UNIVERSITY OF COPENHAGEN

Hybridization and genetic variation in Danish populations of European crab apple (Malus sylvestris)

Larsen, Anders Søndergaard; Asmussen, Conny Bruun; Coart, Els; Olrik, Ditte Christina; Kjær, Erik Dahl

Title: Hybridization and genetic variation in Danish populations of European crab apple (*Malus sylvestris*).

Authors: Anders S. Larsen[a]*, Conny B. Asmussen[b], Els Coart[c], Ditte C. Olrik[a], Erik D. Kjær[a]

[a] Department for Management of Forest Genetic Resources, Danish Centre for Forest, Landscape and Planning, The Royal Veterinary and Agricultural University, Hørsholm Kongevej 11, 2970 Hørsholm, Denmark

[b] The Botanical Section, Department of Ecology, The Royal Veterinary and Agricultural University, Rolighedsvej 21, 1958 Frederiksberg C, Denmark.

[c] Department of Plant Genetics and Breeding, Agricultural Research Centre, Caritasstraat 21, 9090 Melle, Belgium.

*Corresponding author: Tel.: +45 35281653;

Fax: +45 35281517; e-mail address: ansl@kvl.dk (A.S. Larsen)

Abstract

The aim of the present study was to investigate the genetic variation in Danish populations of the endangered European crab apple (*Malus sylvestris*). Special emphasis was given to hybridization between the wild species and its cultivated relative *Malus ×domestica.* A total of 178 wild individuals from four Danish populations were studied along with a reference sample of 29 old cultivars. The genetic variation within and among samples was studied at ten microsatellite marker loci. Additionally, a morphological analysis was carried out to identify hybrids and test for correspondence between phenotypic and genotypic indices of hybridization.

From application of ordination and a model based cluster analysis to the molecular data, two clusters were identified consisting of wild and cultivated individuals respectively. This indicates that pronounced admixture between the two species is not present. At the population level, a high correspondence was found between geographic isolation from *M. ×domestica* and genotypic and morphological indices of hybridization. As expected, isolated populations appeared less affected by hybridization than poorly isolated populations. Isolated 'pure' *M. sylvestris* populations could thus be identified. However, morphological and molecular evidence of hybridization was found to be divergent at the individual level. This is suggestive of some historical introgression into the *M. sylvestris* gene pool, and indicates that relying exclusively on either morphological or molecular characters as diagnostic markers in studies of hybridization between *M. ×domestica* and *M. sylvestris* might lead to fallible results. Combined application of genetic and morphological markers is therefore recommended.

1. Introduction

The European crab apple (*Malus sylvestris* (L.) Mill.) is a small tree of the family Rosaceae, indigenous to the European continent exclusive of the southwestern part and northern Scandinavia. *Malus sylvestris* has become an endangered species in Europe, the main threats being identified as fragmentation and reduction of its habitat as well as hybridization with the domesticated apple (*Malus ×domestica* Borkh.) (Kleinschmit *et al.*, 2000). Both hybridization and habitat modifications are of general importance because they are expected to reduce the long-term viability of the affected species. Hybridization between wild species and their domesticated relatives are likely to reduce the fitness of the wild populations in question and might even lead to their extinction (Lynch, 1991; Rhymer and Simberloff, 1996). Fragmentation and reduction in habitat size can lead to the same results through increased genetic drift and inbreeding depression (National Research Council, 1991). In Denmark a survey of *M. sylvestris* has found a scattered distribution of suggested pure populations based on morphological traits (Ødum, 1968).

In the light of these threats, great efforts are being put into *in situ* and *ex situ* conservation programs for *M. sylvestris* in Denmark and throughout Europe (Graudal *et al.*, 1995; Kleinschmit *et al.*, 2000). It has been suggested that the major aim of such programs should be to create conditions that enhance future evolution of the genetic resource (FAO, 1989; Eriksson *et al.*, 1995). One such condition would be identification and inclusion, in conservation programs, of the greatest possible amount of genetic variation as a buffer towards future changes in the environment of the species (Ledig, 1991; Eriksson *et al.*, 1995). Also, it is important to create undisturbed conditions for effective, natural selection in order to maintain adaptation to local

conditions over time (Kjær *et al*., 2004). Research to unravel the degree of hybridization, intraspecific variation, and inbreeding depression in wild populations of *M. sylvestris* is therefore needed. Such research is further justified by the close relationship between *M. sylvestris* and *M. ×domestica*, which opens up the opportunity of transferring valuable genes from the wild to the cultivated species in commercial breeding programs (Ponomarenko, 1985).

      Malus sylvestris is an outcrossing species (Campbell *et al.*, 1991) with hermaphrodite flowers predominantly pollinated by bees and flies (Syrphidae) (Kugler, 1970). Under Danish conditions *M. sylvestris* is most commonly found in semi-natural grasslands and woodlands with a prehistory of grazing. This is a consequence of the great light-demand of this pioneer species along with the success of large domestic herbivores as seed dispersal vectors. A wide array of extant wild animals is feeding on the fruit under Danish conditions but their success as seed dispersal vectors have been questioned (Buttenschøn and Buttenschøn, 1998).

      Earlier studies of hybridization between the two related species have primarily focused on morphological characters (e.g. Remmy and Gruber, 1993). Wagner (1996) presents a list of 45 morphological characters for differentiation between the two species along with suggestions of characteristic states for each character. Of the wide range of characters suggested, Wagner (1998a) finds that fruit diameter, diameter of long shoots and pubescence (in the order listed) has the greatest power in discriminating between the two species and their hybrids. However, for practical studies the most prominent character analyzed appear to be pubescence of abaxial leaf surfaces (eg. Wagner, 1998b; Coart *et al.*, 2003). Thus, contrary to flowers and fruits, leaves can be

studied for an extensive period of time. In addition, flowers and fruits are often absent on trees growing in dense forests.

Morphologically based studies have found great phenotypical variability with numerous occurrences of phenotypes intermediate between the two hybridizing species *M. sylvestris* and *M. ×domestica*. This variation has led to the hypothesis that the range of closely related Eurasian species are actually one panmictic species (Kleinschmit *et al.*, 2000). The hypothesis is plausible due to the weak genetic barriers to interspecific and intergeneric hybridization within the subfamily Maloideae in Rosaceae (Korban, 1986). Recent research using microsatellite and amplified fragment length polymorphism (AFLP) markers do, however, question this hypothesis by showing only limited admixture between Belgian and German populations of *M. sylvestris* and the close relative *M. ×domestica* (Coart *et al.*, 2003). Isozyme analyses of German *M. sylvestris* populations adds further support to the hypothesis that *M. ×domestica* and *M. sylvestris* still make up two separate gene pools (Wagner *et al.*, 2004).

The objectives of the present study were: (1) to investigate the degree of admixture in Denmark between the genomes of *M. ×domestica* and *M. sylvestris* in the wild, (2) to study the level of genetic differentiation between wild Danish populations of putative *M. sylvestris* and, (3) to investigate the relationship between molecular and morphological markers. Our general hypothesis was that phenotypic variability hampers the exact delineation of *M. sylvestris* from its hybrids with *M. ×domestica* based on morphology, and genetic studies may therefore improve conservation efforts substantially by allowing identification of first generation hybrids. Such studies are also likely to shed light upon important issues such as phenotypic variation and plasticity,

common ancestry, and convergent evolution that could be alternatives to hybridization and introgression events in explaining the observed variation in morphological characters (Linder *et al.*, 1998; Jarvis and Hodgkin, 1999). Understanding these issues will improve the possibility of obtaining a heterogenous breeding material for landscape planting activities as well as the ongoing conservation programs concerning this species.

For the present study, microsatellite markers were chosen due to their applicability in earlier studies of hybridization (Muir *et al.*, 2000; Beaumont *et al.*, 2001; Coart *et al.*, 2003), codominance, and ease of interpretation in studies of population differentiation (Balloux and Lugon-Moulin, 2002), high level of diversity, and reproducibility allowing for future comparisons of European data (Jones *et al.*, 1997). Potentially, homoplasy can pose a problem in delineating the relationship between taxonomic units (Goldstein & Pollock, 1997). Nevertheless, concordance between results obtained using AFLP and microsatellite markers in the aforementioned study of Coart *et al.* (2003) suggests that this is not an issue for the present study. Morphological characters were included in the study 1) as a criteria to distinguish between the wild and cultivar species and 2) to investigate the relationship between phenotypical and molecular indices of hybridization.

2. Materials and methods

2.1 *Sampling*

Four wild populations of putative *M. sylvestris* were chosen for this study, covering a range of population sizes and supposed range of genetic isolation from *M. ×domestica* (Table 1). Each sampling location was taken to be a population in

reproductive terms covering geographical areas in the range of 0.1 to 1 square kilometer. Populations were separated by 100 to 300 kilometers. Within each population a maximum of 50 individuals were selected by random sampling. That is, the trees were sampled so that the density of sampled individuals reflects the population density. The number of individuals sampled at location Noerbjerg was limited by the population size.

Table 1 Details of studied populations. Name, code, and UTM coordinates of the sampling locations are given along with the number of individuals sampled, *n*, and the estimated population size, *N*. Also, a short description indicating the isolation from *M. ×domestica* is provided.

Location	Code	Coordinates	*n*	*N*	Description of location
Hestehave	HE	56º16/10º27	50	App. 300	Old semi-natural forest with a long history of human influence. Specimens of *M. ×domestica* were found along roads and at picnic sites.
Kaas	KA	56º37/8º41	50	App. 400	Oak scrub and semi-natural grassland on minor peninsula. Within a distance of one kilometer, orchards of *M. ×domestica* have been in place for approx. a century.
Kosteskov	KO	54º46/11º41	50	App. 175	Yong semi-natural forest (100-150 yrs) in an area that has seen intensive commercial fruit growing for the last century.
Noerbjerg	NO	57º11/9º45	28	40	Remote oak scrub approx. 150 yrs of age. Region in general unfavorable of fruit growing.
Cultivars	CU	-	29	-	-

Cultivars were obtained from the Pometum of the Royal Veterinary and Agricultural University, Denmark, with the main emphasis on old cultivars and cultivars locally important in the sampling areas of putative *M. sylvestris*. This strategy was chosen based on its inclusion of the most likely parents of potential hybrids among the wild populations studied. In addition, the cultivar 'Elstar' was included in the molecular study as a standard genotype for comparison of results between laboratories and gels.

2.2. DNA extraction

For DNA analyses leaves were collected from mature trees in the selected populations. Material was collected from the terminal part of long shoots during August 2003. Leaf material collected in the field was dried immediately using silica gel as prescribed by Chase and Hills (1991) and stored at room temperature until DNA extractions were completed.

The following modified version of the CTAB method (Doyle and Doyle, 1987) was used for extraction of DNA. Tissue was treated with liquid nitrogen and ground on a vortex. Subsequently, *c.* 15 mg was lysed at 60ºC in 0.6 mL extraction buffer containing 55mM CTAB, 0.1 M Tris HCl, 20mM EDTA, 1.4 M NaCl, and 0.2% 2-mercaptoethanol (pH 8.0). After 30 min 0.6 mL of chloroform/isoamylalcohol (24:1) was added and the samples were centrifuged for 15 min at 4ºC (2,600 g). The supernatant was digested with 6 μg RNAse A followed by addition of 0.3 mL of 2-propanol and precipitated overnight at –18ºC. The DNA was pelleted by centrifugation at 4ºC for 10 min (12,400 g), washed in 70% ethanol, and resuspended in TE buffer containing 10 mM Tris HCl and 1 mM EDTA (pH 8.0).

DNA quality was tested on 0.8% agarose gels run at 100 mV for 45 min. Subsequently, DNA content was measured using an Eppendorf BioPhotometer, and the samples were diluted to 5 ng/μL.

2.3. Microsatellite data

In the present study, ten primer pairs (listed in Table 4) were selected based on clearness of obtained banding patterns, mapping to different linkage groups, and high degree of polymorphism. Loci were amplified individually in 15 μL reactions

containing 25 ng template DNA using TEMPase (Ampliquon) Hot Start DNA Polymerase and buffer. Reactions were carried out in 1.5 mM MgCl$_2$, 0.2 mM of each dNTP, and 0.2 μM of each of forward and reverse primers using 0.75 U *Taq*. Polymerase chain reaction (PCR) was undertaken in Applied Biosystems GeneAmp PCR System 2700 and 9700 under the following conditions: 95°C for 15 min, followed by 35 cycles of 94°C for 40 s, T_a for 30 s, and 72°C for 30 s. A final extension of 72°C for 10 min was included. A T_a of 50°C was used for primer pairs CH01f02, CH01h10, and CH02c06; 60°C for CH02c06; and 55°C for the remainder.

Forward primers of CH02b07, CH02c06, CH02c11, CH01f02, CH04g04, and NZ23g04 were labeled with WeIIRED fluorescent near infrared dyes (D2-PA, D3-PA, or D4-PA). Multiplex electrophoresis was carried out on a CEQ 2000 Genetic Analysis System and data analyzed using CEQ 2000 software (both Beckman Coulter). Forward primers of CH01h01, CH01h10, CH02b12, and NZ28f04 were labeled with IRD 700 or IRD 800 near infrared dye and multiplex electrophoresed on a Gene Readier 4200 followed by analysis using GENE IMAGIR software (both Li-Cor, Inc.)

2.4. Morphological and phenological data

For the present study a suite of five morphological characters were selected. These were; 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. All characters were described in August 2003. Pubescence characters were described on a scale from 0 to 3 according to Remmy and Gruber (1993):

0: Hairless

1: Sparsely hairy; hairs restricted to the main nerves

2: Moderately hairy; easily recognizable hairs on the entire leaf surface

3: Felted

A score of 0 or 1 is usually accepted for individuals of *M. sylvestris* and a score of 3 is typical for *M. ×domestica* whereas hybrids are reported to be intermediate (i.e. a score of 2) (Remmy and Gruber, 1993). Fruits of *M. sylvestris* are green to yellow and less than 3.5 cm in diameter. Red blotches and fruits greater than 3.5 cm are indicative of hybridization (Wagner, 1996). Fruiting was sparse in the populations studied and fruit characters were only recorded for 75 wild individuals of the 178 sampled. Shoot diameter were included initially because of its suggested high discriminating power (Wagner, 1998a), but was later omitted because it was found difficult to isolate the genetic from the environmental effects influencing this character.

In spring 2005, flowering in populations HE and KA along with nearby orchards of *M. ×domestica* (*c.* 5 and 1 km respectively) was scored on a 0 to 3 scale based on the proportion of flowers open on the observed individual:

0. No open flowers

1. < 10% open

2. 10% - 90% open

3. > 90% open

Open flowers at a given time were counted using a pair of binoculars. The total number of flowers on each tree for estimation of proportion of open blossoms were estimated at peak flowering. Records were made five times over a period of three and a half week (cf. Figure 4). A total of 100 wild individuals and 15 different cultivars were recorded

(including 'Belle de Boscoop', 'Filippa', 'Flaskeæble', 'Ingrid Marie', 'Rød Ananas' and ten cultivars of uncertain identity). These cultivars represent a range from early to late flowering periods (MAAF, 1973).

2.5. Statistical analyses

The average genetic diversity over loci within populations was estimated by calculation of mean number of alleles per locus (A), observed heterozygosity (H_O), gene diversity (H_E) according to Nei (1978), and Wright's (1978) fixation index, F_{IS}, using POPGENE 1.32 (Yeh *et al.*, http://www.ualberta.ca/~fyeh/download.htm).

Tests for departure from Hardy-Weinberg proportions (Guo and Thompson, 1992) and gametic disequilibrium were carried out by permutation of allele frequency tables using the software GENEPOP 3.4 (Raymond and Rousset, 1995) and FSTAT 2.9.3 (Goudet, http://www2.unil.ch/izea/softwares/fstat.html), respectively. *F*-statistics (Weir and Cockerham, 1984) were calculated for the group of wild genotypes along with pairwise estimates of genetic differentiation, θ, among all populations (FSTAT). Differentiation between pairs of populations was further investigated by computation of estimates of Nei's (1978) standard genetic distance corrected for small sample sizes, D_S, (MICROSAT; Eric Minch, http://hpgl.stanford.edu/projects/microsat/). Significance of differentiation between pairs of populations out of the complete dataset was tested by permutation of alleles among populations (FSTAT). Additionally, Michalakis and Excoffier's (1996) estimator, Φ_{ST}, based on the stepwise mutation model was calculated and tested (Hardy *et al.*, 2003) for significant deviation from θ using SPAGEDI 1.2 (Hardy and Vekemans, 2002).

Potential genetic structure of the data set was analyzed by ordination of the genotypes by means of non-metric multidimensional scaling (NMDS) (cf., e.g. Manly (1994)) based on Bowcock *et al.*'s (1994) allele sharing distance, D_{AS} (MICROSAT), between pairs of genotypes. Coordinates from principal coordinate analysis (PCO) were used as an initial configuration. Principal component analysis (PCA) was used to make the axes in the final configuration align with the major axes of variation. Multivariate analyses and illustrations were carried out using NTSYSpc 2.1 (Rohlf, University of New York).

Table 2 Morphological description. Numbers for pubescence are percentages of trees with pubescence scores of 0 / 1 / 2 / 3 respectively. For comparison, the score for *M. ×domestica* is 3 for all characters.

Population	Pubescence (%)			
	L-LS[a]	L-SS[b]	LS[c]	PI[d] (%)
HE	66 / 32 / 2 / 0	66 / 32 / 2 / 0	60 / 24 / 14 / 2	18
KA	90 / 10 / 0 / 0	88 / 12 / 0 / 0	86 / 8 / 6 / 0	6
KO	80 / 18 / 2 / 0	78 / 22 / 0 / 0	92 / 0 / 8 / 0	8
NO	72 / 28 / 0 / 0	66 / 34 / 0 / 0	97 / 3 / 0 / 0	0

[a] L-LS: leaves, long shoots; [b] L-SS: leaves, spur shoots; [c] LS: long shoots; [d] PI: pubescence-based hybrid character index as described in the text

A model based Bayesian clustering approach for probabilistic assignment of individuals to an inferred number of clusters with simultaneous estimation of the unknown allele frequencies within them was applied to the genotypic data set. The approach allows identification of individuals that are assigned jointly to two or more clusters, and thereby suggested to be hybrid candidates. Assignment of individuals and inference of allele frequencies is performed so that departures from Hardy-Weinberg and gametic disequilibrium within inferred clusters are minimized. Estimates, q_k, of the proportional assignment of individuals, to each of the inferred populations, $k = 1…K$ are given. The software STRUCTURE 2.1 (Pritchard *et al.*, 2000), which is based on a Markov Chain Monte Carlo method was used for this purpose.

Another feature of this software is the option of using prior information of sampling origin for the individuals studied. This is useful for identifying individuals of migrant origin (Pritchard *et al.*, 2000). Using this option, the prior probability of an individual belonging to its sampling location is $q = 1 - m$, with m being the probability of an individual or gene being an immigrant.

Table 3 Fruit bearing. For each population the percentage of individuals bearing fruit in the year of sampling. Also shown is the proportion of these fruit bearing individuals with fruits indicative of hybrid origin.

Population	Fruit bearing (%)	Hybrid (%)
HE	62	35
KA	24	17
KO	38	42
NO	39	0

Due to the easy applicability and importance attached to pubescence in studies of hybridization between *M. ×domestica* and *M. sylvestris*, a pubescence index, PI, was constructed. The value 1 was assigned to individuals having a score of 2 or above for any of the pubescence characters studied. All other individuals were assigned the value 0. Fruit characters were treated separately.

Table 4 Description of the loci studied and summary of allelic variation. For each locus are shown the number of individuals scored, n, gene diversity according to Nei (1978), H_E, and the number of alleles observed, A. Mean for gene diversity (95%) and the number of alleles are shown in bold.

Locus	A	Size range	Wild populations			Cultivars		
			n	A	H_E	n	A	H_E
NZ28f04[a]	8	90 – 110	178	7	0.690	29	7	0.774
NZ23g04[a]	14	83 – 116	178	12	0.646	29	12	0.851
CH01f02[b]	17	162 – 219	178	17	0.862	29	12	0.898
CH01h01[b]	20	103 – 148	178	20	0.854	29	7	0.842
CH01h10[b]	18	87 – 140	178	17	0.851	29	9	0.804
CH02b07[c]	14	94 – 131	178	12	0.774	29	8	0.780
CH02b12[b]	18	104 – 149	177	18	0.748	29	11	0.854
CH02c06[b]	26	202 - 288	178	26	0.886	29	13	0.903
CH02c11[c]	15	207 – 239	178	13	0.803	29	11	0.878
CH04g04[c]	17	148 – 203	178	15	0.837	29	13	0.874
				18.8	**0.795±0.057**		**9.8**	**0.846±0.033**

Marker loci from [a]: Guilford *et al.* (1997), [b]: Gianfranceschi *et al.* (1998), [c]: Liebhard *et al.* (2002)

3. Results

3.1. Morphological description

In total, 41 wild individuals showed morphological signs of hybridization in one or more of the studied characters. Omitting fruit characters that were only available for 42% of individuals left 16 individuals to be identified morphologically as hybrids (i.e. by pubescence; PI = 1).

When comparing populations, NO showed the lowest proportion of hybrids as inferred by pubescence and fruit characters and HE or KO the highest depending on the character studied (Table 2 and 3).

When analyzing the data set, it was noted that pubescence of leaves (from long or spur shoots) and pubescence of long shoots were often poorly correlated when

Table 5 Measures of genetic variation within populations. Arithmetic means are given for the collective wild sample. No significant departures from Hardy-Weinberg proportions were observed for individual populations. Estimated population size is given by *N*.

Population	*N*	*A*	H_O	H_E	F_{IS}
HE	300	12.6	0.794	0.815	0.017
KA	400	9.2	0.754	0.746	-0.012
KO	175	10.8	0.731	0.761	0.030
NO	40	8.6	0.796	0.791	-0.029
Wild, mean	**-**	**10.3**	**0.769**	**0.778**	**-0.002**
CU	-	10.2	0.821	0.846	0.014

looking at the same individual. Also, out of 20 individuals bearing fruit indicative of hybrid ancestry, only five could be classified as hybrids from their pubescence scores. On the other hand there was close agreement between the pubescence of leaves of long and spur shoots.

3.2. Genetic variation

All loci studied were highly polymorphic as defined by the number of observed alleles, A, and gene diversity, H_E (Table 4). Part of the high allele counts was attributable to 26 alleles that occurred only once. Fifteen private alleles with frequencies exceeding 5% were found in the present study. Two of these were found in the sample of cultivars and thirteen in the sample of *M. sylvestris*. None of the private alleles found in *M. sylvestris* were restricted to one single wild population. Compared to the cultivars, the wild populations had as an average over loci, a higher allele cont, A, but slightly lower gene diversity, H_E. Both the higher number of private alleles and higher A is expected considering the larger sample size of the latter (178 wild individuals versus 29 cultivars). The lower estimate of H_E in the wild compared to the cultivar sample could possibly be explained by the development of many Danish cultivars from intended or unintended crosses between cultivars from diverse origins.

Table 6 The distribution of Nei's (1978) D_S (lower diagonal) and θ (upper diagonal) in pairwise comparisons between populations. All estimates are highly significant ($P < 0.001$) according to the permutation based exact test described in the text.

	HE	KA	KO	NO	CU
HE		0.034	0.017	0.008	0.079
KA	0.122		0.058	0.030	0.126
KO	0.060	0.209		0.021	0.113
NO	0.034	0.102	0.076		0.096
CU	0.539	0.772	0.697	0.648	

The test for pairwise gametic equilibrium between loci performed within populations showed no significance. No significant departures from Hardy-Weinberg proportions were detected at the population level using sequential Bonferroni correction (Holm, 1979). All populations showed high genetic variation when measured as an average over loci (Table 5). Low measures of F_{IS} were obtained suggesting that mating is random in the wild populations. It was noted, however, that the estimate for

population NO was relatively low (-0.029) which could be explained by the lack homozygotes which is to be expected in a small population (NO consists of only 40 individuals) of a self incompatible species. The analysis of allele frequency differences among wild populations showed only low but highly significant differentiation (θ = 0.030 (0.023-0.039 95%); $P < 0.0001$). The calculated Φ_{ST} turned out not significantly different from the allele identity based estimate, θ, implying that the effect of stepwise like mutations on differentiation of populations is negligible compared to migration and/or genetic drift. Alternatively the stepwise mutation model does not apply to the studied loci (Hardy *et al.*, 2003).
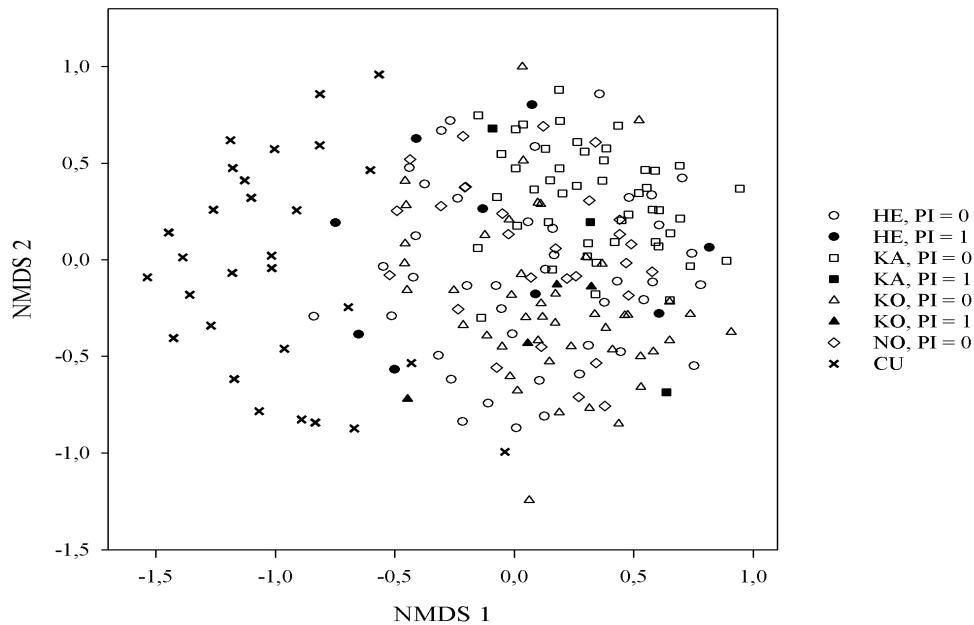
Figure 1 Ordination of individuals based on pairwise allele sharing distances, D_{AS}. Labeling according to sampling location and morphology as given by the pubescence index (PI) described in the text.

In accordance with the low overall differentiation between wild populations, the pairwise distance measures revealed only small differences between pairs of wild populations (Table 6). This was evident from both θ and D_S. As expected,

the greatest differentiation was observed between the cultivars and the wild populations, with HE showing the lowest and KA the highest differentiation. KA was also the population most highly differentiated in pairwise comparisons between wild populations.
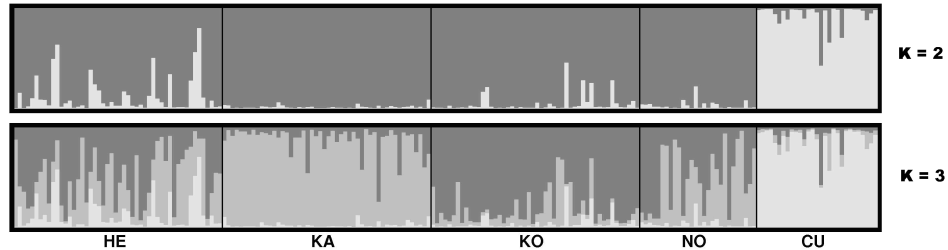

Figure 2 Proportional assignment, q_k, of individual genotypes to each of the $K = 2$ (top) and $K = 3$ (bottom) inferred clusters in the model based admixture analysis. Individuals are grouped by sampling location (diagram produced using DISTRUCT (Rosenberg, http://www.cmb.usc.edu/~noahr/distruct.html)).

3.3. Ordination

An ordination plot was constructed for the present account based on allele sharing distances, D_{AS}, between pairs of individuals by labeling points according to sampling location and pubescence index (Fig. 1). The first and second axis of the final NMDS-configuration accounted for 59% and 41% of the variation described by the method respectively.

On the one hand the NMDS plot show that wild and cultivated individuals fall into two separate clusters with somewhat distinct delineation. On the other hand it is evident that neither sampling locations for wild populations nor the two groups of the pubescence index are differentiated in the plot. These observations suggest that there is little genetic differentiation between wild populations but that wild Danish populations still form a gene pool that can be distinguished by genotypic means from *M. ×domestica*. Classifying points by the individual morphological characters studied (plots

not shown) did not provide any additional information. Strikingly, one cultivar clustered clearly with the wild populations and three individuals from HE clustered with the cultivars.
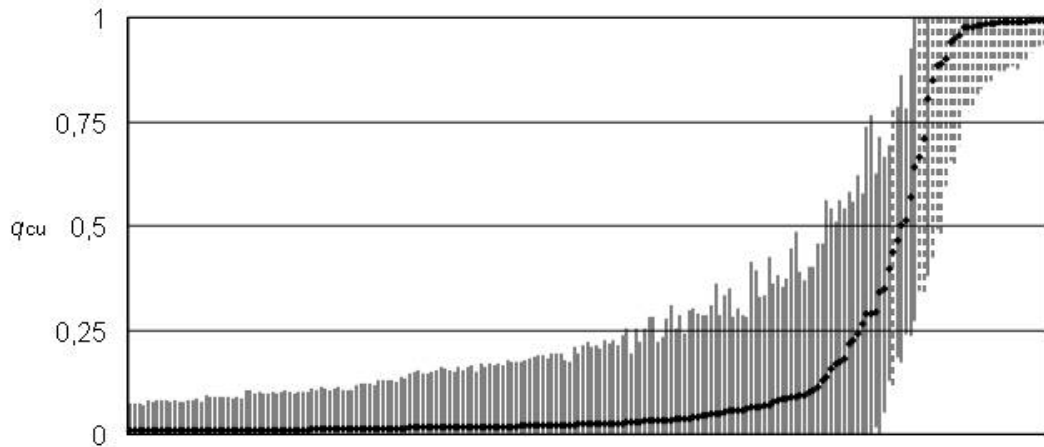

Figure 3 Illustration of the probabilistic assignment of individuals to the inferred cultivar cluster, q_{CU}. Values are plotted against their ranks. Vertical lines indicate the 95% probability intervals for individual values of q_{CU}. Full lines represent individuals collected in the wild; broken lines represent cultivars.

3.4. Genetic admixture analysis

Based on the likelihood output from the genetic admixture analysis using no prior information on sample origin, $K = 3$ clusters turned out to best describe the distribution of individuals within the collective sample. Despite having a slightly lower likelihood, $K = 2$ was selected for the further analyses because assignment of the individuals to two clusters offered a more clear biological interpretation (Fig. 2).

Estimated proportions of individual genotypes belonging to the inferred cultivar population, q_{CU}, was plotted hierarchically along with 95% probability intervals (Fig. 3). This measure showed a strong correlation ($r = 0.82$) with the value of the first NMDS-axis from ordination and suggests a good correspondence between the two methods. It is evident that the studied individuals fall into two main categories made up

of wild and cultivated individuals respectively with some individuals falling in between. All cultivars except one, had probability ranges including q_{CU} = 1.00 and all, except three, had values of q_{CU} higher than 0.85. In order to investigate if the individuals falling between the two main categories are of hybrid origin or just individuals with little information, the 85% cut off for the cultivar sample was taken to delineate a cultivar category and a 15% cut off to delineate a wild category. By doing so 26 individuals were assigned to the cultivar category, 23 (including three cultivars) to the intermediate category and 158 were assigned to the wild category.

Table 7 Summary of inferred categories. Categories inferred from the assignment of individuals in STRUCTURE. Number of individuals falling within each of the three categories is given for wild samples and cultivars respectively. For each category is shown the proportion (%) of individuals showing indication of hybrid ancestry from pubescence and fruit characters.

	Inferred category		
Sampling origin	**Wild**	**Intermediate**	**Cultivar**
Wild samples	158	20	0
Cultivars	0	3	26
Morphology			
Pubescence (PI=1)	6	30[*]	100
Fruit characters (hybrid)	18	64[*]	100

* Percentage given excludes cultivar samples in this category.

When comparing the phenotypic appearance of individuals falling into the intermediate and the wild category respectively, it was evident that a much higher proportion of hybrid typical phenotypes were present in the first category (Table 7). Although the morphological indices of hybridization indicates a higher proportion of hybrids in the intermediate category, it should be noted that the wider probability intervals of q_{CU} for this category are indicative of limited information and therefore some of these individuals might in fact belong to either the wild or the cultivar category.

When considering the distribution of q_{CU} within populations (Fig. 2) it is evident that the individuals of intermediate values of q_{CU} were unevenly distributed. Thus, of the 20 wild individuals in this category thirteen came from HE, six from KO,

one from NO and none from KA. This distribution of genotypically inferred hybrids between populations corresponds well with the findings of the morphological study. That is, HE appears to be greatly influenced by hybridization whereas NO appears to be unaffected. KA and KO are intermediate (cf. Table 2 and 3).

Table 8 Probabilities of individuals being immigrants or having recent immigrant ancestors from the *M. ×domestica* sample. The four most strongly inferred hybrids according to their probability of 'No immigrant ancestry' are shown.

Geographic origin	PI[*]	No immigrant ancestry	Immigrant	Immigrant parent	Immigrant grandparent
KO	1	0.006	0.000	0.696	0.297
HE	1	0.006	0.000	0.751	0.243
HE	1	0.009	0.051	0.845	0.095
HE	1	0.036	0.011	0.591	0.362

*: PI: pubescence based hybrid character index as described in the text.

The clustering approach using prior information on sampling location was applied for identification of individuals of migrant CU origin. Only four individuals from the wild samples could be strongly inferred to be of hybrid origin by their low probability of no immigrant ancestry. All four had a PI of 1 (Table 8). Notably, one cultivar was inferred equally likely to be a migrant of second and third generation (both *c.* 30%) from the wild cluster.

3.5. Flowering period

Great variation was found in duration and timing of flowering periods for wild individuals (Fig. 4) with individuals growing in sheltered locations having the shortest and earliest flowering period. Of the studied wild individuals 25 bore no flowers. When comparing wild individuals and cultivars from nearby orchards only slight overlap in flowering periods were found. Thus, peak flowering differed by *c.* one to two weeks between individuals of *M. sylvestris* and *M. ×domestica*.
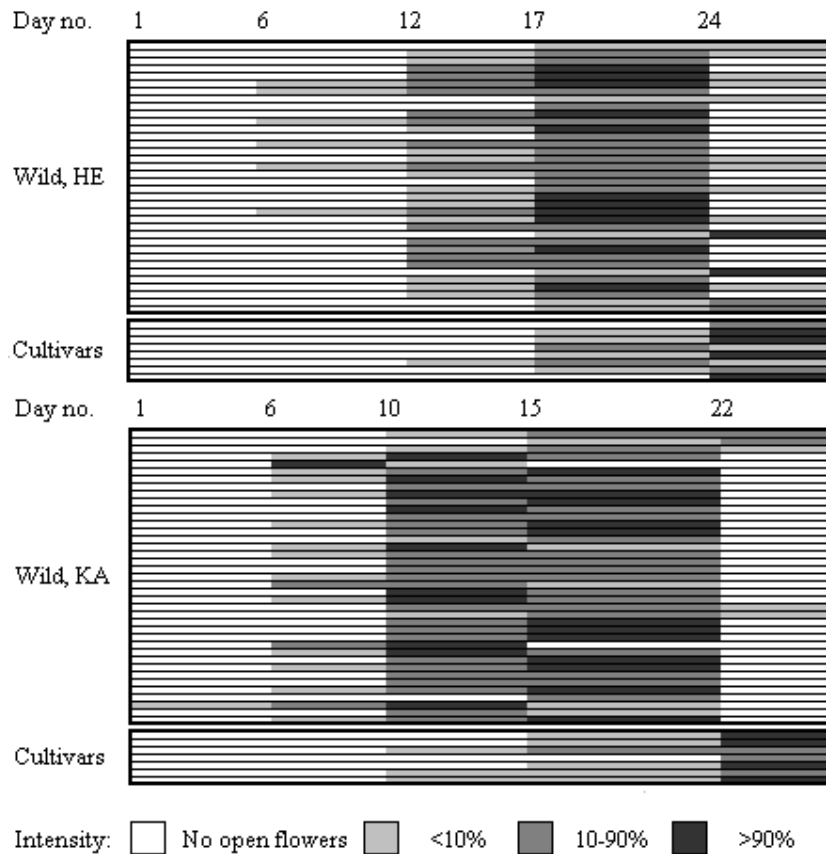
Figure 4 Flowering periods for wild and cultivated individuals at sampling location HE (top) and KA (bottom). Day number one was 05 May 2005 (HE) and 06 May 2005 (KA) respectively. Intensity of flowering is given in percentage of the observed number of flower buds on each tree.

4. Discussion

4.1. Hybridization between Malus sylvestris and Malus ×domestica

The case of *M. sylvestris* and *M. ×domestica* is one of the numerous examples of hybridization between pairs of closely related tree species in Europe, the study of which, is potentially very challenging due to a low number, and often great overlap of diagnostic morphological characters (Winfield *et al.*, 1998; Triest *et al.*, 1999; Nielsen and Olrik, 2001; Jensen *et al.*, 2003). Another challenge in these kinds of

studies is the frequent lack of information about timescale of hybridization and extent of historical introgression between the species studied. Therefore, in the present study analysis of molecular data focused on methods requiring no genotypic references for the parent species. Further, unravelling the relationship between species tends to be complicated when there is a great overlap in allele size ranges (Estoup *et al.*, 1999) as has been the case here.

From the present study it appears that the gene pools of the studied species can be separated by genotypic means as output from both the model and distance based analyses applied. Furthermore, the model based analysis only identified relatively few individuals as hybrids by intermediacy in genotype or migrant ancestry. The finding of two separate gene pools, consistent with the anticipated species identity, in this study is noteworthy in the light that it has been suggested that *M. sylvestris* and *M. ×domestica*, along with their Eurasien relatives, are one panmictic species (Kleinschmit *et al.*, 2000). Although dependent upon the character studied, the morphological study supported to some extent the finding that hybridization has not been so rampant that it has caused total admixture between the two species. However, this is in line with the findings of Coart *et al.* (2003) studying hybridization between the two species in Belgium and Germany. In this study they found 3 hybrids out of a total of 76 putative wild individuals studied.

Finding two separate gene pools suggests that hybridization between the two species is not as frequent in Denmark as might have been expected. In fact, only four wild individuals of recent hybrid ancestry were found when studying migration. To some degree, however, the model is dependent upon the specified migration rate which in the model is assumed identical for the studied populations. Using migration rates in

the range of 0.1 to 1% yielded consistent results indicating that the model is appropriate for the analyzed data set (Pritchard *et al.*, 2000).

The fact that the studied microsatellite loci show both Hardy-Weinberg proportions and gametic equilibrium within populations, further supports the finding that hybridization between *M. ×domestica* and *M. sylvestris* is not likely to be a process of any significant scale at the present time as this would have created deviations from the expected genotypic frequencies and gametic phases (Nei and Li, 1973; Flint-Garcia *et al.*, 2003).

4.2. Factors influencing hybridization

Strikingly, a large variation in the occurrence of hybrids as assessed from both molecular and morphological indications existed among populations. This observation is consistent with the inferred isolation from *M. ×domestica* at the regional and local scale of the sampling locations. It is therefore suggested that an important barrier to hybridization is isolation by distance. Thus, if the seed and pollen dispersal vectors are unlikely to move between the habitats of the two species, hybridization is likely to be limited. Although pollination by insects has traditionally been considered unlikely to cause long-distance gene flow (Slatkin, 1985) it appears that such events are likely to appear more frequently than hypothesized so far (Smouse and Sork, 2004).

The present study suggests that differences in flowering phenology could be an additional explanation of the apparent low frequency of hybridization found in natural populations. Thus, the studied cultivars of *M. ×domestica* were found to flower later than their wild relatives, although some overlap was observed. However, further studies of this issue are needed in order to draw final conclusions. In particular, it

should be established whether the observed difference in flowering period is consistent over a longer period. Phenological barriers were also found to hinder hybridization of closely related *Malus* species in America by Dickson *et al.*, (1991).

Outcrossing depression (Triest *et al.*, 1999) has also been suggested as a possible explanation when discovering an apparent lack of hybridization between closely related species under natural conditions and could be applicable to the case of *M. ×domestica* and *M. sylvestris*. On the other hand, gametophytic incompatibility seems unimportant in this respect, considering the crossability within subfamily Maloideae in general (Korban, 1986) and between the two studied species in particular (A. S. Larsen, unpublished observations).

4.3. Concordance between morphological and molecular characterization at the population and individual level

Although consistent at the population scale, morphological and molecular evidence of hybridization was found to be divergent at the individual level. In analyzing individuals of intermediate values of q_{CU}, it is evident that the proportion of morphologically inferred hybrids in this category was markedly higher than in the cluster of wild individuals. However, in numerical terms the greatest number of morphologically inferred hybrids was found in the cluster of wild individuals defined by molecular genotype. This is consistent with introgression and is further supported by the observed inconsistency at the individual level between morphological characters in defining the hybrid status of the individuals. The latter is to be expected when hybrids are backcrossed to their parent population since many morphological characters in plants are unlinked and characterized by simple genetic inheritance (Rieseberg and

Ellestrand, 1993; Rieseberg, 1995). Undetectable introgression was also suggested by Beaumont *et al.* (2001) as an explanation for the great morphological variation observed within a group of genotypically designated wildcats using the same methods as in the present study. An alternative explanation would be that phenotypical variation within the species *M. sylvestris* is larger than has been acknowledged so far.

Unlinked morphological characters and simple inheritance are also likely explanations for the occurrence of wild phenotypes in the intermediate category suggestive of hybrid ancestry. That is, morphological observations do not always correspond with molecular data in studies of hybridization (Nason *et al.*, 1992; Rieseberg and Ellstrand, 1993; Perron and Bousquet, 1997; Triest *et al.*, 1999; Watano *et al.*, 2004). Ontogenetic variation in morphological characters and phenotypic plasticity within *M. sylvestris* and its hybrids could further complicate the interpretation of individual phenotypes and cause misclassification. These issues are especially problematic because of the complicated interactions between them, which potentially make morphological statements dubious when no information is available on plant age and development (Coleman *et al.*, 1994).

In addition to these biological explanations the issue of information in the studied microsatellite loci could potentially be of great importance in explaining discrepancies in geno- and phenotypes at the individual level. Thus, increasing the genotypic information decreases the variation on q_k and the pairwise genetic distances in the model based clustering and ordination respectively (Gower, 1966; Pritchard *et al.*, 2000). This means that imprecise genotypic inference could also be an issue in the observed discrepancies between morphological and molecular characters. Correlation between the two types of characters at the population scale, however, suggests that the

overall picture of hybridization and introgression as well as its relative distribution between the studied populations is plausible.

A noteworthy result came out of analyzing the cultivar sample. Thus, the cultivar sample showed neither gametic disequilibrium nor deviation from Hardy-Weinberg proportions. This is interesting when considering the hypothesis that today's cultivars are the descendants of a limited number of domesticated individuals derived from a large geographical region and maintained by vegetative propagation for millennia (Juniper *et al.*, 1999).

4.4. Genetic diversity within and among wild populations

Estimates of genetic diversity in the wild populations as estimated by gene diversity (H_E = 0.78), was in line with that found for *M. sylvestris* by Coart *et al.* (2003) for Belgian populations using microsatellite data (H_E = 0.72). However estimates of heterozygote deficiency (F_{IS} = -0.002) and differentiation between populations (θ = 0.030) deviated somewhat from that reported by Coart *et al.* (2003) (F_{IS} = 0.105; F_{ST} = 0.06). A possible explanation for this deviation could be the fact that the latter study included substantially larger sampling areas (1 – 20 km^2) than the present study (0.1 – 1 km^2) thereby introducing a Wahlund effect. The estimate of microsatellite based H_E in the present study is similar to that of other outcrossing species of the temperate region of Europe (H_E = 0.79 for *Fraxinus excelcior* L. (Heuertz *et al.*, 2003); H_E = 0.86 for *Quercus petrea* (Matt.) Liebl. (Jensen *et al.*, 2003)). Being an outcrossing species, it is not surprising that a low excess of homozygotes (F_{IS} = -0.002) is observed. The present scattered distribution of populations and the fact that it is an insect pollinated species would suggest a potential differentiation between populations (Hamrick *et al.*, 1992;

Myking, 2002). The low level of differentiation between populations ($\theta = 0.030$) is, however, in accordance with results reported for species with similar life history traits (*Prunus avium* L. ($G_{ST} = 0.014$) (Mariette *et al.*, 2004) and *Sorbus aucuparia* L. ($\theta = 0.043$) (Bacles *et al.*, 2004) both using isozyme markers). This finding is suggestive of a high rate of gene flow between wild populations.

High gene flow between wild Danish populations is supported by the fact that none of the private alleles in *M. sylvestris* ($p > 0.05$) were restricted to single populations and that the microsatellite data did not support the stepwise mutation model. Being a pioneer species, this gene flow between populations is likely to have occurred by repeated events of extinction and recolonization as proposed by Slatkin (1985). Under this hypothesis, the maintenance of a high gene diversity would be attributable to the long juvenile period of trees reducing the founder effect by allowing colonization by a great number of immigrants before initial seed set (Austerlitz *et al.*, 2000; Mariette *et al.*, 2004). Long distance seed dispersal between established populations might have played only a minor role in maintaining gene diversity and reducing differentiation, as germination and establishment of *M. sylvestris* is hampered in a closed stand (Buttenschøn and Buttenschøn, 1998). Also introgressive gene flow from *M. ×domestica* will hinder loss of genetic diversity within and among wild populations. However, the great variation in introgression across populations observed in this study suggests that this cannot be the only explanation.

4.5. Consequences for the establishment of a conservation program for Malus sylvestris

Despite the low level of differentiation and lack of private alleles that suggests some level of historical gene flow, the dynamic changes in the Danish

landscape over the recent centuries means that one cannot assume an equilibrium situation. The present migration rates may be significantly different today. Also, even though the Danish population of *M. sylvestris* appears genetically viable when considering the high gene diversity observed in the studied populations a contentious point must be made because this may not assure that the population remains so (Young *et al.*, 1996). That is, because flowering is poor in the closed forests and scrubs, which are presently the most frequent habitat for major populations of *M. sylvestris*, variability of the offspring might be low and the degree of inbreeding depression is likely to be elevated (Charlesworth and Charlesworth, 1987). Also, the present study only estimated differentiation in neutral molecular markers. Although little differentiation was found between wild populations such differences are often poorly correlated with the differentiation in adaptive traits, and disregard will potentially lead to outbreeding depression if not taken into account when establishing conservation programs for endangered species (McKay and Latta, 2002; Storz, 2002).

In conclusion, the finding of only very few individuals indicative of recent hybridization in the present study suggests that hybridization between *M. ×domestica* and *M. sylvestris* is likely to be rare under Danish conditions. On the other hand, introgression from *M. ×domestica* into the gene pool of the indigenous *M. sylvestris* is shown to have taken place during the 800 years of apple growing in Denmark. However, it seems evident that there is great regional variation in the extent to which this has occurred. Further, even in the most admixed populations is it possible to find morphologically authentic individuals of wild apple trees, which along with the existence of two genotypically distinct gene pools, suggests that assimilation of the *M. sylvestris* gene pool is not as advanced as has been feared. Also, the present study

illustrates limitations in using both molecular and morphological characters for identification of hybrids at the individual level. This, however, is likely to be alleviated by increasing the number of marker loci and morphological characters respectively. On the other hand, the correspondence between molecular genotype and phenotype at the population level suggests that the present approach is a useful tool for identification of conservation stands in the future. In relation to the morphological characters studied, it turned out that pubescence of leaves from long and spur shoots were almost completely similar so that only one of these characters should be included. Also, interpreting records of long shoot diameters, which were initially included in the study, appeared to be hard due to unclear environmental effects. Finally, on condition that sufficiently varied offspring can be gained, the low present rate of hybridization along with great genetic variation found in the wild populations suggests that a sensible strategy for mobilizing the wild gene pool for use in future plantings would be the establishment of seedling seed orchards. Such orchards should be established from populations predicted to be little affected by hybridization and introgression as could be established from applying a combined molecular and morphological approach as in the present study.

provision of plant material and Vinnie Deichmann, Nancy Mergan, and Sabine Van Glabeke for excellent technical assistance.

References

Austerlitz F, Mariette S, Machon N, Gouyon PH, Godelle B (2000) Effects of colonization processes on genetic diversity: Differences between annual plants and tree species. Genetics 154: 1309-1321

Bacles CFE, Lowe AJ, Ennos RA (2004) Genetic effects of chronic habitat fragmentation on tree species: the case of *Sorbus aucuparia* in a deforested Scotttish landscape. Mol Ecol 13: 573-584

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. Mol Ecol 11: 155-165

Beaumont MA, Barratt EM, Gotelli D, Kitchener AC, Daniels MJ, Pritchard JK, Bruford MW (2001) Genetic diversity and introgression in the Scottish wildcat. Mol Ecol 10: 319-336

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368: 455-457

Buttenschøn RM, Buttenschøn J (1998) Population dynamics of *Malus sylvestris* stands in grazed and ungrazed, semi-natural grasslands and fragmented woodlands in Mols Bjerge, Denmark. Ann Bot Fenn 35: 233-246

Campbell CS, Greene CW, Dickinson TA (1991) Reproductive biology in subfam. Maloideae (Rosaceae). Syst Bot 16: 333-349

Charlesworth B, Charlesworth D (1987) Inbreeding depression and its evolutionary consequences. Annu Rev Ecol Syst 18: 237-268

Chase MW, Hills HH (1991) Silia gel: An ideal material for field preservation of leaf samples for DNA studies. Taxon 40: 215-220

Coart E, Vekemans X, Smulders MJM, Wagner I, van Huylenbroeck J, van Bockstaele E, Roldán-Ruiz I (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

Coleman JS, McConnaughay KDM, Ackerly DD (1994) Interpreting phenotypic variation in plants. Trends Ecol Evol 9: 187-191

Dickson EE, Kresovich S, Weeden NF (1991) Isozymes in North American *Malus* (Rosaceae): Hybridization and species differentiation. Syst Bot 16: 363-375

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19: 11-15

Eriksson G, Namkoong G, Roberds J (1995) Dynamic conservation of forest tree gene resources. Forest Genetic Resources, FAO 23: 2-8

Estoup A, Cornuet JM, Rousset F, Guyomard R (1999) Juxtaposed microsatellite systems as diagnostic markers for admixture: theoretical aspects. Mol Biol Evol 16: 898-908

FAO (1989) Plant Genetic Resources. Their conservation in situ for human use. FAO, Rom

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Annu Rev Plant Biol 54: 357-374

Gianfranceschi L, Seglias N, Tarchini R, Komjanc M, Gessler C (1998) Simple sequence repeats for the genetic analysis of apple. Theor Appl Genet 96: 1069-1076

Gower JC (1966) Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325-338

Graudal L, Kjær ED, Canger S (1995) A systematic approach to conservation of genetic resources of trees and shrubs in Denmark. For Ecol Manag 73: 117-134

Goldstein DB, Pollock DD (1997) Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. Journal of Heredity 88: 335-342

Guilford P, Prakash S, Zhu JM, Rikkerink E, Gardiner S, Bassett H, Forster R (1997) Microsatellites in *Malus X domestica* (apple): Abundance, polymorphism and cultivar identification. Theor Appl Genet 94: 249-254

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 48: 361-372

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. New Forest 6: 95-124

Hardy OJ, Charbonnel N, Fréville H, Heuertz M (2003) Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. Genetics 163: 1467-1482

Hardy OJ, Vekemans X (2002) SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes 2: 618-620

Heuertz M, Vekemans X, Hausman JF, Palada M, Hardy OJ (2003) Estimating seed vs. pollen dispersal from spatial genetic structure in the common ash. Mol Ecol 12: 2483-2495

Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6: 65-70

Jarvis DI, Hodgkin T (1999) Wild relatives and crop cultivars: detecting natural introgression and farmer selection of new genetic combinations in agroecosystems. Mol Ecol 8: 159-173

Jensen JS, Olrik DC, Siegismund HR, Lowe AJ (2003) Population genetics and spatial autocorrelation in an unmanaged stand of Quercus petraea in Denmark. Scand J Forest Res 18: 295-304

Jones CJ, Edwards KJ, Castaglione S *et al.* (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. Mol Breeding 3: 381-390

Juniper BE, Watkins R, Harris SA, Tobutt KR, Alston FH (1999) The origin of the apple. Acta Horticulturae 484: 27-33

Kjær ED, Amaral W, Yanchuk A, Graudal L (2004) Strategies for conservation of forest genetic resources. In: FAO, FLD, IPGRI (2004) Forest genetic resources conservation and management. Vol 1: Overview, concepts and some systematic approaches. International Plant Genetics Ressources Institute, Rome, pp 5-24

Kleinschmit J, Stephan R, Wagner I (2000) Wild fruit trees (Prunus avium, Malus sylvestris and Pyrus pyraster) genetic resources conservation strategy. EUFORGEN, International Plant Genetic Resources Institute. http://www.ipgri.cgiar.org/networks/euforgen/Networks/Noble_Hardwoods/Strategies/WildFruitConsStrategy.htm.

Korban SS (1986) Interspecific hybridization in Malus. Hortscience 21: 41-48

Kugler H (1970) Blütenökologie. 2nd ed. Gustav Fischer Verlag, Stuttgart

Ledig FT (1991) The role of genetic diversity in maintaining the global ecosystem. In: FAO (1991) Forests, a Heritage for the Future: 10th World Congress, FAO, Rome, pp 71-78.

Liebhard R, Gianfranceschi L, Koller B, Ryder CD, Tarchini R, van de Weg E, Gessler C (2002) Development and characterisation of 140 new microsatellites in apple (*Malus x domestica* Borkh.). Molecular Breeding 10: 217-241

Linder CR, Taha I, Seiler GJ, Snow AA, Rieseberg LH (1998) Long-term introgression of crop genes into wild sunflower populations. Theor Appl Genet 96: 339-347

Lynch M (1991) The Genetic Interpretation of Inbreeding Depression and Outbreeding Depression. Evolution 45: 622-629

MAAF (1973) Flowering periods of tree and bush fruits. Ministry of Agriculture, Fisheries and Food, Technical bulletin No. 26. Her Majesty's Stationery Office, London.

Manly BFJ (1994) Multivariate statistical methods - a primer. 2nd ed. Chapman & Hall, London.

Mariette S, Lefranc M, Legrand P, Taneyhill D, Frascaria-Lacoste N, Machon N (2004) Genetic variability in wild cherry populations in France. Effects of colonizing processes. Theor Appl Genet 94: 904-908

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. Trends Ecol Evol 17: 285-291

Michalakis Y, Excoffier L (1996) A genetic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics 142: 1061-1064

Muir G, Fleming CC, Schlötterer C (2000) Species status of hybridizing oaks. Nature 405: 1016

Myking T (2002) Evaluating genetic resources of forest trees by means of life history traits - a Norwegian example. Biodivers Conserv 11: 1681-1696

Nason JD, Ellstrand NC, Arnold ML (1992) Patterns of hybridization and introgression in populations of oaks, manzanitas and irises. Am J Bot 79: 101-111

National Research Council (1991) Managing global genetic resources. Forest trees. National Academy Press, Washington D.C.

Nei M (1978) Estimation of Average Heterozygosity and Genetic Distance from A Small Number of Individuals. Genetics 89: 583-590

Nei M, Li W-H (1973) Linkage disequilibrium in subdivided populations. Genetics 75: 213-219

Nielsen J, Olrik DC (2001) A morphometric analysis of *Prunus spinosa*, *P. domestica* ssp. *insititia*, and their putative hybrids in Denmark. Nord J Bot 21: 349-363

Ødum S (1968) The distribution of trees and shrubs in Denmark. Bot Tidsskr 64 (in Danish with English abstract and legends)

Perron M, Bousquet J (1997) Natural hybridization between black spruce and red spruce. Mol Ecol 6: 725-734

Ponomarenko VV (1985) Wild apples - a useful genetic source for breeding. Sadovodstvo 5: 13-14

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945-959

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86: 248-249

Remmy K, Gruber F (1993) Untersuchungen zur Verbreitung und Morphologie des Wild-Apfels (*Malus sylvestris* (L.) Mill.). Mitteilungen der Deutchen Dendrologischen Gesellschaft 81: 71-94 (English summary)

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. Annu Rev Ecol Syst 27: 83-109

Rieseberg LH (1995) The role of hybridization in evolution: old wine in new skins. Am J Bot 82: 944-953

Rieseberg LH, Ellstrand NC (1993) What Can Molecular and Morphological Markers Tell Us About Plant Hybridization. Crit Rev Plant Sci 12: 213-241

Slatkin M (1985) Gene flow in natural populations. Annu Rev Ecol Syst 16: 393-430

Smouse PE, Sork VL (2004) Measuring pollen flow in forest trees: an exposition of alternative approaches. For Ecol Manag 197: 21-38

Storz JF (2002) Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. Mol Ecol 11: 2537-2551

Triest L, de Greef B, de Bondt R, van Slycken J (1999) RAPD of controlled crosses and clones from the field suggests that hybrids are rare in the *Salix alba-Salix fragilis* complex. Heredity 84: 555-563

Wagner I (1996) Zusammenstellung morphologischer Merkmale und ihrer Ausprägungen zur Unterscheidung von Wild- und Kulturformen des Apfel- (*Malus*) und des Birnbaumes (*Pyrus*). Mitteilungen der Deutchen Dendrologischen Gesellschaft 82: 87-108

Wagner I (1998a) Artenschuts bei Wildapfel. Die Blattbehaarung von 116 Apfelklonen auf zwei Samenplantagen. Forst und Holz 53: 40-43 (English summary)

Wagner I (1998b) Evaluierung der Wildformen von Apfel und Birne. Schriften aus der Forstlichen Fakultät der Universität Göttingen und der Niedersächsischen Versuchsanstalt 125; 68-82

Wagner I, Schmitt HP, Maurer W, Tabel U (2004) Isozyme polymorphism and genetic structure of *Malus sylvestris* (L.) Mill. Native in western areas of Germany with respect to *Malus* X *domestica* Borkh. Acta Horticulturae 663: 545-550

Watano Y, Kanai A, Tani N (2004) Genetic structure of hybrid zones between *Pinus pumila* and *P. parviflora* var. *pentaphylla* (Pinaceae) revealed by molecular hybrid index analysis. Am J Bot 91: 65-72

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. Evolution 38: 1358-1370

Winfield MO, Arnold GM, Cooper F, Ray ML, White J, Karp A, Edwards KJ (1998) A study of genetic diversity in *Populus nigra* subsp. *betulifolia* in Upper Severn area of the UK using AFLP markers. Mol Ecol 7: 3-10

Wright S (1978) Variability within and among natural populations, vol. 4. The University of Chicago Press, Chicago

Young A, Boyle T, Brown L (1996) The population genetic consequences of habitat fragmentation for plants. Trends Ecol Evol 11: 413-418