



HAL
open science

A multifaceted overview of apple tree domestication

Amandine Cornille, Ferran Antolín, Elena García, Cristiano Vernesi, Alice Fietta, Otto Brinkkemper, Wiebke Kirleis, Angela Schlumbaum, Isabel Roldán-Ruiz

► **To cite this version:**

Amandine Cornille, Ferran Antolín, Elena García, Cristiano Vernesi, Alice Fietta, et al.. A multifaceted overview of apple tree domestication. *Trends in Plant Science*, 2019, 24 (8), pp.770-782. 10.1016/j.tplants.2019.05.007 . hal-02357023

HAL Id: hal-02357023

<https://hal.science/hal-02357023>

Submitted on 9 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **A multifaceted overview of apple tree domestication**

2

3 Cornille Amandine^{1*}, Ferran Antolín², Elena Garcia³, Cristiano Vernesi^{4a}, Alice Fietta⁴, Otto
4 Brinkkemper⁵, Wiebke Kirleis⁶, Angela Schlumbaum² and Isabel Roldán-Ruiz^{7,8}

5

6 1. GQE– Le Moulon, INRA, Univ. Paris-Sud, CNRS, AgroParisTech, Université Paris-Saclay,
7 91190, Gif-sur-Yvette, France

8 2. Integrative Prehistory and Archeological Science (IPNA/IPAS), Department of Environmental
9 Sciences, University of Basel. Spalenring 145 CH-4055 Basel (Switzerland)

10 3. Department of Horticulture, University of Arkansas, Fayetteville. AR, USA

11 4. Department of Biodiversity and Molecular Ecology, Research and Innovation Centre -
12 Fondazione Edmund Mach, via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy

13 5. Cultural Heritage Agency, PO Box 1600, 3800 BP Amersfoort, The Netherlands.

14 6. Institute for Prehistoric and Protohistoric Archeology/Graduate School Human Development in
15 Landscapes, Christian-Albrechts-University Kiel, Kiel, Germany.

16 7. Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Plant Sciences Unit,
17 Caritasstraat 39, B-9090 Melle, Belgium.

18 8. Ghent University, Faculty of Sciences, Dept. of Plant Biotechnology and Bioinformatics,
19 Technologiepark 71, 9052 Ghent, Belgium.

20

21 ^acurrent address: Dept. of Sustainable Agroecosystems and Bioresources, Research and Innovation
22 Centre - Fondazione Edmund Mach, via E. Mach 1, 38010 S. Michele all'Adige (TN), Italy

23

24

25 *Corresponding author: amandine.cornille@inra.fr

26

27

28

29 **Abstract**

30 Apple is an iconic tree and major fruit crop worldwide. It is also a model species for the study of the
31 evolutionary processes and the genomic basis underlying the domestication of clonally propagated
32 perennial crops. Multidisciplinary evidence from across Eurasia has documented the pace and
33 process of cultivation of this remarkable crop. While population genetics and genomics have
34 revealed the overall domestication history of apple across Eurasia, untangling the evolutionary
35 processes involved, archeobotany has helped to document the transition from gathering and using
36 apples to the practice of cultivation. Further studies, integrating archeogenetics and –genomics
37 approaches, will definitively bring new insights about key traits involved in apple domestication.
38 Such knowledge has potential to boost innovation in present-day apple breeding.

39

40

41

42

43

44 **Keywords:** Adaptive introgression, self-incompatibility, rootstock, pathogen, Silk Road, targeted
45 sequencing, plant remains, paleogenetics.

46

47

48 **Glossary**

49 **Archeobotany:** The study of plant remains (seeds, fruits, wood, leaves, pollen, etc.) found in
50 archeological deposits with the aim of reconstructing the paleo-environment and the management of
51 plant resources (including the domestication of wild plants), their uses, and the waste disposal
52 practices of ancient societies.

53 **Archeo/paleogenetics and –genomics:** Research fields that use the genetic information gathered
54 from archeological plant or animal remains (ancient DNA; aDNA) to address questions about the
55 evolution, domestication and history of use and cultivation of plants and animals (e.g. [1–8]).

56 **Balancing selection:** Selection regimes that result in increased genetic diversity relative to neutral
57 expectations in populations of living organisms.

58 **Crabapple [9]:** Wild apple species that usually blossom profusely and produce small, acidic fruits.
59 The word crab comes from the Old English ‘crabbe’ meaning bitter or sharp tasting. Many
60 crabapples are cultivated as ornamental trees. In Western Europe the term crabapple is often used to
61 refer to *Malus sylvestris* (the European crabapple), in the Caucasus to *M. orientalis* (the Caucasian
62 crabapple) and, in Siberia, to *M. baccata* (the Siberian crabapple). The native North American

63 crabapples are *M. fusca*, *M. coronaria*, *M. angustifolia*, and *M. ioensis*. *Malus sieversii*, the main
64 progenitor of the cultivated apple, is usually not referred to as a ‘crabapple’.

65 **Grafting:** A technique of vegetative propagation whereby tissues of two plants are intentionally
66 joined and allowed to grow together. The upper part of the combined plant (which contributes the
67 aerial parts and produces fruit) is called the scion and the lower part (which provides the roots) is
68 called the rootstock.

69 **Introgression:** The transfer of genomic regions from one species into the gene pool of another
70 species through an initial hybridization event followed by repeated backcrosses.

71 **Taphonomy:** A number of agents that determine the presence and preservation quality of items
72 found in sedimentary deposits.

73 **The cultivated apple:** The only domesticated species in the genus *Malus*, here called *Malus*
74 *domestica*. *Malus domestica* is often referred to as *Malus* × *domestica*, the “×” indicating a hybrid
75 origin. However, from what we know so far, the cultivated apple did not originate from a hybrid
76 speciation event. The cultivated apple originated from the Central Asian wild apple *M. sieversii* and
77 was later introgressed by other wild species present along the ancient Silk Road trading routes. We
78 therefore prefer *M. domestica*.

79 **Water flotation:** Laboratory technique used to process sediment samples from archeological
80 deposits using the principle of density where a light fraction and a heavy fraction are separated
81 inside the flotation tank. Usually both fractions are dried after the separation process is finished.
82 When waterlogged sediments are processed, the fractions must be stored in water at a temperature
83 at around 4°C.

84

85

86 **The history of apples: tangled, gnarly and otherwise hard to unravel**

87

88 Recent multi-proxy studies have unraveled the complex histories of seed-propagated crops (e.g.
89 maize [10]), but equivalent research in long-lived, clonally propagating perennials is still lacking.
90 One reason is the difficulty of retrieving suitable ancient remains; another is that genomic resources
91 have only been deployed relatively recently in this type of crop, in comparison of seed-propagated
92 crops. Attempts to combine the perspectives from studies in different disciplines, to make multiple-
93 proxy interpretations of the domestication history of apple are rare. As a consequence, answers to
94 many questions are still left for future research.

95 Apple (*Malus domestica* Borkh.) is an iconic tree and a major fruit crop in Eurasia that
96 supports many local economies. Thousands of apple cultivars are grown worldwide to produce
97 high-quality fruit for the fresh market, a range of beverages and other food products [11] (Box 1).
98 The past decade has seen tremendous progress in the reconstruction of apple history due to the
99 increasing availability of genetic and phenotypic data and archeobotanical information for both wild
100 and cultivated *Malus* species. It is now assumed that the cultivation origins of the apple tree are in
101 the Tian Shan mountains located in Central Asia [9] (Box 1). In classical times, apple cultivation
102 expanded along the Silk Road trade routes linking Asia and Europe and spread from there [9,12,13].
103 The introduction of **grafting** techniques facilitated the fixation and propagation of superior
104 genotypes derived from open pollination (i.e. ‘chance seedlings’), and caused a revolution in the
105 history of apple cultivation [13,14] (Box 2). For some apple species, including the closest wild
106 relatives to the cultivated apple, the combination of molecular markers and phenotypic traits that
107 were initially used to build the systematics of the genus *Malus*, has allowed to confirm that some
108 morphological groups are truly genetically differentiated [15–24]. However, the origin of this crop
109 and its domestication history - in particular the relationships between cultivated and wild apples - is
110 not yet completely understood.

111 Botanical and **archeobotanical** evidence, together with genetic and genomic data, has laid
112 the groundwork for a multidisciplinary re-evaluation of the biogeography of wild apples and the
113 history of the cultivated species. Our previous review [9] highlighted the genetics and ecology of
114 apple domestication and its impact on wild apple evolution. In this review, results obtained using
115 botany, **archeobotany**, **archeogenetics/genomics** and genetics/genomics are merged to start
116 mapping the extensive history of the cultivated apple. We conclude this review with a discussion of
117 how multidisciplinary research, including recent technological advances in genomics, can be
118 applied to further study of the domestication process of the cultivated apple.

119

120 **Using archeobotany and the potential of archeogenetics and archeogenomics to reconstruct**
121 **the history of cultivation of apples**

122

123 By fingerprinting present-day apple samples using DNA markers (i.e. microsatellites), the pace and
124 process of apple domestication and diversification has been reconstructed. These data support the
125 view that **the cultivated apple** was domesticated from the wild apple *Malus sieversii* (Ledeb.) M.
126 Roem. in Central Asia and was then transported westwards along trade routes [9]. Along the way,
127 the cultivated apple hybridized with other wild apple species including *Malus sylvestris* (L.) Mill,
128 the European **crabapple** [16,25]. Further data gathering using **archeo/paleogenetics and –**
129 **genomics** approaches is however still relatively slow [26]. This is due to the shortage of suitable
130 ancient apple remains, slowing the data generation for comparative studies between past and
131 present-day apple samples [9]. Full documentation of this process of domestication, expansion and
132 introgressive hybridization (Box 1) requires suitable amounts and a wide diversity of archeological
133 apple remains from covering the range of the cultivated apple in Eurasia. But ancient human
134 populations did not store apple seeds on purpose at the same scale as this was done for seed-
135 propagated crops. As a result, the number of apple seeds in archeological sites therefore remains
136 low in comparison. In addition, ancient apple seed remains from Central Asian archeological sites
137 dating to around 8-10,000 years Before the Common Era (BCE) (Box 3) are scarce due to
138 unfavorable preservation conditions and lack of excavations in certain regions. Additional data
139 could be gleaned by performing new campaigns to gather vegetal remains using **water-based**
140 **sieving** techniques in some relevant sites that are already being excavated, e.g. [27]. Another
141 possible source is archived ancient apple seed remains already present in collections (archaeological
142 or herbarium) that could be (re-)investigated using DNA-based techniques.

143

144 *Archeobotanical insights on apple cultivation*

145

146 Early apple use by humans started with the gathering of wild apples, then proceeded to increasingly
147 managed wild stands (including cultivation with protection) and ultimately the intentional
148 cultivation of apple trees. However, our knowledge of the evolution of these practices in Eurasia
149 remains patchy at best [28] (Box 3). The knowledge gaps can be explained by a lack of apple plant
150 remains (seeds, fruits, wood or pollen) in archeological sites in the region where apple
151 domestication and diffusion took place. The recovery of apple remains in archeological sites
152 depends heavily on preservation conditions and **taphonomic** agents. Apple seed testa (which have
153 the greatest chance of being identified to genus or species level) are abundant and frequent in sites
154 with waterlogged deposits (lakeshore settlements, wells, etc.) preserved under anoxic conditions

155 (Figure 1). In contrast, dry sites rarely contain apple seed testa remains, as they can only be
156 preserved in a charred or mineralized state. So far, waterlogged archeological deposits have
157 primarily been investigated in Europe. It is relatively common to find apple halves and some whole
158 apples (Figure 1) that were accidentally charred during drying in ovens prior to storage in Europe.
159 These remains would be ideal material for investigating changes in fruit size over time (Figure 1),
160 but this has not yet been attempted. Such charred remains can also be found in dry sites, where their
161 detection depends primarily on the intensity of sampling and the use of **water flotation** to retrieve
162 small plant remains from archeological sediments.

163 This combination of factors has resulted in a higher number of apple seed remains from
164 Europe in comparison to those from Asia. We can now reconstruct apple use before and after the
165 arrival of cultivated apples in Europe (Box 3). Unfortunately, the few finds from Kazakhstan,
166 Mongolia, Nepal, Tajikistan and Turkmenistan only date to the last 3,000 years [27,29,30]. This
167 prevents us from tracing earlier cultivation or domestication in this area, which is supposed to be
168 the home of apple cultivation [9,31]. It is also difficult to distinguish between the remains from wild
169 and domestic forms of apple based on the morphological characteristics of seeds, fruits, pollen or
170 wood alone [32,33] (Figure 1 and Box 3). Often identification to the genus level is not possible,
171 leading to descriptions such as *Malus/Pyrus* or even Maloideae (also referred to as Pomoideae tribe
172 - we will use the term Maloideae hereafter) in case of wood or pollen. In only one medieval site in
173 the Nepalese Johng Valley, two seeds could be determined as *M. baccata* [34]. New digital image
174 analysis techniques such as those applied to the Rosaceae family, combined with genetic markers
175 and genome data (nuclear or chloroplastic, see below), might prove useful for the Maloideae and
176 could improve the available data at the genus level [35].

177

178 *Ancient DNA investigations of apple remains*

179

180 Only two studies have used ancient DNA (aDNA) techniques to infer the history of apple gathering,
181 use and cultivation; both studies used data from European sites [28]. First, Schlumbaum *et al.* [28],
182 investigated waterlogged bulk samples of testa fragments from two Neolithic and two Roman sites
183 in the Alpine foreland using chloroplast and nuclear markers. Chloroplast aDNA, present in many
184 copies in one cell, has more chances to survive than single copy nuclear aDNA [36], but nuclear
185 aDNA bi-parental inheritance has the potential to reveal signatures of hybridization between
186 different species, sub-species and even cultivars. The use of chloroplast markers was unsuccessful
187 in that study, but a nuclear ITS1 region (internal transcribed spacer region 1 of the ribosomal DNA)
188 was successfully typed in two Roman samples, confirming that they belonged to the genus *Malus*.
189 The retrieved aDNA ITS1 sequences were identical to each other and were shared with wild *M.*

190 *sylvestris* and *M. sieversii* as well as with apple cultivars. Recent DNA studies also reported similar
191 low power to discriminate among species [12,31,37]. This genome region was therefore not
192 informative to discriminate among apple species, in particular between *M. domestica* and its close
193 relatives *M. sieversii* and *M. sylvestris*. However, the study by Schlumbaum *et al.* [28]
194 demonstrated that it is feasible to use waterlogged remains to at least identify the taxonomic status
195 of seed remains as belonging to genus *Malus* vs. *Pyrus*.

196 More recently, Fietta *et al.* [38] used aDNA techniques to characterize 23 ancient apple
197 seeds from Europe, covering a large time scale and geographical range: ten archeological sites
198 spanning a period from the Bronze age (3300 years BCE) to the Middle Ages (1492 Anno Domini,
199 “AD” hereafter) collected in locations distributed from England to Estonia. A small target ITS1
200 region (about 80-150 bp) [39] was typed, but only four specimens out of the 23 (i.e. amplification
201 success of 17%) delivered useful data. Among these, two specimens had been retrieved at a Middle
202 Ages archeological site in Tartu (Estonia) and two at a late Roman (AD 300-400) archeological site
203 in London (UK). The two Estonian samples formed a monophyletic group with the available
204 GenBank ITS1 sequences belonging to *M. sylvestris*, *M. domestica*, *Malus pumila* Mill., *Malus*
205 *hupehensis* (Pamp.) Rehder, and *Malus caspiensis* Langenf., with a high similarity (98% *blastn*
206 sequence similarity) with a Russian specimen of *M. sylvestris*. One of the specimens from the
207 London site also showed a close similarity with the Russian *M. sylvestris* sample. Again, ITS
208 markers did not distinguish between cultivated and wild apples, but did allow the taxonomic
209 identification of the genus *Malus* of waterlogged seed remains.

210 These studies demonstrate that morphological and genetic identification results coincide and
211 that morphological criteria for the distinction between *Malus* and *Pyrus* are reliable, at least in well-
212 preserved samples. However, these two studies also highlight that higher-resolution genetic markers
213 are needed to be able to distinguish between wild and cultivated apple. Such markers are also
214 needed to draw comparisons with present-day apple samples using population genomics and genetic
215 approaches.

216

217 *Archeogenomics holds great promise*

218

219 Recently, advances in **archeogenomics** have opened new perspectives for aDNA studies in woody
220 plant perennials and will likely offer new opportunities for the investigation of the origin and
221 domestication of the cultivated apple. Plant aDNA extraction protocols, e.g. [40], and the use of
222 next generation sequencing technologies allow the sequencing and reconstruction of plant genomes
223 (e.g. [5,10]). This is especially true for DNA extracted from ancient apple remains, as it is currently
224 accepted that either waterlogged or desiccated plant materials, including wood, are more suitable

225 for archeogenomics [41,42] than charred plant remains [43]. Many new techniques are introduced
226 in aDNA studies. Sequence capture or target enrichment is particularly elegant and effective [39]
227 because it reduces contamination risks and can target several genomic regions simultaneously.
228 DNA regions both in the nuclear and plastidial genomes, known to be informative when identifying
229 different *Malus* species and cultivars, can be targeted. Such archeogenomics data, combined with
230 the increasing number of genomic resources available for the genus *Malus* (see [24] for a general
231 review), will enable tracing and follow-up of polymorphisms and alleles in functional genes that are
232 known to have played a role in domestication. Examples are genes involved in water and nutrient
233 use, dormancy, flowering time, fruit ripening, color, acidity, or disease resistance [24].

234 In summary, current archeogenomics techniques open perspectives for significant progress in
235 reconstructing the pace and process of apple domestication. The success of this endeavor will
236 depend on strong collaboration between population genomicians, archeobotanists and archeo-
237 genomicians. First they will have to work together to retrieve biological apple remains from key
238 archeological sites such as those located in Central Asia and along the Silk Road, and then to apply
239 state-of-the-art sequencing and population genomics methods to the found remains.

240

241 **Genomic insights into apple history**

242

243 For decades already, population genetics and genomics studies have provided insights into the
244 history of the cultivated apple and its wild relatives, as well as some clues about the genomic basis
245 of domestication. However, genomic evidence has only been generated for a few traits relevant to
246 apple domestication and cultivation (e.g. fruit size and color). Little is known about the role of the
247 crop-wild hybridization that occurred between the cultivated apple and its wild relatives in the
248 adaptation of the cultivated apple to different environments in Eurasia and beyond. Knowledge of
249 the untapped genetic diversity available in wild apple relatives is highly relevant to modern crop
250 improvement [44]; recent advances in genomics hold great promise in this area.

251

252 *Apple diversification and domestication: from population genetics to genomics*

253

254 Population genetics approaches have generated insights into the origins of the cultivated apple, its
255 domestication history and its spread by human populations, and the biogeography of its wild
256 relatives, i.e. *M. sieversii* and *M. sylvestris*. In a recent review we have summarized the available
257 knowledge on these topics based on the use of classical genetic markers such as microsatellites [9].
258 Genetic analyses support a Central Asian origin for cultivated apple, together with a large
259 secondary contribution from the European wild apple (*M. sylvestris*) to the genetic makeup of the

260 crop. These genetic studies revealed that wild apple species display weak spatial population
261 structure, reflecting high levels of gene flow and high levels of introgression from the domesticated
262 apple [9,45]. Also within the cultivated gene pool, weak levels of structure have been detected
263 [25,46–50]. Genetic research based on a few marker loci has thus revealed a major role of
264 hybridization during the domestication of the apple tree.

265 Recently, next-generation DNA sequencing technologies have made it possible to generate
266 annotated genome assemblies of different genotypes and to produce vast amounts of DNA sequence
267 data for a relatively large sample of wild and cultivated apple genotypes, enabling (i) the
268 exploration of the genomic changes that have taken place during domestication and breeding, and
269 (ii) the analysis of wild apple species evolution and diversification. For example, a *de novo*
270 assembly of a doubled-haploid derived from the cultivar Golden Delicious (line GDDH13) using
271 third-generation sequencing technologies has provided a high quality reference genome [51] that
272 complements the previously published genome sequence of Golden Delicious [52]. Interestingly,
273 the exploration of the GDDH13 reference genome revealed a burst of transposable elements whose
274 timing coincided with the uplift of the Tian Shan Mountains. An additional high quality genome
275 assembly of the anther-derived homozygous HFTH1 revealed extensive genomic variation
276 compared to the GDDH13 reference genome, largely explained by the activity of transposable
277 elements [53]. In particular, the authors pinpointed a long terminal repeat retrotransposon insertion
278 (“redTE”) upstream of a gene (MdMYB1) known to be a core transcriptional factor of anthocyanin
279 biosynthesis and associated with red skin color. Repetitive elements are therefore at the core of the
280 cultivated apple genomic architecture, and might have played a central role in the evolution of this
281 genus as demonstrated in other crops such as maize and sunflower [54,55]. Further, analysis of
282 patterns of DNA methylation showed that epigenetic factors may contribute to agronomically
283 relevant aspects such as fruit development [51]. Also recently, Duan et al. [12] used genome re-
284 sequencing approaches to generate a detailed genome variation map of 117 apple accessions from
285 24 species, and comprising 35 samples of cultivated apple (scions and rootstocks), 39 samples from
286 recognized wild contributor species (10 *M. sylvestris* and 29 *M. sieversii*), and 21 from other species
287 neglected in previous studies. This genome-wide dataset has provided insights that complement
288 those generated by studies based on small sets of genetic markers regarding phylogenetic
289 relationships among the cultivated apple and its wild relatives (for an overview see [9]). Genomic
290 regions determining fruit firmness and flavor were detected using a genome-wide association study,
291 and the effect of domestication and selection on these regions was described [12]. All of the
292 gathered evidence supported a model of apple fruit size evolution comprising two major events: one
293 occurring prior to domestication and the other after/during domestication or recent breeding.
294 Similarly, but without the use of genome-wide data, Yao et al. [56] found that a microRNA whose

295 expression is associated with fruit size was fixed in cultivated apple and in its wild progenitors with
296 large fruits. This indicates that selection for fruit size was initiated before domestication. Genomic
297 analyses therefore suggest that apple differs from annual crop models, such as maize [57], where
298 selection for most key agronomic traits likely occurred after domestication [52]. To date, these
299 genomic studies of apple have not integrated an estimate of the timing of selection in the context of
300 a pre-existing or post-domestication adaptation to a new environment (e.g. climate, pathogens,
301 humans). Nor have they tested whether such genomic regions are the result of crop-wild
302 introgressions following apple domestication.

303 Combined genomics and archeogenomics analyses of present-day and ancient samples (e.g.
304 charred apples or apple halves found in archeological sites), respectively, could provide temporal
305 multi-proxy data that would make it possible to decipher the origin and further selection of traits
306 relevant for apple breeding such as dormancy, regulation of flowering, fruit size, ripening, color,
307 acidity or disease resistance (see [24] for a general review).

308

309 *Introgression as an essential player in apple diversification*

310

311 While inference from quantitative genomics and phylogenomics contributes significantly to our
312 understanding of apple evolution, other issues may be better investigated at the population level. In
313 particular, as explained above, crop-wild hybridization is at the core of apple evolution. Such
314 hybridizations followed by successful establishment of wild beneficial alleles into domesticated
315 populations and *vice versa* can occasionally trigger adaptation. For example, adaptive introgression
316 from highland teosintes has contributed to maize highland adaptation [57]. However, to date no
317 studies have been published on the adaptive consequences of crop-wild introgressions in apple and
318 their genomic basis. The development of new population genomic inference methods makes it
319 possible to identify the dynamics of gene flow and positive selection between populations [24,58–
320 61]. In combination with a comprehensive sampling of both local wild populations and of landraces
321 cultivated along the Silk Road, this offers opportunities to explore the timing, genomic architecture,
322 and regimes of selection acting in the introgressed regions in wild and cultivated apple.

323 The role of adaptive introgression driven by **balancing selection** at self-incompatibility loci
324 in favoring crop-wild pollination and introgression during domestication also deserves further
325 investigation using next-generation sequencing [62]. It is widely assumed that natural selection
326 should favor introgression for alleles at genes evolving under multi-allelic balancing selection
327 including self-incompatibility in plants [63]. Apples are characterized by gametophytic self-
328 incompatibility (GSI) which shows two distinct components: one that determines the pistil
329 specificity (S-RNase gene) and another that determines the pollen specificity (SFBB gene), called

330 S-genes. The locus that contains the genes determining GSI specificity is called the S-locus. The
331 studies that investigated the evolution at the GSI system in *Malus* (and other Rosaceae) have non-
332 exhaustively described hundreds of alleles of the S-RNase and SFBB genes obtained by PCR-based
333 genotyping, Sanger sequencing, or DNA gel blotting for a limited set of accessions [64–72].
334 However, the architecture of the S-locus, its variability and the respective selective regimes acting
335 on it across wild and cultivated apple are still little explored. The extremely high nucleotide
336 diversity, structural rearrangements and repetitive sequences are the reason why traditional
337 sequence analysis is impractical [73–75]; however, see Veeckman et al. [76] for methods and tools
338 to reconstruct loci with a high level of complexity in highly heterozygous species using next-
339 generation sequencing and advanced bioinformatics tools. To decipher the role of the S-locus in the
340 adaptive introgression during apple domestication, one must get a detailed view of polymorphism of
341 the whole region from geographically widely distributed wild and cultivated apple samples.
342 Recently, using RNA sequencing, Pratas et al. [77] identified 24 SFBB genes expressed in anthers
343 and determined their gene sequence in nine *M. domestica* cultivars. Whole-genome information has
344 also become publicly available for apples, including the high quality reference genome and the re-
345 sequenced genomes for more than 100 wild and cultivated apple accessions. It is therefore timely to
346 take advantage of these data to explore species-wide polymorphism of the whole S-locus region and
347 its role in the evolution history of apples during domestication. Such investigations would
348 complement archeogenomics data that is difficult to use for the reconstruction of the S-locus, as that
349 would require high quality and long fragment data.

350

351 *Knowledge of apple domestication and diversification underpins future breeding*

352

353 Crop wild relatives (CWRs) are an invaluable resource for apple breeding [78]. However, the wild
354 genetic resources remain largely unexplored [44]. For instance, *M. baccata* (L.) Borkh., the Siberian
355 crab apple, is one of the main genetic resources for breeding programs, particularly in China,
356 because of its excellent resistance to cold stress and apple scab [79]. Likewise, *Malus floribunda*
357 Siebold ex Van Houtte, *M. hupehensis*, *Malus zumi* (Matsum.) Rehder, *Malus × robusta* (CarriŠre)
358 Rehder and *Malus sieboldii* (Regel) Rehder could provide valuable genetic variation for developing
359 elite cultivars and rootstocks that would be better adapted to specific biotic and abiotic stress
360 conditions [80]. The use of CWRs is also particularly relevant when considering the important role
361 played by hybridization and introgression in the domestication process of *M. domestica* [9]. Several
362 disease resistance genes implemented in commercial apple breeding were introgressed from a wild
363 *Malus* species. Well-known examples are the *Rvi6* gene from *M. floribunda*, which confers
364 resistance to apple scab [81] or the *FB_MR5* gene of *M. × robusta*, which confers resistance to fire

365 blight [81,82]. In the *M. sylvestris* - *M. domestica* pair, the lack of strong reproductive barriers or
366 spatial distance between apple orchards and *M. sylvestris* populations [17] has resulted in the spread
367 of alleles from the cultivated gene pool to populations of the wild species [17,18]. This has
368 generated a number of unmanaged wild apple *M. sylvestris* populations in Europe with introgressed
369 crop alleles [17,18]. These trees may have become unintended repositories of alleles that are no
370 longer available in breeding germplasm, or they may represent an untapped reservoir of adaptive
371 alleles. To date, the lack of cost-efficient methods to screen CWR genomes at high resolution to
372 identify genomic regions associated to traits of relevance in breeding has hindered the use of CWRs
373 in apple breeding.

374 Recent technological innovations have generated in-depth knowledge of the genes
375 underlying specific adaptations to relevant biotic and abiotic factors. These innovations hold great
376 promise to further accelerate this discovery process. For instance, next generation sequencing
377 technologies have been used to unravel the genes underlying resistance to pathogens, such as blue
378 mold (*Penicillium expansum*) infection in *M. sieversii* compared to *M. domestica* [83] and to
379 identify QTLs (Quantitative Trait Loci) controlling resistance to this pathogen in *M. sieversii* [84].
380 These studies illustrate the efficient use of genomic tools to understand domestication and selection
381 in apple and to identify useful genetic variation in wild apple species [85,86] that might open new
382 perspectives for breeding.

383 A major bottleneck is the total lack of systematic, range-wide and genome-wide collections
384 for any of the wild apple relatives (*M. sieversii*, *M. sylvestris*, *M. orientalis* Uglitzk, *M.*
385 *niedzwetzkyana* Dieck) and more ancestral species (*M. baccata*, *M. floribunda*, *M. florentina*
386 (Zuccagni) C.K.Schneid.)). Sampling these taxa will offer clear opportunities to: (i) explore the
387 evolutionary history of apple; (ii) underpin the genomics of adaptation in vegetatively propagated
388 perennials; and (iii) tackle challenges of modern apple breeding by characterizing the genomic
389 variation underlying fruit-related traits and apple adaptation to the biotic and abiotic environments it
390 encountered during its spread along the Silk Road. Archeobotany and archeogenomic studies of
391 apple seed remains from different periods of time will also perfectly complement present-day
392 genomics datasets to bring new insights regarding key traits involved in domestication and recent
393 diversification.

394

395 **Concluding remarks**

396

397 The botany, taxonomy and genetics of apples has been a tantalizing research subject for some time
398 [13,31]. Interdisciplinary investigations have allowed (at least partial) elucidation of some of the
399 processes underlying apple domestication and the history of wild apples in Europe, where apple

400 remains have been systematically recovered from archeological deposits since the beginning of the
401 Holocene. These studies have demonstrated that apple is a suitable study model to unravel a process
402 of divergence and adaptation with gene flow [9]. However, major issues remain to be resolved (see
403 Outstanding Questions). Current limitations include potential difficulties in species identification of
404 archeological remains and the lack of archeobotanical samples in and around the area of origin in
405 Central Asia. However, all of the modern technologies can now be applied to address these issues:
406 easier genome retrieval from modern individuals, genomic approaches on ancient waterlogged and
407 desiccated plant remains including wood, and novel morphometric approaches in archeobotany. We
408 therefore anticipate that genomics and archeogenomics will assist in deciphering apple evolution.
409 Retrieval of representative samples from the centers of domestication in Central Asia and centers of
410 diversification in Europe and in the Caucasus remains an important concern. The risk of loss of wild
411 apple populations in their area of origin and beyond due to human encroachment and climate
412 change has become alarming, underscoring the urgency of investigating this iconic crop. These
413 findings should contribute to sustainable management of cultivated apple germplasm for future
414 breeding programs and the conservation of wild apple populations that are often located in
415 developing countries and subject to increasing anthropogenic threats.

416

417 **Acknowledgements**

418

419 We thank the *Institut Diversité Ecologie et Evolution du Vivant* (IDEEV), the laboratoire
420 d'Excellence BASC and the ATIP-Avenir for funding, the archeobotanists who kindly provided data
421 for our database: Maria Lityńska-Zajac, Marianne Høyem Andreasen, Tanja Märkle, Helmut Kroll,
422 Kelly Reed, Julian Wiethold, Anne de Vareilles, Daniele Arobba, Angela Kreuz, Barbara Zach,
423 Mette Marie Hald, Bénédicte Pradat, Núria Rovira, Koen Deforce, Magdalena Moskal-del-Hoyo,
424 Örne Akeret, Tjasa Tolar, Milena Primavera and Tania Zerl. We also thank Miriam Levenson
425 (ILVO) for excellent English Language editing and review.

426

427 **References**

- 428 1 Besnard, G. *et al.* (2017) On the origins and domestication of the olive: a review and perspectives. *Ann. Bot.* DOI:
429 10.1093/aob/mcx145
- 430 2 Castillo, C.C. *et al.* (2016) Rice, Beans and Trade Crops on the Early Maritime Silk Route in Southeast Asia.
431 *Antiquity* 90, 1255–1269
- 432 3 da Fonseca, R.R. *et al.* (2015) The origin and evolution of maize in the Southwestern United States. *Nat. Plants* 1,
433 14003
- 434 4 Jaenicke-Després, V. *et al.* (2003) Early Allelic Selection in Maize as Revealed by Ancient DNA. *Science* 302,
435 1206
- 436 5 Nathan, W. *et al.* (2018) Ancient DNA reveals the timing and persistence of organellar genetic bottlenecks over
437 3,000 years of sunflower domestication and improvement. *Evol. Appl.* 0, 1–16
- 438 6 Salse, J. (2012) In silico archeogenomics unveils modern plant genome organisation, regulation and evolution.
439 *Curr. Opin. Plant Biol.* 15, 122–130

- 440 7 MacHugh, D.E. *et al.* (2017) Taming the past: ancient DNA and the study of animal domestication. *Annu. Rev.*
441 *Anim. Biosci.* 5, 329–351
- 442 8 Orlando, L. *et al.* (2015) Reconstructing ancient genomes and epigenomes. *Nat Rev Genet* 16, 395–408
- 443 9 Cornille, A. *et al.* (2014) The domestication and evolutionary ecology of apples. *Trends Genet.* 30, 57–65
- 444 10 Kistler, L. *et al.* (2018) Multiproxy evidence highlights a complex evolutionary legacy of maize in South
445 America. *Science* 362, 1309
- 446 11 Morgan, J. and Richards, A. (1993) *The Book of Apples*, Brogdale Horticultural Trust, Ebury Press.
- 447 12 Duan, N. *et al.* (2017) Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit
448 enlargement. *Nat. Commun.* 8, 249
- 449 13 Juniper, B.E. and Mabberley, D.J. (2006) *The Story of the Apple*, Imber Press, Inc.
- 450 14 Gardiner, S.E. *et al.* (2007) Fruits and Nuts: Apple. 4 (Kole, C., ed), pp. 1–62, Springer Berlin Heidelberg
- 451 15 Coart, E. *et al.* (2003) Genetic variation in the endangered wild apple (*Malus sylvestris* (L.) Mill.) in Belgium as
452 revealed by amplified fragment length polymorphism and microsatellite markers. *Mol. Ecol.* 12, 845–857
- 453 16 Coart, E. *et al.* (2006) Chloroplast diversity in the genus *Malus*: new insights into the relationship between the
454 European wild apple (*Malus sylvestris* (L.) Mill.) and the domesticated apple (*Malus domestica* Borkh.). *Mol.*
455 *Ecol.* 15, 2171–2182
- 456 17 Cornille, A. *et al.* (2015) Anthropogenic and natural drivers of gene flow in a temperate wild fruit tree: a basis for
457 conservation and breeding programs in apples. *Evol. Appl.* DOI: 10.1111/eva.12250
- 458 18 Feurtey, A. *et al.* (2017) Crop-to-wild gene flow and its fitness consequences for a wild fruit tree: Towards a
459 comprehensive conservation strategy of the wild apple in Europe. *Evol. Appl.* 10, 180–188
- 460 19 Gharghani, A. *et al.* (2009) Genetic identity and relationships of Iranian apple (*Malus domestica* Borkh.) cultivars
461 and landraces, wild *Malus* species and representative old apple cultivars based on simple sequence repeat (SSR)
462 marker analysis. *Genet Resour Crop Ev* 56, 829–842
- 463 20 Richards, C. *et al.* (2009) Genetic diversity and population structure in *Malus sieversii*, a wild progenitor species
464 of domesticated apple. *Tree Genet. Genomes* 5, 339–347
- 465 21 Volk, Gayle M. *et al.* (2014) The vulnerability of US apple (*Malus*) genetic resources. *Genet. Resour. Crop Evol.*
466 DOI: 10.1007/s10722-014-0194-2
- 467 22 Zhang, C. *et al.* (2007) Genetic structure of *Malus sieversii* population from Xinjiang, China, revealed by SSR
468 markers. *J. Genet. Genomics* 34, 947–955
- 469 23 Zhang, Q. *et al.* (2012) Evaluation of Genetic Diversity in Chinese Wild Apple Species Along with Apple
470 Cultivars Using SSR Markers. *Plant Mol. Biol. Report.* 30, 539–546
- 471 24 Peace, C.P. *et al.* (2019) Apple whole genome sequences: recent advances and new prospects. *Hortic. Res.* 6, 59
- 472 25 Cornille, A. *et al.* (2012) New insight into the history of domesticated apple: secondary contribution of the
473 European wild apple to the genome of cultivated varieties. *PLoS Genet* 8, e1002703
- 474 26 Fuller, D.Q. (2018) Long and attenuated: comparative trends in the domestication of tree fruits. *Veg. Hist.*
475 *Archaeobotany* 27, 165–176
- 476 27 Spengler, R.N. *et al.* (2018) Arboreal crops on the medieval Silk Road: Archaeobotanical studies at Tashbulak.
477 *PLOS ONE* 13, e0201409
- 478 28 Schlumbaum, A. *et al.* (2012) Towards the onset of fruit tree growing north of the Alps: Ancient DNA from
479 waterlogged apple (*Malus* sp.) seed fragments. *Ann. Anat. - Anat. Anz.* 194, 157–162
- 480 29 Miller, N.F. (1999) Agricultural development in western Central Asia in the Chalcolithic and Bronze Ages. *Veg.*
481 *Hist. Archaeobotany* 8, 13–19
- 482 30 Spengler, R.N. (2019) *Fruit from the Sands: The Silk Road Origins of the Foods We Eat*, University of California
483 Press.
- 484 31 Harris, S.A. *et al.* (2002) Genetic clues to the origin of the apple. *Trends Genet* 18, 426–430
- 485 32 Jacomet, S. *et al.* (1989) *Archäobotanik am Zürichsee. Ackerbau, Sammelwirtschaft und Umwelt von neolithischen*
486 *und bronzzeitlichen Seeufersiedlungen im Raum Zürich. Ergebnisse von Untersuchungen pflanzlicher*
487 *Makroreste der Jahre 1979-1988.*, 67
- 488 33 Schweingruber, F.H. (1979) Wildäpfel und prähistorische Äpfel. *Archaeo-Physika*
- 489 34 Knörzner, K.-H. (2000) 3000 years of agriculture in a valley of the High Himalayas. *Veg. Hist. Archaeobotany* 9,
490 219–222
- 491 35 Ucchesu, M. *et al.* (2017) First finds of *Prunus domestica* L. in Italy from the Phoenician and Punic periods (6th–
492 2nd centuries bc). *Veg. Hist. Archaeobotany* 26, 539–549
- 493 36 Wales, N. *et al.* (2016) The limits and potential of paleogenomic techniques for reconstructing grapevine
494 domestication. *J. Archaeol. Sci.* 72, 57–70
- 495 37 Robinson, J.P. *et al.* (2001) Taxonomy of the genus *Malus* Mill. (Rosaceae) with emphasis on the cultivated
496 apple, *Malus domestica* Borkh. *Plant Syst Evol* 226, 35–58
- 497 38 Fietta, A. *et al.* (2014) , EvolvApple: studying apple domestication through the genome scan of ancient seeds. , in
498 *Settore AGR/07 - Genetica Agraria*, Basel (CH)
- 499 39 Savelyeva, E.N. *et al.* (2013) Analysis of sequences of ITS1 internal transcribed spacer and 5.8S ribosome gene
500 of *Malus* species. *Russ. J. Genet.* 49, 1175–1182
- 501 40 Wales, N. and Kistler, L. (2019) Extraction of Ancient DNA from Plant Remains. In *Ancient DNA: Methods and*
502 *Protocols* (Shapiro, B. *et al.*, eds), pp. 45–55, Springer New York

- 503 41 Brown, T.A. *et al.* (2015) Recent advances in ancient DNA research and their implications for archaeobotany.
504 *Veg. Hist. Archaeobotany* 24, 207–214
- 505 42 Wagner, S. *et al.* (2018) High-Throughput DNA sequencing of ancient wood. *Mol. Ecol.* 27, 1138–1154
- 506 43 Nistelberger, H.M. *et al.* (2016) The efficacy of high-throughput sequencing and target enrichment on charred
507 archaeobotanical remains. *Sci. Rep.* 6, 37347
- 508 44 Zhang, H. *et al.* (2017) Back into the wild—Apply untapped genetic diversity of wild relatives for crop
509 improvement. *Evol. Appl.* 10, 5–24
- 510 45 Amirchakhmaghi, N. *et al.* (2018) First insight into genetic diversity and population structure of the Caucasian
511 wild apple in the Hyrcanian forest (Iran) and its resistance to apple scab and powdery mildew. *Genet. Resour.*
512 *Crop Evol.* in press,
- 513 46 Lassois, L. *et al.* (2016) Genetic Diversity, Population Structure, Parentage Analysis, and Construction of Core
514 Collections in the French Apple Germplasm Based on SSR Markers. *Plant Mol. Biol. Report.* 34, 827–844
- 515 47 Omasheva, M.Y. *et al.* (2017) To what extent do wild apples in Kazakhstan retain their genetic integrity? *Tree*
516 *Genet. Genomes* 13, 52
- 517 48 Pereira-Lorenzo, S. *et al.* (2017) Analysis of the genetic diversity and structure of the Spanish apple genetic
518 resources suggests the existence of an Iberian genepool. *Ann. Appl. Biol.* 171, 424–440
- 519 49 Urrestarazu, J. *et al.* (2016) Analysis of the genetic diversity and structure across a wide range of germplasm
520 reveals prominent gene flow in apple at the European level. *BMC Plant Biol.* 16, 1–20
- 521 50 Vanderzande, S. *et al.* (2017) Genetic diversity, population structure, and linkage disequilibrium of elite and local
522 apple accessions from Belgium using the IRSC array. *Tree Genet. Genomes* 13, 125
- 523 51 Daccord, N. *et al.* (2017) , The doubled haploid ‘Golden Delicious’ apple reference genome and epigenome. ,
524 *Nat. Genet.*, 1099–1106
- 525 52 Velasco, R. *et al.* (2010) The genome of the domesticated apple (*Malus x domestica* Borkh.). *Nat Genet* 42,
- 526 53 Zhang, L. *et al.* (2019) A high-quality apple genome assembly reveals the association of a retrotransposon and
527 red fruit colour. *Nat. Commun.* 10, 1494
- 528 54 Badouin, H. *et al.* (2017) The sunflower genome provides insights into oil metabolism, flowering and Asterid
529 evolution. *Nature* 546, 148
- 530 55 Diez, C.M. *et al.* (2014) Three Groups of Transposable Elements with Contrasting Copy Number Dynamics and
531 Host Responses in the Maize (*Zea mays* ssp. *mays*) Genome. *PLOS Genet.* 10, e1004298
- 532 56 Yao, J.-L. *et al.* (2015) , A microRNA allele that emerged prior to apple domestication may underlie fruit size
533 evolution. , *The Plant Journal*, 417–427
- 534 57 Brandenburg, J.-T. *et al.* (2017) Independent introductions and admixtures have contributed to adaptation of
535 European maize and its American counterparts. *PLOS Genet.* 13, 1–30
- 536 58 Racimo, F. *et al.* (2017) Signatures of Archaic Adaptive Introgression in Present-Day Human Populations. *Mol.*
537 *Biol. Evol.* 34, 296–317
- 538 59 Skov, L. *et al.* (2018) Detecting archaic introgression using an unadmixed outgroup. *PLOS Genet.* 14, 1–15
- 539 60 Roux, C. *et al.* (2016) Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic
540 Divergence. *PLOS Biol.* 14, e2000234
- 541 61 Medina, P. *et al.* (2018) Estimating the Timing of Multiple Admixture Pulses During Local Ancestry Inference.
542 *Genetics* 210, 1089–1107
- 543 62 Burgarella, C. *et al.* (2018) Adaptive introgression: an untapped evolutionary mechanism for crop adaptation.
544 *bioRxiv* DOI: 10.1101/379966
- 545 63 Rieseberg, L.H. *et al.* (1995) Chromosomal and genic barriers to introgression in *Helianthus*. *Genetics* 141,
546 1163–1171
- 547 64 Minamikawa, M. *et al.* (2010) Apple S locus region represents a large cluster of related, polymorphic and pollen-
548 specific F-box genes. *Plant Mol. Biol.* 74, 143–154
- 549 65 Okada, K. *et al.* (2013) Isolation and characterization of multiple F-box genes linked to the S 9 - and S 10 -RNase
550 in apple (*Malus × domestica* Borkh.). *Plant Reprod.* 26, 101–111
- 551 66 Aguiar, B. *et al.* (2015) Convergent Evolution at the Gametophytic Self-Incompatibility System in *Malus* and
552 *Prunus*. *PLOS ONE* 10, e0126138
- 553 67 Ashkani, J. and Rees, D.J.G. (2016) A Comprehensive Study of Molecular Evolution at the Self-Incompatibility
554 Locus of Rosaceae. *J. Mol. Evol.* 82, 128–145
- 555 68 Larsen, B. *et al.* (2016) A high-throughput method for genotyping S-RNase alleles in apple. *Mol. Breed. New*
556 *Strateg. Plant Improv.* 36, 24–24
- 557 69 Dreesen, R.S.G. *et al.* (2010) Analysis of *Malus* S-RNase gene diversity based on a comparative study of old and
558 modern apple cultivars and European wild apple. *Mol. Breed.* 26, 693–709
- 559 70 Vieira, J. *et al.* (2010) Evolutionary patterns at the RNase based gametophytic self - incompatibility system in
560 two divergent Rosaceae groups (Maloideae and *Prunus*). *BMC Evol. Biol.* 10, 200
- 561 71 Ma, X. *et al.* (2017) Identification, genealogical structure and population genetics of S-alleles in *Malus sieversii*,
562 the wild ancestor of domesticated apple. *Heredity* 119, 185
- 563 72 De Franceschi, P. *et al.* (2018) Characterization of 25 full-length S-RNase alleles, including flanking regions,
564 from a pool of resequenced apple cultivars. *Plant Mol. Biol.* 97, 279–296

- 565 73 Goubet, P.M. *et al.* (2012) Contrasted Patterns of Molecular Evolution in Dominant and Recessive Self-
566 Incompatibility Haplotypes in *Arabidopsis*. *PLOS Genet.* 8, 1–13
- 567 74 Tsuchimatsu, T. *et al.* (2017) Patterns of Polymorphism at the Self-Incompatibility Locus in 1,083 *Arabidopsis*
568 *thaliana* Genomes. *Mol. Biol. Evol.* 34, 1878–1889
- 569 75 Guo, Y.-L. *et al.* (2011) Evolution of the S-Locus Region in *Arabidopsis* Relatives. *Plant Physiol.* 157, 937–946
- 570 76 Veeckman, E. *et al.* (2018) Overcoming challenges in variant calling: exploring sequence diversity in candidate
571 genes for plant development in perennial ryegrass (*Lolium perenne*). *DNA Res.* DOI: 10.1093/dnares/dsy033
- 572 77 Pratas, M.I. *et al.* (2018) Inferences on specificity recognition at the *Malus*×*domestica* gametophytic self-
573 incompatibility system. *Sci. Rep.* 8, 1717
- 574 78 Migicovsky, Z. and Myles, S. (2017) Exploiting Wild Relatives for Genomics-assisted Breeding of Perennial
575 Crops. *Front. Plant Sci.* 8, 460
- 576 79 Bus, V.G.M. *et al.* (2011) Revision of the Nomenclature of the Differential Host-Pathogen Interactions of
577 *Venturia inaequalis* and *Malus*. *Annu. Rev. Phytopathol.* 49, 391–413
- 578 80 Wang, M.-R. *et al.* (2018) Cryobiotechnology of apple (*Malus* spp.): development, progress and future prospects.
579 *Plant Cell Rep.* DOI: 10.1007/s00299-018-2249-x
- 580 81 Hough, L. (1944) A survey of the scab resistance of the foliage on seedlings in selected apple progenies. *Proc*
581 *Amer Soc Hort Sci*
- 582 82 Peil, A. *et al.* (2007) Strong evidence for a fire blight resistance gene of *Malus robusta* located on linkage group
583 3. *Plant Breed.* 126, 470–475
- 584 83 Ballester, A.-R. *et al.* (2017) Transcriptomic Response of Resistant (PI613981–*Malus sieversii*) and Susceptible
585 (“Royal Gala”) Genotypes of Apple to Blue Mold (*Penicillium expansum*) Infection. *Front. Plant Sci.* 8, 1981
- 586 84 Norelli, J.L. *et al.* (2017) Genotyping-by-sequencing markers facilitate the identification of quantitative trait loci
587 controlling resistance to *Penicillium expansum* in *Malus sieversii*. *PLOS ONE* 12, e0172949
- 588 85 Malnoy, M. *et al.* (2016) DNA-Free Genetically Edited Grapevine and Apple Protoplast Using CRISPR/Cas9
589 Ribonucleoproteins. *Front. Plant Sci.* 7, 1904
- 590 86 Voytas, D.F. and Gao, C. (2014) Precision Genome Engineering and Agriculture: Opportunities and Regulatory
591 Challenges. *PLOS Biol.* 12, 1–6
- 592 87 Charrier, G. *et al.* (2015) Effects of environmental factors and management practices on microclimate, winter
593 physiology, and frost resistance in trees. *Front. Plant Sci.* 6, 259
- 594 88 Charrier, G. *et al.* (2017) Monitoring of freezing dynamics in trees: a simple phase shift causes complexity. *Plant*
595 *Physiol.* DOI: 10.1104/pp.16.01815
- 596 89 Hartmann, H.T. *et al.* (1990) *Plant Propagation: Principles and Practices*, Pearson.
- 597 90 De Franceschi, P. *et al.* (2012) Molecular bases and evolutionary dynamics of self-incompatibility in the Pyrinae
598 (Rosaceae). *J. Exp. Bot.* 63, 4015–4032
- 599 91 Gaut, B.S. *et al.* (2015) Genomics and the Contrasting Dynamics of Annual and Perennial Domestication. *Trends*
600 *Genet.* 31, 709–719
- 601 92 Zohary, D. and Spiegel-Roy, P. (1975) Beginnings of fruit growing in the Old World. *Science*
- 602 93 Rom, C.R. *et al.* (1990) Size controlling apple rootstocks affect growth, spur quality, foliar nutrition and
603 productivity. *Compact Fruit Tree* 23, 17–21
- 604 94 Jackson, J.E. (2003) *The Biology of Apples and Pears*, Cambridge.
- 605 95 Autio, W. *et al.* (2011) , Performance of several semi-dwarfing rootstocks with “Fuji” and “McIntosh” as Scion
606 cultivars in the 1999 NC-140 semi-dwarf apple rootstock trials. , in *IX International Symposium on Integrating*
607 *Canopy, Rootstock and Environmental Physiology in Orchard Systems*, Belgium, pp. 327–334
- 608 96 Ferree, D.C. and Warrington, I.J. (2003) *Apples: botany, production, and uses*,
- 609 97 Preston, A.P. (1966) Apple Rootstock Studies: Fifteen Years’ Results with Malling-Merton Clones. *J. Hortic. Sci.*
610 41, 349–360
- 611 98 Pearsall, D. (2015) *Paleoethnobotany, Third Edition*. New York: Routledge., Third edition. New York: Routledge.
- 612 99 Brozio, J.P. *et al.* (2014) A Middle Neolithic well from Northern Germany: a precise source to reconstruct water
613 supply management, subsistence economy, and deposition practices. *World Reshaped Pract. Impacts Early*
614 *Agrar. Soc.* 51, 135–153
- 615 100 Kirleis, W. (2018) *The cultural significance of plants: Past Landscapes: The Dynamics of Interaction Between*
616 *Society, Landscape, and Culture*, Haug *et al.* (eds.). Sidestone Press, Leiden.
- 617 101 Hoban, S. *et al.* (2016) Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future
618 Directions. *Am. Nat.* 188, 379–397
- 619 102 Antolin, F. *et al.* (2017) *Archäobotanik (Makroreste): Ergebnisse der Untersuchungen der Flächenproben im*
620 *Hinblick auf Taphonomie, Ackerbau, Sammelwirtschaft und Umwelt*, Bleicher, N., Harb, C. Band 3:
621 Naturwissenschaftliche Untersuchungen, Zürich und Egg.
- 622 103 Bieniek, A. and Lityńska-Zajac, M. (2001) New finds of *Malus sylvestris* Mill. (wild apple) from Neolithic sites
623 in Poland. *Veg. Hist. Archaeobotany* 10, 105–106
- 624 104 Martin, A.C. and Barkley, W.D. (1961) *Seed identification manual*, University of California Press.
- 625
- 626

627 **Box 1. On the cultivation of apples, and its past and present use worldwide**

628 The commercial apple (*Malus domestica* Borkh.), sometimes referred to as the “Grand King of Fruit
629 Crops”, is a deciduous species belonging to the Rosaceae family. It is cultivated mostly in
630 temperate zones and consumed either fresh or in food products worldwide. Thanks to its
631 (eco)physiological properties and grafting techniques, the species can survive in cold environments
632 by reducing the risk of damage during autumn and spring frost events [87,88].

633 Known for its fruit qualities and characteristics, it has been cultivated since ancient times
634 [13]. The so-called Silk Road, a network of ancient trade routes from Asia to Europe in existence
635 from as 3500 BCE to the mid-1400s, promoted the spread of apples into the European continent.
636 The Greek philosopher and botanist Theophrastus (ca. 320 BCE) studied the apples brought to
637 Greece by Alexander the Great and described six apple cultivars and cultural practices such as
638 general tree care and grafting. It is believed that the Romans introduced some of today’s apple
639 cultivars to the British Isles. Canadian settlers introduced this fruit to North America, with the first
640 commercial apple orchard in the U.S.A. established ca. 1600 [13].

641 Apples are now consumed worldwide as either fresh fruit or as the basis for various food
642 products (e.g. dried, juice, cider, purees, etc.). Over 89 million tons of apples are produced annually
643 worldwide, making this fruit the third most popular in the world
644 (<http://www.fao.org/faostat/en/#data/QC>). The largest producer is China, representing around 40%
645 of the world’s total production, followed by the United States, with over 4.5 million tons per year.
646 In Europe, Italy, France, Germany and Poland are the biggest apple producers
647 (www.yara.us/agriculture/crops/apple/key-facts/world-apple-production/). Product quality is
648 ensured by intensive cultural and pest management practices as well as harvesting at proper
649 maturity, storage, and shipping under highly managed conditions.

650

651 **Box 2. The rootstock and the scion: Key components of large-scale apple cultivation**

652 To obtain the uniformity required for profitable production, the commercial apple tree is a
653 composite biological unit: a combination of rootstock, scion, and sometimes an interstem. Such
654 rootstock/scion/interstem combination is created by either **grafting** or budding [89]. These asexual
655 propagation practices are essential for commercial production, because apple seedlings derived
656 from a single tree are genetically and phenotypically highly heterogeneous due to their forced cross-
657 pollination due to self-incompatibility [77,90]. The cultivated apple tree is therefore a complex
658 biological entity that requires a vast array of knowledge from scientists, consultants and growers to
659 ensure a high quality product for the market.

660 Invented in East Asia about 4,000 years ago, the grafting of scion onto a rootstock created a
661 revolution for apple spread and cultivation [9,91,92] The use of dwarfing rootstocks has facilitated

662 the establishment of high-density orchards (2,000 to 3,500 trees/hectare). It has also been shown
663 that more dwarfing rootstocks induce the scion cultivar to flower and bear fruit earlier than less
664 dwarfing rootstocks [93–95]. Hypotheses of the biological mechanisms involved in scion size
665 control and precocity include hormone translocation and changes in fluxes of nutrients, water, and
666 assimilates between rootstock and scion [94]. Higher light-use efficiency in high-density orchards
667 can have a positive effect on fruit size and color as well as on fruit production. In general, the more
668 dwarfing the rootstocks, the better they perform in high-density planting, generating earlier
669 economic returns [96].

670 The interaction between the rootstock and the scion is complex. For example, the most
671 commonly used M.9 rootstock (dwarfing) has the highest ratio of crop to scion when compared to
672 more vigorous rootstocks [97]. However, choice of rootstock can also influence productivity [93]
673 and other factors such as precocity, yield, environmental and edaphic adaptability, light
674 interception, and disease and pest resistance [97,98]. Scion selection is an existential question when
675 selecting for desired fruit qualities and tree characteristics. Important traits influenced by scion
676 include fruit quality (size, shape, color, sweetness, tartness, firmness, flesh color, etc.), yield,
677 disease resistance, time of harvest, and storage potential [93,97].

678 Worldwide there are an estimated 6,000 to 10,000 scion cultivars [49]. Frequently, new
679 cultivars arise as spontaneous mutations, often called ‘bud sports’ (i.e. ‘Red Delicious’), but several
680 private and public breeding programs also create new apple cultivars via crossing and selection.

681

682 **Box 3. Some archeobotanical evidence of apple cultivation**

683 A first glimpse into apple use and cultivation in Europe is coming into view. We have compiled
684 archeobotanical data from about 600 sites located in 34 countries to give an overview of the use of
685 apple in prehistoric and historic times (Figure I). The compilation is based on findings of apple in
686 Eurasia, identified as *Malus* or *Malus/Pyrus* (or Maloideae). In these studies, the small plant
687 remains from soil samples are retrieved using **water flotation** techniques [98].

688 Most remains of *M. domestica* identified are from the Roman period (ca. 2,100-1,500 BCE)
689 onwards (Figure I). The increase observed since the Roman period could be a result of the spread of
690 the cultivation of cultivated apple in central and northern Europe. The important representation of
691 wild apple remains during the Neolithic period can be interpreted as an evidence of the efforts made
692 to manage the landscapes around the sites in order to increase the productivity of certain trees such
693 as hazel, apple or oak, and to protect them from domestic animals (e.g. [99]). *Malus sylvestris* needs
694 a lot of light, thus it was unlikely to be common in prehistoric forests in Europe; instead it would
695 have been found on floodplains, hedges and thickets [100]. One good example for this is the Zürich-
696 Parkhaus Opéra site (dated to ca. 5,100 BCE), where the excellent preservation conditions made it

697 possible to calculate that between 190 and 400 kg of apples were gathered yearly per domestic unit
698 in average over a period of 15 years [101]. These amounts clearly exceed the normal productivity of
699 a natural woodland environment in the Lake Zürich region. For some of the northernmost regions of
700 Europe it has been proposed that early farmers could have spread wild apples [100], which would
701 also indicate a high degree of management of *M. sylvestris*. The land use system changed
702 significantly afterwards, with larger areas of cultivated land and less access to woodland. The
703 consumption of wild apples also appears to have decreased (note, however, that this observation
704 may be biased due to the state of research and **taphonomy** issues). A significant increase of apple
705 consumption probably only started with domesticated apple cultivation (see also [100]).
706

707 **Figure captions**

708

709 **Figure I. Presence of apple remains recorded from archeobotanical studies a. in different**
710 **types of preservation conditions, b. for different regions, c. for different chronological periods.**

711 Similar temporal trends are shown in the different regions, with a high number of records for the
712 Neolithic period (ca. 8,000-4,500 BCE), a decrease in the Bronze and Iron ages (ca. 4,500-2,100
713 BCE), followed by an increase in Roman times and a generally good record of medieval and post-
714 medieval finds.

715

716 **Figure 1. Photography and sizes of seed remains of apples.** From left to right and from top to
717 bottom: Seed of *Malus sylvestris* (Zürich-Parkhaus Opéra), pericarp of *Malus/Pyrus* (Zürich-
718 Parkhaus Opéra, form (Photo credits: G. Haldimann) [102]; charred crabapple (Photo credits: W.
719 Kirleis); charred apple half (Photo credits: W. Kirleis). Fruit size is used as a criterion to distinguish
720 between wild and domesticated apple in Europe. Charring has an important effect on fruit size, as
721 fruits can become ca. 1/3 smaller after charring [33]. Before charring, wild apples were 20-36 mm
722 wide and 18-36 mm long (weight 5-16 g). After charring, wild apples measure 15-25 mm wide and
723 15-26 mm long. Most charred archeological remains of wild apples fit within these dimensions and
724 the results agreed with an average of 27 mm in diameter for charred wild apple [103]. Within this
725 genus, the differentiation of seed morphology between species has been little explored, but the
726 morphological differentiation of seeds between *Malus* sp. and *Pyrus* sp. is relatively well defined in
727 the archeobotanical literature. Seeds of apple are wider and rounded in the apex, in contrast to those
728 of pear. They also have fibrous lengthwise striations on their surface (e.g. [104]) that make them
729 well recognizable in both charred and uncharred states (Figure I). Pericarp fragments can be found
730 in archeological sites as well. In addition, the inner wall of the endocarp (also called pericarp or
731 carpel in the literature) of apple seeds is shiny, and presents undulating fibers on the inner side
732 (Figure I). This allows identification of apple pericarp fragments found in archeological sites.

733

734

735