The domestication and evolutionary ecology of apples

Trends in Genetics

Cornille, A.; Giraud, T.; Smulders, M.J.M.; Roldán-Ruiz, I.; Gladieux, P.

https://doi.org/10.1016/j.tig.2013.10.002

Review

The domestication and evolutionary ecology of apples

Amandine Cornille[1,2], Tatiana Giraud[1,2], Marinus J.M. Smulders[3], Isabel Roldán-Ruiz[4], and Pierre Gladieux[1,2]

[1] Centre National de la Recherche Scientifique (CNRS), Unité Mixte de Recherche (UMR) 8079, Bâtiment 360, 91405 Orsay, France
[2] Université Paris Sud, UMR 8079, Bâtiment 360, 91405 Orsay, France
[3] Wageningen UR Plant Breeding, Wageningen University & Research Centre, Wageningen, The Netherlands
[4] Growth and Development Group, Plant Sciences Unit, Institute for Agricultural and Fisheries Research (ILVO), Melle, Belgium

The cultivated apple is a major fruit crop in temperate zones. Its wild relatives, distributed across temperate Eurasia and growing in diverse habitats, represent potentially useful sources of diversity for apple breeding. We review here the most recent findings on the genetics and ecology of apple domestication and its impact on wild apples. Genetic analyses have revealed a Central Asian origin for cultivated apple, together with an unexpectedly large secondary contribution from the European crabapple. Wild apple species display strong population structures and high levels of introgression from domesticated apple, and this may threaten their genetic integrity. Recent research has revealed a major role of hybridization in the domestication of the cultivated apple and has highlighted the value of apple as an ideal model for unraveling adaptive diversification processes in perennial fruit crops. We discuss the implications of this knowledge for apple breeding and for the conservation of wild apples.

The genus *Malus* as a model system for understanding fruit tree evolution

Apple (*Malus domestica* Borkh) is one of the most important fruit crops in temperate regions worldwide in terms of production levels (http://faostat.fao.org/), and it occupies a central position in folklore, culture, and art [1]. Dessert apples are popular because of their taste, nutritional properties, storability, and convenience of use. Apple-based beverages such as ciders have been consumed for centuries by the peoples of Eurasia, even before the advent of the cultivated apple. Humans have been exploiting, selecting, and transporting apples for centuries, and several thousand apple cultivars have been documented [2]. Much of the diversity present in domesticated apples is currently maintained in germplasm (see Glossary) repositories and amateur collections, including a broad spectrum of cultivars of largely unknown history and pedigree [2,3]. Apples are self-incompatible, and it must soon have become apparent that offspring grown from seed frequently did not resemble the mother apple tree. This high degree of variability of the progeny, combined with a long juvenile phase, probably complicated and slowed the artificial selection of interesting phenotypes by early farmers. The introduction of vegetative propagation by grafting, and the selection of dwarfed apple trees for use as rootstocks, were thus key events in the history of apples, facilitating the

Corresponding author: Cornille, A. (amandine.cornille@gmail.com).
Keywords: QTL; archaeological record; crabapple; genome; microsatellites; *Malus domestica*.

0168-9525/$ – see front matter
 http://dx.doi.org/10.1016/j.tig.2013.10.002

Glossary

Crabapple: wild apple species, usually producing profuse blossom and small acidic fruits. The word crab comes from the Old English 'crabbe' meaning bitter or sharp tasting. Many crabapples are cultivated as ornamental trees and their apples are sometimes used for preserves. In Western Europe the term crabapple is often used to refer to *Malus sylvestris* (the European crabapple), in the Caucasus this term refers to *Malus orientalis* (the Caucasian crabapple) and, in Siberia, to *Malus baccata* (the Siberian crabapple). The native North American crabapples are *Malus fusca, Malus coronaria, Malus angustifolia,* and *Malus ioensis. Malus sieversii*, the main progenitor of the cultivated apple, is usually not referred to as a 'crabapple'.
Diachronic: changing over time.
Germplasm: collection of genetic resources for a species.
Hybridization: the process of interbreeding between individuals from different gene pools.
Introgression: the transfer of genomic regions from one species into the gene pool of another species through an initial hybridization event followed by repeated backcrosses.
Isolation by distance: pattern of correlation between genetic differentiation and geographic distance, resulting from a balance between genetic drift and dispersal.
Panmictic group: random mating population.
QTL mapping: identification of genomic regions governing a phenotypic trait by the statistical association of molecular markers with the phenotypic trait in a population derived from a cross between two or more parental lines (a 'segregating population') or in a population of unrelated individuals (an 'association mapping population').
Rootstocks: part of a plant (usually the underground part) onto which a graft is fixed. In apples, the genotypes of the rootstocks used often result in dwarf apple trees.
Self-incompatible: a plant incapable of self-fertilization because its own pollen is prevented from germinating on the stigma or because the pollen tube is blocked before it reaches the ovule.
Simple sequence repeats (SSR): in other words microsatellites, the repetition in tandem of di-, tri-, or tetranucleotide motifs.
SSR marker (or microsatellite marker): a marker based on PCR amplification and separation of SSR alleles with differences in repeat length.
Vegetative propagation: natural or assisted asexual reproduction in plants through the regeneration of tissues or plant organs from one plant part or tissue. Natural vegetative propagation methods include bulbs or corms, for example. Horticultural vegetative propagation may involve stem or root cuttings or grafting. Grafting is often used for the propagation and maintenance of cultivars of interest in perennial fruit crops.

selection and propagation of superior genotypes derived from open-pollination progenies (i.e., 'chance seedlings') [4], and improving control over the size of mature apple trees [1]. Despite the immense diversity of cultivars available, apple production worldwide is now largely based on the cultivation of a few dozen ornamental and edible cultivars, grafted onto less than a dozen different clonal rootstocks, with high levels of chemical inputs.

Wild *Malus* taxa constitute a useful source of variation for apple breeding. They are also of ecological importance as a source of food for wildlife or as components of hedges in agricultural landscapes and, as such, they make a significant contribution to ecosystem services. The use of these natural resources in breeding and the development of effective conservation programs require a good understanding of the genetic relationships within and between cultivated and related wild *Malus* taxa. More generally, the cultivated apple and its wild relatives constitute a suitable model system for studies of the domestication of long-lived perennial crops, about which much less is known than for seed-propagated annual crops.

Recent population studies have provided insight into the origins of the cultivated apple, its domestication history, and its spread by human populations, the status and genetic integrity of its wild relatives, and the response to selection of these long-lived perennials. Here, we review recent progress in studies of the origin and diversity of cultivated apple and its wild relatives. We also explore future prospects for increasing our understanding of the nature and genetic architecture of the phenotypic changes associated with apple domestication across the wide range of habitats and environmental conditions in which wild apples grow.

The status and genetic integrity of cultivated and wild apple taxa

The apple belongs to the Rosaceae family, as do other major temperate fruit tree species (e.g., pear, apricot, peach, plum, cherry, and almond). The genus *Malus* consists of about 30 species and several subspecies, but the taxonomy of this genus is complex, unclear, and likely to be revised in the future [5]. Species delimitation in apples is hampered by the paucity of diagnostic morphological features [6], the fact that records are lacking concerning the geographic range of some taxa, and the prevalence of hybridization between species. Hybridization among apples is facilitated by a lack of interspecific reproductive barriers, self-incompatibility, and the cultivation of apple in areas in which wild apple populations occur naturally [7–11].

The cultivated (or 'sweet') apple is usually designated as *Malus pumila* Mill. or *M. domestica* Borkh., but many other synonyms, now considered illegitimate, have been used. The name *M. pumila* seems to be more legitimate in terms of the rules of botanical nomenclature, but *M. domestica* is the name most widely used, and this name will be used here. *M. domestica* belongs to the section *Malus*, one of the five sections within the genus *Malus*, and is a diverse crop species containing a large number of wild-introgressed individuals and feral populations [12]. The cultivated apple was initially domesticated from the

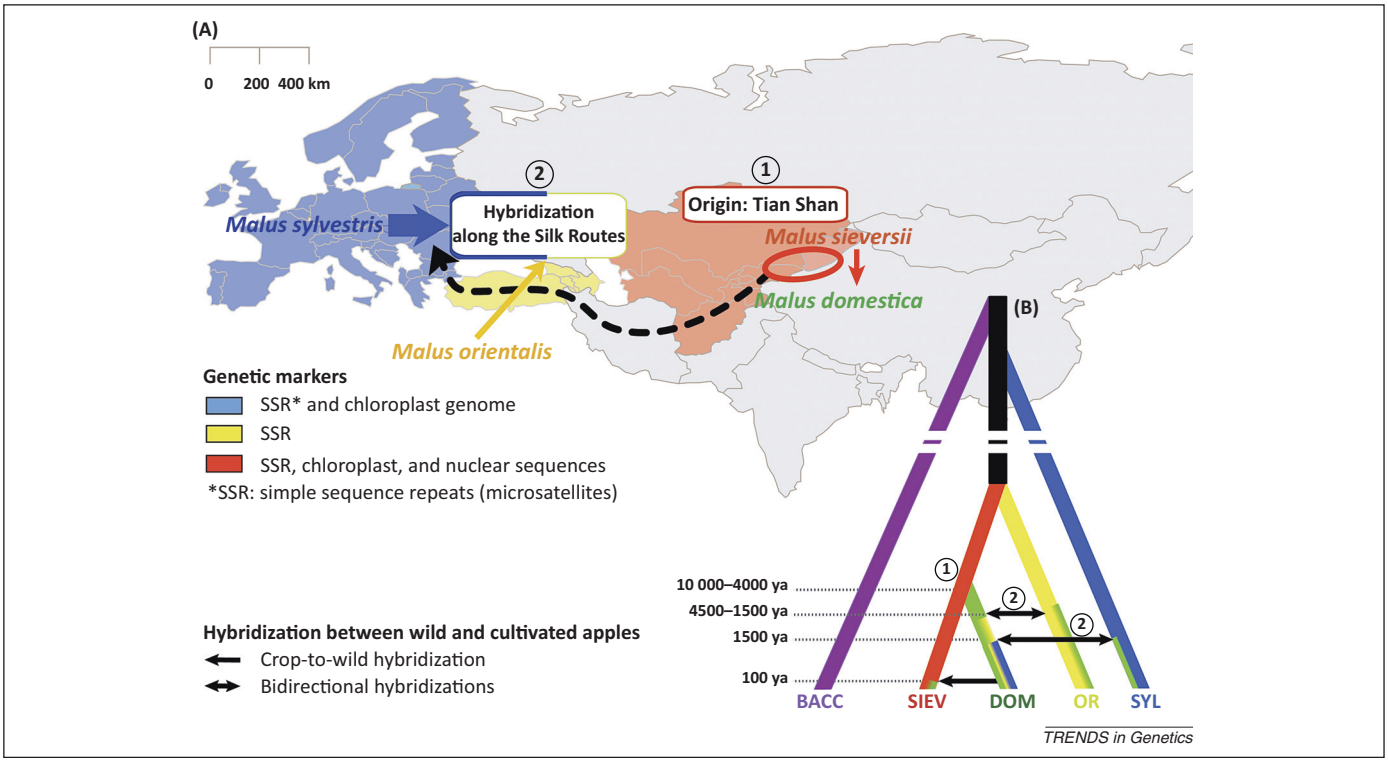

Figure 1. Evolutionary history of the cultivated apple. **(A)** This history was revealed by recent population studies using different types of molecular markers for evolutionary inferences. (1) Origin in the Tian Shan Mountains from *Malus sieversii,* followed by (2) dispersal from Asia to Europe along the Silk Route, facilitating hybridization and introgression from the Caucasian and European crabapples. Arrow thickness is proportional to the genetic contribution of various wild species to the genetic makeup of *Malus domestica*. **(B)** Genealogical relationships between wild and cultivated apples. Approximate dates of the domestication and hybridization events between wild and cultivated species are detailed in the legend. Abbreviations: BACC, *Malus baccata*; DOM, *M. domestica*; OR, *Malus orientalis*; SIEV, *M. sieversii*; SYL, *Malus sylvestris*; ya, years ago.

wild apple *Malus sieversii* (Ldb.) Roem in the Tian Shan Mountains in Central Asia [11,13–15]. People then took the domesticated apples westwards along the great trade routes known as the Silk Route (Figure 1), where they came into contact with other wild apples, such as *Malus baccata* (L.) Borkh. in Siberia, *Malus orientalis* Uglitz. in the Caucasus, and *Malus sylvestris* Mill. in Europe (Box 1). These three crabapple species and *M. sieversii* are considered to be the closest relatives of *M. domestica*, with which they are fully interfertile [12]. The four wild apple relatives are widely distributed across temperate areas in Eurasia where they often grow as low-density populations (except for *M. sieversii*) in a wide range of habitats and environmental conditions (Box 1).

Several recent studies have used molecular markers to investigate the diversity, structure, and dispersal capacities of wild apple relatives (Box 2) other than *M. baccata*, which has been little studied to date. They have provided highly relevant and timely information, helping to identify and protect germplasm sources for apple breeding for traits

Box 1. Ecology of wild relatives of the cultivated apple

Three decades of field sampling expeditions in Eurasia by different research groups have provided a clearer picture of the geographic distributions and ecology of wild apple relatives (Figure I). The four wild apple species, *Malus sylvestris*, *Malus orientalis*, *Malus sieversii*, and *Malus baccata*, are mostly pollinated by bees and flies [36,51]. Diverse wild animals, including mammals and large birds, feed on the fruits, but their respective efficiencies as seed-dispersal vectors are unknown [1,36]. The wild apple species bear plentiful fragrant white flowers and red to yellow fruits of about 1–6 cm in diameter (Figure I) that are variable in shape, color, and taste. The fruits of *M. orientalis*, *M. baccata* (http://www.efloras.org/), and *M. sylvestris* are edible and were probably already used before the spread of cultivated apple [1].

M. sieversii (Ldb.) Roem grows at intermediate elevations (typically 900–1600 m) in the mountainous regions of Central Asia, and is extremely variable in growth habit, height, fruit quality, and fruit size [1,10,35]. Trees growing in this area, on their own roots, can bear fruit approaching the size of commercial cultivars (i.e., >60 mm diameter), with excellent characteristics that are highly appreciated locally. Currently, *M. sieversii* remains in only a few small intact forests of no more than several hundred hectares in extent in the Tian Shan, at the border of Kazakhstan and China, in southern Kazakhstan, with further small stands in Kyrgyzstan (Figure I). It is also reported to occur sporadically across an area comprising Turkmenistan, Uzbekistan, Tajikistan, and northeastern Afghanistan (Figure I). *M. sieversii* populations experienced severe overharvesting during the Soviet era, and the remaining populations are further threatened by forest destruction [18].

M. sylvestris Mill. occurs across Western and Central Europe [52] (Figure I). The European crabapple has a high light requirement, resulting in its occurrence mostly at forest edges, in farmland hedges, and on marginal sites. It grows on almost all soils but prefers the wet edge of the forest (http://www.euforgen.org/). This species has suffered from the abandonment of coppicing, a traditional forest management practice designed to rejuvenate old trees or shrubs by cutting them to ground level, thus promoting the regeneration of new stems from the base. The European crabapple is now considered endangered in Belgium, the Czech Republic, and Germany [52].

M. orientalis Uglitzh occurs in the Caucasian region (Turkey, Armenia, Georgia, and southwestern Russia) but little is known about its ecology (Figure I).

M. baccata occurs across Siberia and South Asia (India, Pakistan, and Nepal) and is probably relatively cold-tolerant, although little is known about its ecology (Figure I).

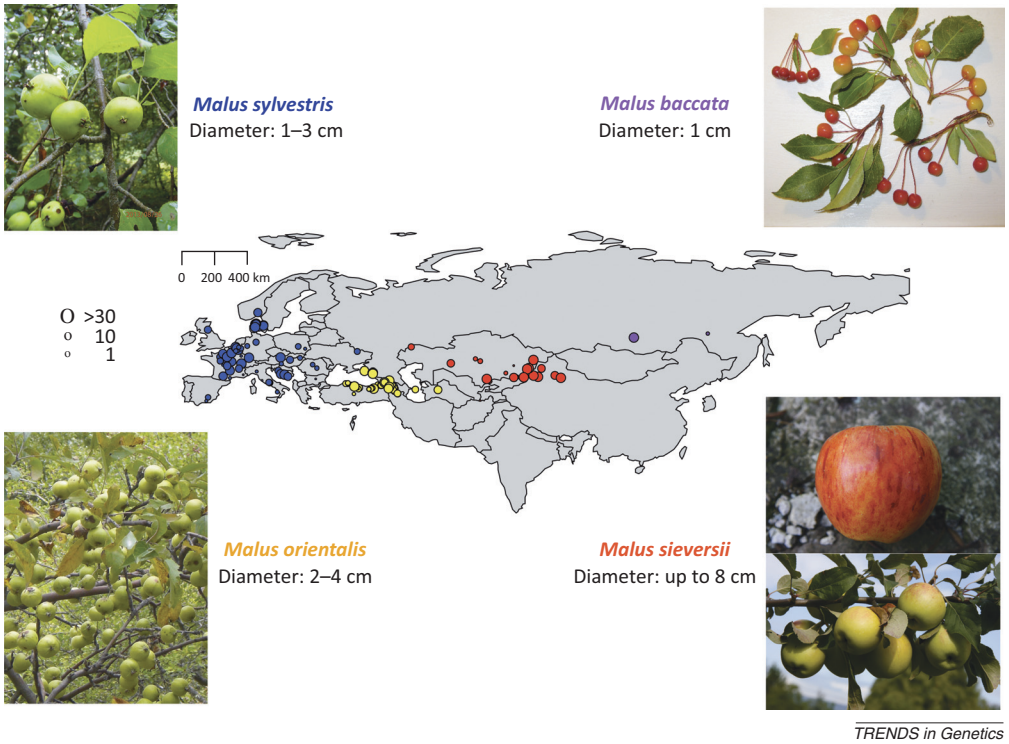

Figure I. Distribution and key morphological features of wild apples. The distribution of species can be inferred from the geographic origin of accessions included in published studies: *Malus sylvestris* (blue) [15,17,20,34,36,36], *Malus orientalis* (yellow) [15,17,20–22], *Malus sieversii* (red) [5,14,15,17–20,53], and *Malus baccata* (purple) [15]. Disk areas are proportional to the number of accessions. Pictures of the fruit of the different wild apples are provided, as well as their respective diameters.

Box 2. Phylogeography, population structure, and dispersal of wild apples

The population structure of the four wild apple species has recently been investigated with microsatellites [17,19,34]. A weak spatial genetic population structure was detected for each species, consistent with the high level of interpopulation gene flow. Estimates of historical gene flow have suggested that wild apple relatives have high dispersal capacities, but further investigations of contemporary and historical pollen and seed dispersal distances are required.

The European crabapple, *Malus sylvestris*, is the wild apple relative with the most thoroughly investigated biogeographic history. *M. sylvestris* experienced range contraction in glacial refugia in Southern Europe during the last glacial maximum. At the beginning of the Holocene, when the climate became warmer in Europe, it recolonized Northern Europe. Three main populations have been identified – in Western Europe, around the Carpathian Mountains, and in the Balkan Peninsula – with admixture in their suture zones [17]. Further investigation, including high-density sampling in France, has revealed an alternative structure, with five distinct clusters in Italy, western France/Great Britain/Belgium, eastern France, Sweden/Denmark/Norway, and the Balkans (Cornille *et al.*, unpublished).

Malus sieversii exists as a main population spread over Central Asia and a smaller population (101 individuals) in the Tian Shan Mountains [20]. Comparison of the population structure of *M. sieversii* with that of one of its major fungal pathogens, *Venturia inaequalis*, revealed similar patterns of genetic differentiation in the *M. sieversii* forests of the eastern mountains of Kazakhstan [54]. This correspondence suggests possible co-structuring of the populations of the main apple progenitor, *M. sieversii*, and its chief pathogen. The use of another dataset (949 individuals) [19] revealed stronger population differentiation in Kazakhstan than previously reported [20], with two large populations and two others with narrow distributions. The difference in sampling density (101 versus 949 individuals, respectively) may account for these apparent differences in population structure.

Malus orientalis displays a weak north–south spatial genetic structure, with three distinct populations: a large population corresponding to most of the Armenian samples and two more narrowly distributed populations, one in the Southern Caucasus (Turkey and southern Armenia) and the other at more northerly latitudes (in Russia).

The population structure of *Malus baccata* has not yet been investigated, mainly due to the lack of samples across its distribution.

The four wild relatives may have geographic distributions contiguous with those of other wild species [1], but detailed information about the distribution of these other wild populations is lacking.

such as resistance to pathogens (e.g., *Venturia inaequalis*, the fungus causing apple scab, [16]). Population genetic analyses of large collections of wild apple relatives have uncovered signatures of postglacial recolonization from distinct glacial refugia in *M. sylvestris* [17], and high levels of diversity and complex population structures in *M. sieversii* [10,18–20] and *M. orientalis* [20–23]. Genetic data have also revealed weak isolation by distance (Box 2). Ubiquitous and high levels of introgression from domesticated apple were found in three of the four wild apple relatives (Box 3), revealing a threat to their genetic integrity. Thus, wild apple relatives clearly have a high dispersal capacity, and the widespread cultivation of domesticated apples in temperate latitudes has promoted extensive crop-to-wild gene flow. The mating system of apple species, characterized by a self-incompatibility mechanism that ensures outbreeding, may have played a crucial role in these introgressions. This mechanism is controlled by a single locus, the *S* locus, with a large number of alleles, up to several hundred in some species. In apples, individual trees can only mate with other individuals with different *S* alleles. The self-incompatibility of apples may thus have favored interspecific hybridization and introgression not only by forcing outcrossing, but also through adaptive introgression, because *S* locus alleles that introgress from a closely related species, if rare in the recipient species, can rapidly spread by positive selection [24,25].

Population genetic methods are increasingly used to resolve taxonomic uncertainties or to identify misclassified germplasm in apple collections. Admixture analyses have revealed that *Malus kirghisorum*, the wild apple from the forests of Kyrgyzstan, and *Malus asiatica*, the apple formerly cultivated in China, are in fact indistinguishable from *M. sieversii* [14]. Multilocus microsatellite typing of the US Department of Agriculture (USDA)–Agricultural Research Service (ARS)–NPGS (National Plant Germplan System) collection (Geneva, USA) suggested that hybrids and misclassifications may also be common in germplasm repositories [26,27]. Further surveys of the diversity and distribution of 'wild' apples should facilitate the conservation of wild genetic resources *in situ* (e.g., conservation of genetically differentiated populations) or *ex situ* (e.g., establishment of core collections of pure wild individuals, maximizing genetic diversity), ultimately enabling their optimal use [27,28].

The enduring riddle of the genetic makeup of cultivated apples

Do 'domesticated' apples actually exist?

The debate concerning whether a crop or animal 'species' should be considered 'domesticated' is similar to that regarding the species concept [29]. Definitions of the boundaries of domesticated taxa are essential in studies of domestication because the classification of domestication-related traits and of the underlying genes can have a major impact on evolutionary inferences. De Queiroz [6] argued that all modern biologists agree that species correspond to segments of evolutionary lineages evolving independently from each other. The seemingly endless dispute about species concepts stems from the confusion between species definition and species criteria, with different recognition criteria corresponding to different events occurring during lineage divergence, rather than to fundamental differences in what is considered to represent a species [6]. Applying a similar reasoning to the domestication process, domesticated species can be defined as segments of evolutionary lineages diverging from their wild progenitors in response to artificial selection pressures and human control over reproduction. Domestication, similarly to speciation, is often quantitative in nature, and different means of quantifying lineage independence can be used to measure different arbitrary 'stages' along this continuum [29–31]. The later stages of divergence along the domestication continuum are characterized by the fixation of marked discontinuities between crops and their wild ancestors (i.e., 'domestication syndrome') which can be used to classify plants categorically as 'domesticated'. Earlier stages of divergence are more readily detected with rapidly evolving genetic markers.

Box 3. Crop-to-wild gene flow in apple

Self-incompatibility and large dispersal capacities make the genus *Malus* an outstanding system for the study of crop-to-wild and interspecific gene flow. The differentiation between domesticated and wild gene pools makes it possible to track cultivated varieties with genetic tools, facilitating inferences about population history such as movement, population subdivision, hybridization, and introgression.

Multilocus microsatellite typing and Bayesian clustering methods, making use of the differentiation between domesticated and wild gene pools, recently revealed the existence of extensive crop-to-wild gene flow in apples [20,27,34,36,41]. Most studies have focused on *Malus domestica* to *Malus sylvestris* gene flow (Table I). Statistical models investigating the factors (environmental and human) influencing the extent of crop-to-wild introgression in apple (gene flow from *M. domestica* to the European crabapple) have revealed a significant positive effect of the number of apple orchards on crop-to-wild introgression rates in the European crabapple (Cornille *et al.*, unpublished). Thus, human activity influences rates of hybridization and introgression from the cultivated apple to *M. sylvestris*. Recent studies have also investigated crop-to-wild gene flow from modern commercial varieties imported from Western countries to *Malus sieversii* and *Malus orientalis* [15,20]. Introgression from the cultivated apple into the wild progenitor *M. sieversii*, and to a lesser extent into *M. orientalis*, has also been detected. The high potential for crop-to-wild gene flow must be considered in the management of wild apple resources.

Phenotypic data have been reported in parallel with genetic data and have been compared between wild and cultivated apple species (Table I). The results obtained for molecular markers matched, to some extent, the morphological traits studied, but no clear-cut morphological criteria distinguishing between cultivated apple and the European crabapple were identified [34,36]. Comparisons of phenotypic data showed that admixed *M. sieversii* individuals had, on average, larger fruits than 'pure' wild individuals [27]. However, it is difficult to determine the extent to which these admixed individuals are actually feral trees.

Table I. Crop-to-wild gene flow in apple[a]

Species	*N*	Geographic area	DNA marker[b]	Percentage of hybrids	Phenotypic data	Refs
Malus sylvestris	44	Belgium, Germany	SSR, AFLP	6.8	Hairiness of the underside of the leaf	[34]
	178	Denmark	SSR	11.2	Fruit diameter, fruit color, pubescence of terminal part of long shoots, pubescence of abaxial surface of leaves from long shoots, and pubescence of abaxial surface of leaves from spur shoots	[36]
	159	The Netherlands, Belgium		17	–	[41]
	61	Germany, Macedonia		19.7		[27]
	796	Austria, Belgium, Bosnia Herzegovina, Bulgaria, Denmark, France, Germany, Great Britain, Hungary, Italy, Norway, Poland, Romania, Ukraine		36.7	–	[20]
Malus orientalis	46	Russia, Turkey	SSR	13.0		[27]
	217	Turkey, Armenia, Russia		3.2	–	[20]
Malus sieversii	118	Kazakhstan	SSR	14.4	Mean fruit weight, length, and width	[27]
	101	Kazakhstan, China, Kyrgyzstan, Tajikistan and Uzbekistan		14.8	–	[20]

[a]Material features of each study: wild species, sampling size (*N*), geographic area covered, molecular and phenotypic data used, percentage of hybrids.

[b]Abbreviations: AFLP, amplified fragment length polymorphisms; SSR, simple sequence repeats.

The domestication process, which typically involved open pollination and unconscious selection over thousands of years, has been followed by the recent process of 'crop improvement', in which breeders intentionally carry out crosses for the selection of new traits in crop varieties [32]. Furthermore, crop improvement traits are typically variable among the cultivars of a crop [33]. The modern breeding phase can thus complicate studies of the genetic basis of phenotypic changes occurring during the initial domestication phase. However, it is theoretically possible to distinguish between the two phases of the evolution of domesticated plants by inferring the age and origin of alleles contributing to agronomically important traits.

Analyses of population structure based on comprehensive samples of cultivated and wild gene pools have revealed that the cultivated apple forms a distinct, panmictic group, well separated from its Central Asian progenitor, *M. sieversii*, with similar levels of genetic variation in the two groups [15,27]. Divergent selection between wild and cultivated gene pools, extensive selection on domesticates by humans outside the center of origin of the crop, and vegetative propagation have all contributed to a reduction of gene flow between cultivated apples and their progenitors, and the resulting pattern of clear population differentiation provides definitive evidence for the domesticated 'nature' of apples. Several studies have investigated phenotypic variation in close relatives of *M. domestica* [10,23,27,34–36] but, to the best of our knowledge, no attempt has been made to quantify the morphological and physiological changes associated with apple domestication. These phenotypic changes are expected to be weaker in apples and other long-lived perennial fruit crops than in seed-propagated annuals because the lengthy juvenile phase and the use of grafting make domestication more recent, at least in terms of the number of generations [37]. Nevertheless, several traits clearly distinguish

domesticated long-lived perennials from their wild ancestors (see below), and the phenotypic changes associated with domestication may parallel those in annual crops [30].

Archaeological evidence

Several scholars have evaluated the archaeological and historical evidence relating to the origin and use of apples [1–3,12]. Evidence for the collection of wild apples has been found at Neolithic and Bronze Age archaeological sites across Europe. Similar evidence suggesting the early use of apples in Western and Central Asia is lacking, but some 'anomalies' in the distribution of *M. domestica*, such as isolated patches of large sweet apples (i.e., different from native bitter crabapples) in the Caucasus Mountains, the Crimean Peninsula, parts of Afghanistan, Iran, Turkey, and the Kursk region of European Russia, and the large, possibly 3000-year-old apple discovered at Navan Fort (Northern Ireland), might be hallmarks of early apple seedling imports from Central Asia [1]. Advances in ancient DNA technology should facilitate diachronic investigations of the domesticated status of apple seed remains in the Neolithic and Bronze Age archaeological record across Eurasia, providing information about the transition from gathering to cultivation. The recent amplification of ribosomal DNA from apple remains – seed coat (testa) and pericarp – preserved in waterlogged layers from a Roman site in Switzerland suggests that it may be feasible to identify wild and domesticated apple from the archaeological record [38]. Insight into the human behavior responsible for the changes associated with domestication can also be gained by studying current practices such as the exploitation of wild apples in some areas of Western and Central Asia.

Grafting, as a handy way to propagate elite cultivars, has probably played a key role in the spread of apple as a major fruit crop worldwide [1,11]. Grafting is a practice that is thought to have begun about 3800 years ago based on a cuneiform description of budwood importation for grape in Mesopotamia. Indirect evidence has been obtained for the cultivation of apples 3000 years ago in Mesopotamia [12], but the most compelling evidence for horticulture and for apple cultivation dates from the Greek period (i.e., from about the 3rd century BC) [1]. The Romans probably learned apple grafting, cultivation, harvesting, and storage from the Greeks, and brought the production chain technology to the rest of their empire. Dwarf apple trees were known 2300 years ago, and dwarfing rootstocks were probably brought to the West from the Caucasus. In the 19th century, dwarfing clones known as Paradise (or French Paradise) or Doucin (or English Paradise) were common in Europe, before the East Malling Research Station in England created a standardized collection of 10 rootstocks with new names. Two of the original East Malling selections, M9 (Paradis Jaune de Metz) and M7 (Doucin Reinette), are still widely used by horticulturists worldwide.

Genetic evidence

Having originated from *M. sieversii* in Central Asia about 4000 to 10 000 years ago, the cultivated apple then underwent hybridization with its wild relatives during its spread from the Tian Shan Mountains westward along the Silk Route (Figure 1). The different timescales of the domestication process (ancient initial domestication in Central Asia followed by more recent hybridization during the route of spread of apple cultivation) have been reconstructed from panels of markers tracing back different evolutionary timescales (i.e., chloroplast, nuclear, and microsatellite markers). These genetic markers have revealed original patterns of diversity and differentiation within cultivated apple, and non-trivial relationships with its wild relatives. The various steps in the domestication of apple unraveled by the analysis of genetic markers are outlined below.

(i) *Origin in the Tian Shan Mountains of Central Asia: was* Malus sieversii *the progenitor?*

The morphological similarity between *M. domestica* and *M. sieversii*, and the extraordinary diversity of wild apples in Kazakhstan, were first reported by Vavilov [13] and then by Ponomarenko (1983, cited in [39]). These findings were recently confirmed by collaborative collection expeditions involving American, European, and Central Asian scientists [10,14,15,39]. The genetic data were found to be consistent with these morphological observations. Initial phylogenetic analyses of chloroplast and nuclear sequences confirmed a progenitor–descendant relationship, with *M. sieversii–M. domestica* being the pair of species most closely related genetically, based on sequences, and possibly not even distinguishable, but conclusions about the origin of this crop were hampered by the lack of strong statistical support and/or limited sample size [11,14]. Analyses of microsatellite data [i.e., simple sequence repeat (SSR) markers] for a comprehensive collection of wild and cultivated apples have provided new insight, revealing that *M. domestica* forms a distinct, panmictic group, well separated from *M. sieversii* [15,40]. This may indicate that the populations of *M. sieversii* from which the crop was domesticated have yet to be identified or may no longer exist. *Malus sieversii* is unusual in that it grows at high density over large areas, but these forests have been decimated since Vavilov reported their existence. The substantial genetic differentiation between *M. domestica* and *M. sieversii* may also result from a combination of genetic drift and introgressions from other wild apple species, erasing the genetic footprints of the initial contribution of *M. sieversii*. Genomic studies and further sampling of wild apple and cultivar diversity should provide quantitative and qualitative insight into the origin of *M. domestica*, and improve our understanding of the genetic architecture and basis of the phenotypic changes selected during domestication and subsequent crop improvement.

(ii) *Diversification of the domesticated apple along the Silk Route*

Several studies have highlighted the importance of the European crabapple, *M. sylvestris*, in the evolution and diversification of the cultivated apple and, to a lesser extent, that of the Caucasian crabapple *M. orientalis*, through the use of microsatellite markers and nuclear or chloroplast sequences [15,36,41,42].

M. sieversii was probably the initial progenitor, but other wild apple species subsequently contributed to the genetic makeup of cultivated apple through introgressive hybridization, affecting both nuclear and chloroplast DNA, during its dispersal westward along the Silk Route (Figure 1). These introgressions obscure the signal of any initial bottleneck occurring during the original domestication of *M. domestica*. Introgression has been so extensive that *M. domestica* now appears to be more similar genetically to the European crabapple *M. sylvestris* than to the Asian wild apple *M. sieversii* on the basis of microsatellite markers [15]. Most, but not all, apple cultivars now also appear to contain the chloroplast genome of *M. sylvestris* [42]. Genetic evidence has also demonstrated that current cider cultivars, which usually produce smaller fruits that are more bitter than those of dessert cultivars, are not the cultivars genetically closest to *M. sylvestris* [15], as initially hypothesized. This hypothesis was based on the known use of crabapples for the preparation of apple beverages, even before apple cultivation [1], and the astringent, bitter properties of these small apples, ideal for cider-making. Dessert apple cultivars were actually found to be genetically closer to *M. sylvestris* than cider cultivars in studies using nuclear microsatellite markers [15]. Because cider is made and consumed across Eurasia, cider apple cultivars may also have originated in Central Asia where wild apple populations are highly diverse in terms of color, taste, and size.

Breeding methods for fruit tree crops and the phenotypic changes associated with apple domestication

The life-history traits of apples make it harder to fix the desired agronomic traits quickly in this crop than in annual seed-propagated crops. In particular, apples have long generation times and display obligate outcrossing due to their self-incompatibility system. The development of appropriate breeding methods (i.e., open-pollination followed by the grafting of interesting phenotypes) resulted in a model of domestication fundamentally different from that for annual seed-propagated crops.

Theoretically, the introduction of clonal propagation may have rapidly decreased the genetic diversity of cultivated apples because grafting permitted the genetic fixation of a limited set of elite clones that could only diversify by accumulating somatic mutations. However, no evidence of a domestication bottleneck or clonal population structure has been found in apple in studies considering a single representative of each variety [15]. This raises questions about how high levels of diversity have been maintained, and about the lack of a clonal or relatedness structure between cultivars. Long-prevailing apple breeding methods may account for this apparent paradox. If many farmers each independently selected trees producing good fruits from progeny arising from natural pollination ('chance seedlings' [10]), the obligate outcrossing of the crop, the isolation of farms, and differences in taste preferences across regions may have been powerful forces driving the maintenance of high levels of variation. The occasional selection of plants resulting from hybridization with wild relatives would have further increased diversity.

These traditional methods may have contributed to a high level of diversity in the cultivated gene pool, but current breeding practice methods, in which a small set of cultivars is common to the pedigree of many new cultivars, may be eroding the genetic diversity available in the cultivated genepool. For instance, a genetic analysis of a comprehensive sample of commercial cultivars revealed extensive first-, second- and third-degree relationships between popular apple varieties [43], as previously reported for grape [44]. This highlights the need to make better use of the diversity in apple germplasm [28,45]. Conversely, the current situation may also reflect the extensive use of genetic diversity by breeders, but with the cultivars generated replacing each other in the production area rather than being grown alongside older varieties. This is the situation that applies for many vegetable and arable crops in which genetic diversity remains at a high level if considered over a longer period of time.

Apple phenotypes selected by humans include: (i) higher productivity (e.g., number of fruits, fruit size, fruit weight); (ii) fruit quality (e.g., color, shape), flavor (volatile compounds), taste (e.g., fruit acidity, sugar content, polyphenolics), texture (fruit flesh firmness), and conservation (storage behavior); (iii) easier harvesting (e.g., growth habit, plant size); and (iv) greater phenological congruence with agricultural practices (e.g., shorter juvenile phases, synchronicity in blooming or fruit ripening). Crop improvement and cultivar diversification have affected many characteristics, including juvenile phase length, cold requirement, cold tolerance, drought tolerance, fruit tenacity, and disease resistance [3]. Rootstocks may also have been subject to selection for their effects on the productivity of aerial parts and their role in tolerance to edaphic, climatic, and biotic conditions. The domestication and improvement of cultivated apples may have benefited from the high heritability of quality traits [46].

Quantitative trait locus (QTL) mapping has been used to dissect the genetic architecture of several desired traits through crosses between cultivars [47]. However, QTL mapping has not been used to investigate the genetic architecture of domestication traits through wild × domesticated biparental crosses. Given the high costs of developing and maintaining mapping populations in apple, alternatives, such as candidate gene and bottom-up approaches, relying on population genetic analyses and functional information to connect selected (candidate) genes to phenotypes, are likely to be the best options for identifying the genomic regions that have undergone the greatest change during domestication [48,49]. The recently released 'Golden Delicious' genome sequence [14], and the availability of large-scale genotyping tools [50], have made it possible to generate population-scale data for both domesticated and wild apples.

Concluding remarks

Recent studies of the evolutionary history of the domesticated apple and its wild relatives have shown this to be an outstanding biological model for exploring the evolutionary

process at work during domestication. Full-genome resequencing and reference-assisted *de novo* assembly of domesticated and wild apple genomes will undoubtedly provide further insight into the history of this remarkable crop, facilitating the genetic dissection of agronomically and ecologically relevant traits. The main limiting factor may now be the availability of suitable accessions for evolutionary inferences and plant breeding rather than the development of genomic tools, highlighting the need for further field collections. Future sampling campaigns should focus on: (i) worldwide rootstock accessions – the origin of rootstocks is unknown despite their key role in the tolerance of cultivars to abiotic conditions, (ii) local and ancient cultivars from different parts of the world (Eastern Europe, Mediterranean, and Asia), and (iii) wild unsampled species and populations, from Central Asia in particular.

Another unanswered question concerns the adaptive significance of the high levels of wild-to-crop gene flow: has human selection unwittingly favored gene flow from wild species? Do introgressions localize in neutral regions of the crop genome or at regions of functional significance? Regarding crop-to-wild gene flow, has natural selection retained *M. domestica* alleles introgressed into wild apples? Conversely, is introgression deleterious for crabapples, threatening wild populations, or will introgressions from the domesticated species be selected against?

The application of population and comparative genomics to wild and cultivated apples provides new opportunities to investigate the molecular basis and genetic architecture of both (i) agronomic traits potentially selected and introgressed during domestication (e.g., fruit size, color, volatile content), and (ii) ecological traits contributing to the adaptation of wild apple populations to environmental and biotic factors (e.g., temperature, drought, pathogens). Population and comparative genomics may also reveal which traits were selected during the early steps of domestication and which were selected during the modern breeding phase. Breeding programs will probably benefit from the incorporation of candidate alleles identified by evolutionary analyses, and there may be an increase in diversity from exotic germplasm in key genomic regions (such as apparently less diverse candidate regions). The ecological genomics of apples *sensu lato* will, therefore, not only advance our understanding of the history of this fascinating system but will also guide and accelerate the development of new breeding strategies for increasing the volume and quality of apple production in the face of dynamic abiotic and biotic threats.

Acknowledgments

We thank the editor and two referees for comments on a previous version of this manuscript. We thank Ilya Zacharov for photographs of *M. baccata* and Pascal Heitlzer for photographs of *M. sieversii*.

References

1 Juniper, B.E. and Mabberley, D.J. (2006) *The Story of the Apple*, Imber Press
2 Morgan, J. and Richards, A. (2003) *The New Book of Apples*, Brogdale Horticultural Trust, Ebury Press
3 Janick, J. (2005) The origins of fruits, fruit growing, and fruit breeding. *Plant Breed. Rev.* 25, 255–321
4 Gardiner, S.E. *et al.* (2007) Apple. In *Genome Mapping and Molecular Breeding in Plants, Fruits and Nuts* (Kole, C., ed.), pp. 1–62, Springer
5 Robinson, J.P. *et al.* (2001) Taxonomy of the genus *Malus* Mill. (Rosaceae) with emphasis on the cultivated apple, *Malus domestica* Borkh. *Plant Syst. Evol.* 226, 35–58
6 De Queiroz, K. (1999) The general lineage concept of species, species criteria, and the process of speciation. In *Endless Forms: Species and Speciation* (Howard, D.J. and Berlocher, S.H., eds), pp. 49–89, Oxford University Press
7 Korban, S.S. (1986) Interspecific hybridization in *Malus*. *HortScience* 21, 41–48
8 Way, R.D. *et al.* (1991) Apples (*Malus*). *Acta Hort.* 290, 3–46
9 Watkins, R. (1995) Apple and pear. In *Evolution of Crop Plants* (Smartt, J. and Simmonds, N.W., eds), pp. 418–422, Longman
10 Forsline, P.L. *et al.* (2003) Collection, maintenance, characterization and utilization of wild apples of Central Asia. *Hort. Rev.* 29, 1–61
11 Harris, S.A. *et al.* (2002) Genetic clues to the origin of the apple. *Trends Genet.* 18, 426–430
12 Zohary, D. and Hopf, M. (2000) *Domestication of Plants in the Old World*. (3rd edn), Oxford University Press
13 Vavilov, N.I. (1926) Studies on the origin of cultivated plants. *Tr. Byuro Prikl. Bot.* 16, 139–245
14 Velasco, R. *et al.* (2010) The genome of the domesticated apple (*Malus* × *domestica* Borkh.). *Nat. Genet.* 42, 833–839
15 Cornille, A. *et al.* (2012) New insight into the history of domesticated apple: secondary contribution of the European wild apple to the genome of cultivated varieties. *PLoS Genet.* 8, e1002703
16 Lê Van, A. *et al.* (2013) Differential selection pressures exerted by host resistance quantitative trait loci on a pathogen population: a case study in an apple × *Venturia inaequalis* pathosystem. *New Phytol.* 197, 899–908
17 Cornille, A. *et al.* (2013) Post-glacial recolonization history of the European crabapple (*Malus sylvestris* Mill.), a wild contributor to the domesticated apple. *Mol. Ecol.* 22, 2249–2263
18 Zhang, C. *et al.* (2007) Genetic structure of *Malus sieversii* population from Xinjiang, China, revealed by SSR markers. *J. Genet. Genomics* 34, 947–955
19 Richards, C. *et al.* (2009) Genetic diversity and population structure in *Malus sieversii*, a wild progenitor species of domesticated apple. *Tree Genet. Genomes* 5, 339–347
20 Cornille, A. *et al.* (2013) Crop-to-wild gene flow and spatial genetic structure in the closest wild relatives of the cultivated apple. *Evol. Appl.* 6, 737–748
21 Volk, G.M. *et al.* (2008) Genetic diversity and disease resistance of wild *Malus orientalis* from Turkey and Southern Russia. *J. Am. Soc. Hortic. Sci.* 133, 383–389
22 Volk, G.M. *et al.* (2009) Capturing the diversity of wild *Malus orientalis* from Georgia, Armenia, Russia, and Turkey. *J. Am. Soc. Hortic. Sci.* 134, 453–459
23 Höfer, M. *et al.* (2013) Assessment of phenotypic variation of *Malus orientalis* in the North Caucasus region. *Genet. Resour. Crop Evol.* 60, 1463–1477
24 Castric, V. *et al.* (2008) Repeated adaptive introgression at a gene under multiallelic balancing selection. *PLoS Genet.* 4, e1000168
25 Joly, S. and Schoen, D.J. (2011) Migration rates, frequency-dependent selection and self-incompatibility locus in *Leavenworthia* (Brassicaceae). *Evolution* 65, 2357–2369
26 Van Treuren, R. *et al.* (2010) Microsatellite genotyping of apple *Malus domestica* Borkh.) genetic resources in The Netherlands: application in collection management and variety identification. *Genet. Resour. Crop Evol.* 57, 853–865
27 Gross, B.L. *et al.* (2012) Identification of interspecific hybrids among domesticated apple and its wild relatives. *Tree Genet. Genomes* 8, 1223–1235
28 Myles, S. (2013) Improving fruit and wine: what does genomics have to offer? *Trends Genet.* 29, 190–196
29 Larson, G. and Burger, J. (2013) A population genetics view of animal domestication. *Trends Genet.* 29, 197–205
30 Miller, A. and Gross, B.L. (2011) From forest to field: perennial fruit crops domestication. *Am. J. Bot.* 98, 1389–1414
31 Nosil, P. (2012) *Ecological Speciation*, Oxford University Press
32 Gross, B.L. and Olsen, K.M. (2009) Genetic perspectives on crop domestication. *Trends Plant Sci.* 15, 529–537

33 Olsen, K.M. and Wendel, J.F. (2013) A bountiful harvest: genomic insights into crop domestication phenotypes. *Annu. Rev. Plant Biol.* 64, 47–70

34 Coart, E. *et al.* (2003) Genetic variation in the endangered wild apple (*Malus sylvestris* (L.) Mill.) in Belgium as revealed by amplified fragment length polymorphism and microsatellite markers. *Mol. Ecol.* 12, 845–857

35 Dzhangaliev, A.D. (2003) The wild apple tree of Kazakhstan. *Hort. Rev.* 00, 63–303

36 Larsen, A. *et al.* (2006) Hybridization and genetic variation in Danish populations of European crab apple (*Malus sylvestris*). *Tree Genet. Genomes* 2, 86–97

37 Pickersgill, B. (2007) Domestication of plants in the Americas: insights from Mendelian and molecular genetics. *Ann. Bot.* 100, 925–940

38 Schlumbaum, A. *et al.* (2012) Towards the onset of fruit tree growing north of the Alps: ancient DNA from waterlogged apple (*Malus* sp.) seed fragments. *Ann. Anat.* 194, 157–162

39 Luby, J.J. *et al.* (2001) Field resistance to fire blight in a diverse apple (*Malus* sp.) germplasm collection. *J. Am. Soc. Hortic. Sci.* 127, 245–253

40 Gharghani, A. *et al.* (2009) Genetic identity and relationships of Iranian apple (*Malus domestica* Borkh.) cultivars and landraces, wild *Malus* species and representative old apple cultivars based on simple sequence repeat (SSR) marker analysis. *Genet. Resour. Crop Evol.* 56, 829–842

41 Koopman, W.J.M. *et al.* (2007) Linked vs. unlinked markers: multilocus microsatellite haplotype-sharing as a tool to estimate gene flow and introgression. *Mol. Ecol.* 16, 243–256

42 Nikiforova, S.V. *et al.* (2013) Phylogenetic analysis of 47 chloroplast genomes clarifies the contribution of wild species to the domesticated apple maternal line. *Mol. Biol. Evol.* 30, 1751–1760

43 Noiton, D.A.M. and Alspach, P.A. (1996) Founding clones, inbreeding, coancestry, and status number of modern apple cultivars. *J. Am. Soc. Hortic. Sci.* 121, 773–782

44 Myles, S. *et al.* (2011) Genetic structure and domestication history of the grape. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3530–3535

45 Feuillet, C. *et al.* (2008) Cereal breeding takes a walk on the wild side. *Trends Genet.* 24, 24–32

46 Kouassi, A. *et al.* (2009) Estimation of genetic parameters and prediction of breeding values for apple fruit-quality traits using pedigreed plant material in Europe. *Tree Genet. Genomes* 5, 659–672

47 Maliepaard, C. *et al.* (1998) Aligning male and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. *Theor. Appl. Genet.* 97, 60–73

48 Morrell, P.L. and Clegg, M.T. (2007) Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. *Proc. Natl. Acad. Sci. U.S.A.* 104, 3289–3294

49 Ross-Ibarra, J. *et al.* (2007) Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8641–8648

50 Chagné, D. *et al.* (2012) Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLoS ONE* 7, e31745

51 Deguines, N. *et al.* (2012) The whereabouts of flower visitors: contrasting land-use preferences revealed by a country-wide survey based on citizen science. *PLoS ONE* 7, e45822

52 Jacques, D. *et al.* (2009) Natural distribution and variability of the wild apple (*Malus sylvestris*) in Belgium. *Belg. J. Bot.* 142, 39–49

53 Volk, G.M. *et al.* (2005) *Ex situ* conservation of vegetatively propagated species: development of a seed-based core collection for *Malus sieversii*. *J. Am. Soc. Hortic. Sci.* 130, 203–210

54 Gladieux, P. *et al.* (2010) Evolution of the population structure of *Venturia inaequalis*, the apple scab fungus, associated with the domestication of its host. *Mol. Ecol.* 19, 658–674