genetic Resources and Crop Evolution*, *46*(2), 143–147.

Larson, G., Piperno, D. R., Allaby, R. G., Purugganan, M. D., Andersson, L., Arroyo-Kalin, M., Barton, L., Climer Vigueira, C., Denham, T., Dobney, K., Doust, A. N., Gepts, P., Gilbert, M. T. P., Gremillion, K. J., Lucas, L., Lukens, L., Marshall, F. B., Olsen, K. M., Pires, J. C., ... Fuller, D. Q. (2014). Current perspectives and the future of domestication studies. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(17), 6139–6146. https://doi.org/10.1073/pnas.1323964111

Lázaro, A., & Traveset, A. (2009). Does the spatial variation in selective pressures explain among-site differences in seed mass? A test with *Buxus balearica*. *Evolutionary Ecology*, *23*(6), 847. https://doi.org/10.1007/s10682-008-9275-z

Leinonen, T., Cano, J. M., Makinen, H., & Merilä, J. (2006). Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *Journal of Evolutionary Biology*, *19*, 1803–1812. https://doi.org/10.1111/j.1420-9101.2006.01182.x

Leinonen, T., McCairns, R. J., O'Hara, R. B., & Merilä, J. (2013). Qst–Fst comparisons: Evolutionary and ecological insights from genomic heterogeneity. *Nature Review Genetics*, *14*, 179–190. https://doi.org/10.1038/nrg3395

Leinonen, T., O'Hara, R. B., Cano, J. M., & Merilä, J. (2008). Comparative studies of quantitative trait and neutral marker divergence: A meta-analysis. *Journal of Evolutionary Biology*, *21*, 1–17. https://doi.org/10.1111/j.1420-9101.2007.01445.x

MAPAMA. (2019). *Anuario de Estadística Agrícola* (No. 2018). Ministry of Agriculture and Environment of Spain.

Merilä, J., & Crnokrak, P. (2001). Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, *14*, 892–903. https://doi.org/10.1046/j.1420-9101.2001.00348.x

Messer, P., Ellner, S., & Hairston, N. G. (2016). Can population genetics adapt to rapid evolution? *Trends in Genetics*, *32*, 408–418. https://doi.org/10.1016/j.tig.2016.04.005

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

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

Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, *106*, 283–292. https://doi.org/10.1086/282771

Oksanen, J., Blanchet, G., Fiendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, H. M. H., Szoecs, E., & Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-6. https://CRAN.Rproject.org/package=vegan

Oshima, H., Ueno, E., Saito, I., & Matsumoto, H. (2003). Quantitative determination of cyanide in foods by spectrophotometry using picric acid test strips. *Japanese Journal of Food Chemistry and Safety*, *10*(2), 96–100. https://doi.org/10.18891/jjfcs.10.2_96

Pannell, J. R., Auld, J. R., Brandvain, Y., Burd, M., Busch, J. W., Cheptou, P.-O., Conner, J. K., Goldberg, E. E., Grant, A.-G., Grossenbacher, D. L., Hovick, S. M., Igic, B., Kalisz, S., Petanidou, T., Randle, A. M., Casas, R. R., Pauw, A., Vamosi, J. C., & Winn, A. A. (2015). The scope of Baker's law. *The New Phytologist*, *208*(3), 656–667. https://doi.org/10.1111/nph.13539

Patton, C. A., Ranney, T. G., Burton, J. D., & Walgenbach, J. F. (1997). Natural pest resistance of *Prunus taxa* to feeding by adult Japanese beetles: Role of endogenous allelochemicals in host plant resistance. *Journal of the American Society for Horticultural Science*, *122*(5), 668–672. https://doi.org/10.21273/JASHS.122.5.668

Pavan, S., Delvento, C., Mazzeo, R., Ricciardi, F., Losciale, P., Gaeta, L., D'Agostino, N., Taranto, F., Sánchez-Pérez, R., Ricciardi, L., & Lotti, C. (2021). Almond diversity and homozygosity define structure, kinship, inbreeding, and linkage disequilibrium in cultivated germplasm, and reveal genomic associations with nut and seed weight. *Horticulture Research*, *8*(1), 1–12. https://doi.org/10.1038/s41438-020-00447-1

Perea, R., Dirzo, R., San Miguel, A., & Gil, L. (2016). Post-dispersal seed recovery by animals: Is it a plant or an animal-driven process? *Oikos*, *125*(8), 1203–1210. https://doi.org/10.1111/oik.02556

Petit, R. J., & Hampe, A. (2006). Some evolutionary consequences of being a tree. *Annual Review of Ecology Evolution and Systematics*, *37*, 187–214. https://doi.org/10.1146/annurev.ecolsys.37.091305.110215

Richardson, D. M., Pysek, P., Rejmanek, M., Barbour, M. G., Panetta, F. D., & West, C. J. (2000). Naturalization and invasion of alien plants: Concepts and definitions. *Diversity and Distributions*, *6*(2), 93–107. https://doi.org/10.1046/j.1472-4642.2000.00083.x

Richardson, D. M., & Rejmánek, M. (2011). Trees and shrubs as invasive alien species – A global review. *Diversity and Distributions*, *17*(5), 788–809. https://doi.org/10.1111/j.1472-4642.2011.00782.x

Rubio de Casas, R., Barea Marquez, A., Ocaña-Calahorro, F., Balaguer-Romano, R., Gómez, J. M., Schupp, E. W., Sanchez-Perez, R., Guillamon, J., & Zhang, J. (2021). Data from: Phenotypic and Genetic data for 'Evolution of fruit and seed traits during almond naturalization'. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.mcvdnck29

Rubio-Cabetas, M. J. (2016). Almond rootstock overview. *Options Méditerranéenes A*, *119*, 133–143.

Ruiz de la Torre, J., Viñas, J. I. G., Ruiz, J., del Castillo y de Navascués, & Cardo, Ó. G. (2006). *Flora mayor*. Organismo Autónomo Parques Nacionales.

Sánchez-Pérez, R., Belmonte, F. S., Borch, J., Dicenta, F., Møller, B. L., & Jørgensen, K. (2012). Prunasin hydrolases during fruit development in sweet and bitter almonds. *Plant Physiology*, *158*(4), 1916–1932. https://doi.org/10.1104/pp.111.192021

Sánchez-Pérez, R., Howard, W. E., Dicenta, F., Arús, P., & Martínez-Gómez, P. (2007). Mapping major genes and quantitative trait loci controlling agronomic traits in almond. *Plant Breeding*, *126*, 310–318. https://doi.org/10.1111/j.1439-0523.2007.01329.x

Sánchez-Pérez, R., Jørgensen, K., Olsen, C. E., Dicenta, F., & Møller, B. L. (2008). Bitterness in almonds. *Plant Physiology*, *146*(3), 1040–1052. https://doi.org/10.1104/pp.107.112979

Sánchez-Pérez, R., Pavan, S., Mazzeo, R., Moldovan, C., Aiese Cigliano, R., Del Cueto, J., Ricciardi, F., Lotti, C., Ricciardi, L., Dicenta, F., López-Marqués, R. L., & Møller, B. L. (2019). Mutation of a bHLH transcription factor allowed almond domestication. *Science*, *364*(6445), 1095. https://doi.org/10.1126/science.aav8197

Schupp, E. W., Zwolak, R., Jones, L. R., Snell, R. S., Beckman, N. G., Aslan, C., & Shea, K. (2019). Intrinsic and extrinsic drivers of intraspeciic variation in seed dispersal are diverse and pervasive. *AoB Plants*, *11*(6), plz067.

Socias i Company, & Felipe, A. J. (1992). Almond: A diverse germplasm. *HortScience*, *27*(7), 718–863. https://doi.org/10.21273/HORTSCI.27.7.718

Spiegel-Roy, P., & Kochba, J. (1981). Inheritance of nut and kernel traits in almond. *Euphytica*, *30*, 161–174.

Struempf, H. G., Schondube, J. E., & Martínez del Río, C. (1999). The cyanogenic glycoside amygdalin does not deter consumption of ripe fruit by Cedar Waxwings. *The Auk*, *116*(3), 749–758. https://doi.org/10.2307/4089335

Tamura, M., Ushijima, K., Sassa, H., Hirano, H., Tao, R., Gradziel, T. M., & Dandekar, A. M. (2000). Identification of self-incompatibility genotypes of almond by allele-specific PCR analysis. *Theoretical and Applied Genetics*, *101*, 344–349. https://doi.org/10.1007/s001220051489

van Wesemael, B., Rambaud, X., Poesen, J., Muligan, M., Cammeraat, E., & Stevens, A. (2006). Spatial patterns of land degradation and their impacts on the water balance of rainfed treecrops: A case study in South East Spain. *Geoderma*, *133*, 43–56. https://doi.org/10.1016/j.geoderma.2006.03.036

Vander Wall, S. B. (2001). The evolutionary ecology of nut dispersal. *The Botanical Review*, *67*, 74–117. https://doi.org/10.1007/BF02857850

Vander Wall, S. B. (2010). How plants manipulate the scatter-hoarding behaviour of seed-dispersing animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*, 989–997. https://doi.org/10.1098/rstb.2009.0205

Vander Wall, S. B., & Beck, M. (2012). A comparison of frugivory and scatter-hoarding seed-dispersal syndromes. *Botanical Review*, *78*, 10–31. https://doi.org/10.1007/s12229-011-9093-9

Vander Wall, S. B., Dimitri, L. A., Longland, W. S., & White, J. D. M. (2019). Seed value influences cache pilfering rates by desert rodents. *Integrative Zoology*, *14*, 75–86. https://doi.org/10.1111/1749-4877.12358

Wu, D., Lao, S., & Fan, L. (2021). De-domestication: An extension of crop evolution. *Trends in Plant Science*, *26*(6), 560–574. https://doi.org/10.1016/j.tplants.2021.02.003

Zhang, H., Mittal, N., Leamy, L. J., Barazani, O., & Song, B.-H. (2017). Back into the wild – Apply untapped genetic diversity of wild relatives for crop improvement. *Evolutionary Applications*, *10*(1), 5–24. https://doi.org/10.1111/eva.12434

Zohary, D., & Spiegel-Roy, P. (1975). Beginnings of fruit growing in the Old World. *Science*, *187*, 319–327. https://doi.org/10.1126/science.187.4174.319

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.