American Journal of Botany

American Journal of Botany 97(6): 1014–1024. 2010.

Fine-scale comparisons of genetic variability in seed families of asexually and sexually reproducing *Crataegus* (hawthorn; Rosaceae)[1]

Eugenia Y. Y. Lo[2,3,5,7], Saša Stefanović[3], Kermit Ritland[4], and Timothy A. Dickinson[5,6]

[2]Department of Ecology and Evolutionary Biology, Yale University, 21 Sachem Street, New Haven, Connecticut 06511 USA; [3]Department of Biology, University of Toronto Mississauga, 3359 Mississauga Road N, Mississauga, Ontario, Canada L5L 1C6; [4]Department of Forest Sciences, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4; [5]Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, Ontario, Canada M5S 1B1; and [6]Green Plant Herbarium, Royal Ontario Museum, 100 Queen's Park, Toronto, Ontario, Canada M5S 2C6

The reproductive system is one of the key mechanisms that determine genetic diversity at different biological levels. However, few attempts have been made to assess the consequences of asexual reproduction by comparing genetic structure and fecundity of seed families in asexually and sexually reproducing individuals. We have examined two similar hawthorn species, *Crataegus crus-galli* and *C. punctata*, that differ in ploidy level and breeding system. Fecundity (per-fruit seed set) and microsatellite genotypes for five loci were determined in 18 and 26 seed families of *C. crus-galli* and *C. punctata* (totals of 83 and 118 embryos), respectively. Compared with the sexual diploid *C. punctata*, tetraploid *C. crus-galli* shows similar fecundity, but lower genotypic diversity within and between seed families. Reproduction in the tetraploid, while predominantly apomictic, is nevertheless accompanied by outcrossing and selfing. We conclude that in *C. crus-galli* pollen flow between conspecific individuals is limited, and the combination of pollen fertility, self-compatibility, and pseudogamous apomixis provides reproductive assurance in these tetraploids. Reproductive assurance, in turn, may explain the derived floral architecture seen in most North American tetraploid hawthorns. We also discuss analytical approaches for inferring mating-system parameters in tetraploids and for comparing microsatellite variation across ploidy levels.

Key words: Apomixis; family genetic structure; fecundity; embryo genotypes; microsatellite; outcrossing; pollen flow; polyploidy; progeny array; Rosaceae.

Many 20th-century biologists predicted that the founder effect and the lack of segregation and recombination in asexual populations would usually lead to a deficiency in genetic variation and eventually hamper their potential response to selection (Stebbins, 1950; Darlington, 1958; Grant, 1981; Lynch and Gabriel, 1983; Kondrashov, 1993). Molecular data now provide support for contrasting views such as those of Clausen (1954), who drew attention to the facultative nature of asexual reproduction and pointed out that asexual reproduction would merely reduce the frequency with which genetic variants arise. Molecular studies have clearly illustrated varying levels of genetic diversity in apomicts at the population level (e.g., Carino and Daehler, 1999; Esselman et al., 1999; Van Der Hulst et al., 2000; Paun et al., 2006; Hörandl and Paun, 2007; Whitton et al., 2008). However, the influence of asexual reproduction on genetic diversity and fecundity at the level of seed families remains uninvestigated.

Recurrent gene flow through interspecific crosses (Menken et al., 1995; Dijk, 2003; Thompson and Whitton, 2006; Hörandl and Paun, 2007; Nybom, 2007; Talent and Dickinson, 2007b) and occasional sexuality (i.e., meiosis and gamete fusion, either separately or together) have been suggested as mechanisms for creating and multiplying novel genotypes in agamospermous plants (Gustafsson, 1942; Clausen, 1954; Ellstrand and Roose, 1987; Bayer et al., 1990; Thompson et al., 2008; Whitton et al., 2008 and references therein). The frequency of such events will depend on the amount of pollen flow between individuals not only for the endosperm fertilization, which is required in asexual seed formation in pseudogamous apomicts, but also for embryo fertilization when seeds are produced sexually (for terminology describing apomixis used here, see Talent, 2009). However, most studies devoted to population dynamics of asexual plants have addressed gene flow via seed dispersal (e.g., Durand et al., 2000; Garnier et al., 2002; Rogstad et al., 2002; Štorchová et al., 2002; Paun et al., 2006; Thompson and Whitton, 2006). A notable exception is the paper by Oddou-Muratorio et al. (2001) on *Sorbus torminalis*, in which both pollen- and seed-mediated gene flow were found to be important.

The occurrence of asexual reproduction has been correlated with several biological and ecological phenomena. These include the breakdown of self-incompatibility (Mable, 2008; Whitton et al., 2008) and changes in pollinator densities (Ashman et al., 2004), abiotic environments (Michaels and Bazzaz, 1989; Dijk, 2003; Houliston et al., 2006), and reproductive fitness (Meirmans

[1] Manuscript received 31 March 2009; revision accepted 19 March 2010.

The authors thank T. Paton and S. Russell of the Centre for Applied Genomics, The Hospital for Sick Children, Toronto, for the ABI facilities; S. Nguyen for making the seed collections; and J. Chung for extracting seeds and helping with DNA extraction. N. Talent and three anonymous reviewers generously commented on earlier drafts of this paper. Financial support from the Natural Sciences and Engineering Research Council of Canada (grant A3430 to T.A.D. and 326439-06 to S.S.), the Botany Department of the University of Toronto, and the Royal Ontario Museum is gratefully acknowledged, as is an award from the Royal Ontario Museum Foundation to Mark Engstrom and T.A.D. for the purchase of the thermocycler used in part of this work.

[7] Author for correspondence (e-mail: eugenia.lo@yale.edu)

doi:10.3732/ajb.0900091

et al., 2006; Mogie et al., 2007; Voigt et al., 2007; Hörandl, 2008). In relatively long-lived woody plants, the consequences of asexual reproduction at the level of populations are of some importance because taxonomic conclusions drawn from phenetic data over the course of a few decades may be based on only one or a few generations at most. Assumptions about genetic relatedness and reproductive success of individuals appear to influence many taxonomic decisions (Dickinson, 1999). In the case of North American hawthorns (*Crataegus* L.; Rosaceae subtribe Pyrinae), their notorious complexity stems mainly from taxonomists having recognized morphological variants as species on the assumption that local aggregations of individuals represent panmictic populations. Alternatively, these same aggregations can be seen as comprising large numbers of putatively sibling individuals that are inferred to have originated apomictically (Dickinson, 1986). However, direct genetic evidence for this view is lacking; it is based instead on inferences from morphometric (Dickinson and Phipps, 1985, 1986) and flow-cytometric data (Talent and Dickinson, 2007a, b).

Hawthorns are ruderal shrubs and small trees with unspecialized insect-pollinated flowers and fleshy fruits (polypyrenous drupes). Diploid hawthorns resemble other Rosaceae in exhibiting gametophytic self-incompatibility (Dickinson and Phipps, 1986; Dickinson et al., 2007). By contrast, in polyploids asexual production of seeds (agamospermy) is common and results from gametophytic apomixis, in which unreduced female gametophytes develop from somatic cells of the nucellus rather than from megaspores (apospory). The unreduced egg cells develop parthenogenetically to give rise to embryos that are genetically identical to their parent plant (Nogler, 1984). Although embryo development is uniparental, fertilization of polar nuclei by one or both sperm nuclei is prerequisite for endosperm formation (generally known as pseudogamy; Muniyamma and Phipps, 1979, 1984; Dickinson et al., 2007; Talent and Dickinson, 2007a).

To examine pollen flow within populations of hawthorns and test for a difference in genetic variation and per-fruit fecundity between sexual and apomictic individuals, we present genetic data from a comparison of sexually and asexually produced seed families from *Crataegus punctata* Jacq. and *C. crus-galli* L., respectively, two species from eastern North America that have both been ascribed to *C*. sect. *Crus-galli* (Phipps et al., 1990). Specifically, we seek to (1) examine genetic variation of progeny within and between multiple seed families of the two species; (2) estimate rates of outcrossing, selfing, and apomixis so as to infer pollen flow among individuals; and (3) compare seed fecundity (i.e., the ratio of seeds successfully set per fruit) between the two species. In addition, we discuss the applicability of three analytical methods that are helpful when comparing genetic variation with microsatellite data between ploidy levels.

MATERIALS AND METHODS

Plant materials—In Ontario, Canada, *C. crus-galli* L. is polyploid, self-fertile, and facultatively apomictic (Dickinson and Phipps, 1985, 1986), whereas *C. punctata* Jacq. is diploid, self-incompatible, and sexual (Dickinson and Phipps, 1986; Talent and Dickinson, 2007a). Despite differences in the reproductive strategy of the two species, they both have pollen-to-ovule ratios in excess of 2000 (Dickinson and Phipps, 1986). Individual local populations of *C. crus-galli* were found to be morphologically quite homogeneous. At the same time, these populations are markedly differentiated from each other with respect to variation in their flowers, fruits, and leaves (Dickinson and Phipps, 1985; Dickinson, 1986). Such morphological differentiation between populations is considerably less pronounced in *C. punctata* (Dickinson and Phipps, 1986). In this species, within-population morphological variation is greater than between-population variation (Dickinson, 1986).

Seed families were randomly sampled using an "ignorant person" strategy (Ward, 1974) at each of two sites in southern Ontario (Table 1). At site ON04, all sampled individuals were tagged and identified unambiguously as *C. crus-galli* (glossy, unlobed leaves; 10-stamen flowers). At site ON46, all sampled individuals were tagged and identified unambiguously as *C. punctata* (pubescent, slightly lobed leaves with secondary venation markedly impressed on the adaxial surface; 20-stamen flowers). All of the pyrenes in each fruit were opened, so as to extract any seed present, but only a single seed from each fruit was selected for DNA extraction (many pyrenes enclosed no seeds at all). For each tree (i.e., for each seed family), DNA was extracted from four to eight embryos that were expected to share the same maternal genotypes. This number varied with the success of obtaining sufficient yields of DNA from a single embryo. A total number of 83 embryos representing 18 *C. crus-galli* families and 118 embryos representing 26 *C. punctata* families were included in the analyses. *Crataegus crus-galli* at site ON04 and *C. punctata* at site ON46 (Table 1) are well documented by many specimens deposited in the Herbarium of the Biology Department of the University of Western Ontario (UWO) and in the Green Plant Herbarium of the Royal Ontario Museum (TRT) (Dickinson and Phipps, 1985).

Measure of per-fruit fecundity—Flowers of *C. crus-galli* at site ON04 are characterized by the presence of (0-)1-2 gynoecial units (i.e., locules and associated styles), whereas those of *C. punctata* at site ON46 and elsewhere have 2–5 gynoecial units (Dickinson, 1983; T. A. Dickinson, unpublished data). Each gynoecial unit forms a pyrene in the mature fruit that may contain a single seed. A total of 192 and 264 fruits of *C. crus-galli* and *C. punctata*, with 10–37 fruits per tree, were examined for the number of fully developed seeds in each fruit. For each tree, the total number of seeds obtained is plotted against the total number of fruits that were opened.

DNA extraction and microsatellite analyses—Seeds of *C. crus-galli* and *C. punctata* are characterized by a large embryo covered with a thin translucent layer of endosperm usually adhering to the brown seed coat when fully developed. For each seed, the seed coat and the attached endosperm were carefully removed, and the remaining white embryo tissue of about 8–13 mg fresh weight was used for DNA extraction. The extraction protocol described in Lo et al. (2007) was applied to seed tissues, except that all volumes were reduced by half and the final DNA pellet was eluted in 20 μL of water. Because of the limited DNA quantity for each seed sample, only five dinucleotide microsatellite loci were used, and they are located on linkage groups 12 and 14 of *Malus* × *domestica* (Liebhard et al., 2002; Table 2). Details of polymerase-chain-reaction conditions and fragment analyses are described in Lo et al. (2009b). To assess the dosage of the amplified products (observed alleles) in the samples and designate genotypic configurations for tetraploid individuals, we used the microsatellite DNA allele counting-peak ratios method (MAC-PR; Esselink et al., 2004) based on quantitative values for microsatellite allele-amplification peak areas.

Linkage disequilibrium analyses and marker independence—To examine whether the five SSR loci represent a set of independent markers in the *Crataegus* genome, genotype linkage disequilibrium (LD) was tested by Fisher's exact test for each pair of loci (a total of 10 possible pairwise combinations) with GENEPOP, version 3.3 (Raymond and Rousset, 1995), using the Markov chain method with 100 batches and 10 000 iterations per batch. Because multiple tests were involved, the sequential Bonferroni correction was applied to test for significance (Rice, 1989). Because of the computational burden that would result from multiple possible pairwise combinations of alleles in polyploids with Fisher's exact test, LD tests were performed only on 2*x* individuals in our sample. Therefore, because the same sets of SSR loci are used in 2*x* and 4*x* individuals, we assumed that results of LD among loci based on diploids could also be applied to 4*x C. crus-galli*.

Analyses of genetic variation in progeny arrays—Multilocus genotype diversity for each family in *C. crus-galli* (18 families) and *C. punctata* (26 families) was estimated using the software GENODIVE, version 2.0b4 (Meirmans and Tienderen, 2004). The stepwise mutation model (SMM), which is suggested to be a more accurate model for microsatellite mutations (Morris et al., 1995; Fu and Chakraborty, 1998; Hardy et al., 2003; Whittaker et al., 2003; Paun and Hörandl, 2006), was used to determine the frequency distribution of pairwise dissimilarity among embryos of the two taxa. Threshold values have been shown to vary among species, and differences in the threshold setting may

TABLE 1. Descriptive statistics of genetic variation at the five microsatellite loci (see Table 2) for seed families of *Crataegus crus-galli* and *C. punctata.* Diversity indices were estimated at 0%, 2%, 4%, and 6% threshold limits. *G* = number of multilocous genotypes, *PG* = proportion of distinguishable genotypes, and *D* = multilocous genotype diversity. In *C. crus-galli*, *G*, *PG*, and *D* estimates are not significantly different among thresholds. However, in *C. punctata*, they increase significantly when the thresholds increase from 0 to 4%. No significant differences are found between estimates at the thresholds of 4% and 6%.

Species	*C. crus-galli* L.	*C. punctata* Jacq.
Province; county; location; coordinates; site code	Ontario; Regional Municipality of Niagara; Fort George National Historic Park; 43°14′40″N, 79°03′40″W; ON04	Ontario; Perth; Fullarton Twp; 43°18′54″N, 81°10′18″W; ON46
Ploidy level	4*x*	2*x*
Number of families (total progeny size)	18 (83)	26 (118)
Total allele number	24	47
Gene diversity (H_e) averaged over families	0.54 ± 0.15	0.65 ± 0.18
Average *G* per family (at 0%; 2%; 4%; 6% thresholds)	1; 1.12; 1.18; 1.47	2.2; 2.64; 4.04; 4.08
Average *PG* per family (at 0%; 2%; 4%; 6% thresholds)	0.17; 0.19; 0.2; 0.22	0.4; 0.47; 0.73; 0.74
Average *D* per family (at 0%; 2%; 4%; 6% thresholds)	0; 0.04; 0.06; 0.17	0.4; 0.5; 0.81; 0.82

influence estimates of clonal frequency and genetic diversity (Lasso, 2008). For these reasons, we compared diversity measures obtained from four threshold limits when assigning genotypes into clonal lineages. These ranged from the most conserved limit of 0% to the commonly used limits of 2% and 4% (e.g., Arens et al., 1998; Winfield et al., 1998; Van der Hulst et al., 2000; Douhovnikoff and Dodd, 2003) and the limit of 6%, which exceeds the commonly used empirical values. All pairs of embryos below the designated threshold were assigned as clonemates. Once the clones were assigned, the following three parameters were estimated. First, the total number of multilocus genotypes (*G*) was determined. Second, the proportion of distinguishable genotypes (*PG*) was calculated as the number of genotypes divided by sample size (Ellstrand and Roose, 1987). Third, Nei's (1987) genetic diversity corrected for sample size (also known as Simpson's diversity index, *D*) was calculated. This index ranges from zero, where two randomly picked embryos share a single genotype, to one, where embryos are genetically different in a family. Probability values obtained from one-tailed *t*-tests with Bonferroni correction were used to assess whether there is a significant difference between the mean diversity indices of the two taxa under various threshold values.

To increase confidence in the diversity findings, we employed two additional methods to infer genetic relatedness across progeny arrays. One is to construct a tree and use the scatter of embryos within a family as a reflection of intrafamilial dissimilarity. Such a visual display gives a global view of progeny relationships within and between families. The distances that Bruvo et al. (2004) calculated on the basis of a two-phase mutation model for microsatellites and scaled by ploidy level, were used to construct a neighbor-joining tree in PHYLIP, version 3.66 (Felsenstein, 2006), and visualized with TREEVIEW (Page, 1996). The second method is a comparison of Rousset's distances ($\hat{a}$; Rousset, 2000), which were calculated on the basis of allele sharing between all pairs of embryos within a population using SPAGeDI, version 1.2 (Hardy and Vekemans, 2002). This algorithm estimates the probability of a gene being identical by descent between two individuals *i* and *j* with respect to the mean Q_o over all pairs of sampled individuals defined as $\hat{a}_{ij} = (Q_o - Q_{ij})/(1 - Q_o)$ and is an analogue of a kinship coefficient (Rousset, 2000; Vekemans and Hardy, 2004). We compared the relatedness of embryos in three categories—total samples (ALL), between seed families (BS), and within seed family (WS)—for the diploid and tetraploid taxa separately. One-way analysis of variance (ANOVA) with Tukey's HSD test at the 95% confidence limit was performed using XLSTAT, version 2007.4 (Addinsoft, New York, New York, USA), to find whether there was a significant difference in these pairwise $\hat{a}$ values between taxa. Although distances are not completely independent and may not meet the assumptions of an ANOVA, we are making comparisons between taxon samples that are independent, in order to demonstrate the degree of differences as if each distance is an independent observation.

Partition of genetic variation—To determine how genetic diversity is partitioned between and within families, we conducted analysis of molecular variance (AMOVA; Excoffier et al., 1992) with the program GENALEX (Peakall and Smouse, 2006). The codominant SSR data were first converted to a binary data matrix by treating an allele's absence as "0" and its presence as "1." The binary data were then summarized using Jaccard's coefficient (JC), which does not consider the shared absence of an allele as similarity (Legendre and Legendre, 1983). In this way, we obtained an unbiased estimate of pairwise genetic distances between embryos. AMOVA analyses were then performed on the JC matrix with significance tests for 10 000 permutations.

Estimation of selfing and apomixis rates—To examine the reproductive system of the two species in a manner that reflects the genetic variation of their seed families, we estimated multilocus (t_m) and single-locus (t_s) outcrossing rates, biparental inbreeding ($t_m - t_s$, due to matings between closely related individuals), selfing rate (*s*), and the rate of apomixis ("*a*") from the progeny arrays. For diploid *C. punctata*, the program MLTR (Ritland, 2002) was used to estimate t_m and t_s. For tetraploid *C. crus-galli,* these parameters plus "*a*" were estimated with a maximum-likelihood procedure using a modification of the MLTET program for tetraploids (Murawski et al., 1994), wherein "*a*" was additionally estimated. Standard deviations of the parameters were obtained from 1000 bootstraps with families as resampling units. The modified algorithm for jointly estimating rates of apomixis and selfing, and a brief discussion of its properties, are given in Appendix 1.

RESULTS

Marker independence—Linkage disequilibrium with respect to SSR loci in 2*x C. punctata* was shown not to be significant

TABLE 2. Nucleotide sequences and information on microsatellite markers used in this study. We developed primers were based on *Malus* × *domestica* and showed that they are applicable to *Crataegus* species (Liebhard et al., 2002). *C* = total number of alleles detected in *C. crus-galli*, *P* = total number of alleles detected in *C. punctata*, and LG = linkage group. Number in parentheses indicates unique alleles detected in each species.

Locus	Forward primer sequence (5′–3′)	Reverse primer sequence (5′–3′)	Total alleles	*C*	*P*	Size (bp)	Maps on LG no.
CH01F02	ACC ACA TTA GAG CAG TTG AGG	CTG GTT TGT TTT CCT CCA GC	16	10 (4)	13 (6)	145–185	12
CH03A02	TTG TGG ACG TTC TGT GTT GG	CAA GTT CAA CAG CTC AAG ATG A	16	9 (2)	13 (7)	124–168	14
CH04F06	GGC TCA GAG TAC TTG CAG AGG	ATC CTT AAG CGC TCT CCA CA	20	10 (1)	18 (7)	128–190	14
CH04G04	AGT GGA TGA TGA GGA TGA GG	GCT AGT TGC ACC AAG TTC ACA	7	3 (0)	7 (3)	145–159	12
CH05D04	ACT TGT GAG CCG TGA GAG GT	TCC GAA GGT ATG CTT CGA TT	12	5 (2)	11 (6)	166–196	12

($P > 0.05$; after the Bonferroni correction) for the 10 possible pairwise combinations. Although the five loci are distributed over only two linkage groups in *Malus* × *domestica* (Table 2; Liebhard et al., 2002), apparently they are located at a sufficient distance from each other in *C. punctata* that they are unlinked.

Diversity measures and threshold setting—Although the total number of alleles detected in all embryos of 2*x C. punctata* (Table 1) is about two-fold that found in 4*x C. crus-galli*, gene diversity (H_e) of *C. punctata* is not significantly greater than that of *C. crus-galli* ($t = 0.69$, df = 42, $P = 0.06$; Table 1). The average multilocus diversity indices *G*, *PG*, and *D* of the 18 *C. crus-galli* seed families are significantly lower than the average estimates of the 26 *C. punctata* seed families, regardless of which threshold level was set (***G***: $t = 5.24$, df = 42, $P < 0.01$; ***PG***: $t = 5.9$, df = 42, $P < 0.01$; ***D***: $t = 6.04$, df = 42, $P < 0.01$; Table 1). In *C. crus-galli*, the three diversity indices did not indicate significant increase with the rise in thresholds (***G***: $t = 0.76$, df = 34, $P = 0.45$; ***PG***: $t = 1.7$, df = 34, $P = 0.09$; ***D***: $t = 0.78$, df = 34, $P = 0.43$). However, in *C. punctata*, these indices increase significantly by more or less the same magnitude when the threshold values increase from 0 to 4% (***G***: $t = 5.36$, df = 50, $P < 0.003$; ***PG***: $t = 8.54$, df = 50, $P < 0.003$; ***D***: $t = 5.35$, df = 50, $P < 0.003$). No significant differences are found between estimates at the thresholds of 4% and 6% (***G***: $t = 0.1$, df = 50, $P = 0.9$; ***PG***: $t = 0.54$, df = 50, $P = 0.3$; ***D***: $t = 0.26$, df = 50, $P = 0.79$).

Genetic relatedness and partitioning of genetic variation among seed progeny—Seed samples of *C. crus-galli* and *C. punctata* are distinguished as two separate clusters in the neighbor-joining dendrogram (Fig. 1). These clusters differ substantially from each other with respect to topology. Embryos of *C. punctata* are resolved into several small clusters, and members of the same family are scattered in different clusters (Fig. 1A). By contrast, *C. crus-galli* embryos from a given family are likely to be found in the same cluster (Fig. 1B).

ANOVA results indicate that the pairwise Rousset's distances ($\hat{a}$) among all *C. punctata* embryos (ALL; Fig. 2) are significantly greater than those among all of the *C. crus-galli* embryos (ALL; $F = 819.43$, df = 1 and 789, $P < 0.001$; Fig. 2). Negative $\hat{a}$ values observed in *C. crus-galli* indicate a higher interindividual gene identity compared with the overall sampled individual gene identity (Rousset, 2000). Partitioning the distances between embryos into within-seed-family (WS) and between-seed-family (BS) components indicates that these components are similar within species but differ significantly between the two species (WS: $F = 177.09$, df = 1 and 370, $P < 0.001$; BS: $F = 235.11$, df = 1 and 370, $P < 0.001$; Fig. 2). These results are consistent with AMOVA results that indicated that ≤79% of the variation is partitioned within families of *C. punctata*, but only 31% within families of *C. crus-galli*. The amount of between-family variation in *C. punctata* (21%) is about one third that in *C. crus-galli* (69%).

Breeding system—Multilocus (t_m) and average single-locus (t_s) estimates of the outcrossing rate in *C. punctata* are high (0.96 and 0.84; Table 3), which suggests predominant outcrossing and, thus, frequent pollen flow between parent trees. Moreover, the estimates of the rates of biparental inbreeding and selfing are small (0.12 and 0.04; Table 3), revealing little or no evidence of inbreeding. The rate of apomixis is not estimated for *C. punctata* because such estimation is not implemented in the program MLTR for diploids. Nevertheless, apomixis is seemingly insignificant (plus selfing < 0.04) given the high outcrossing rate (0.96) among their embryos. This is not surprising in the Rosaceae, or at least in *Crataegus*, because apomixis is found almost only in polyploids and not in diploids (Dickinson et al., 2007; Talent and Dickinson, 2007a). On the other hand, the MLTET results indicate that the rates of both outcrossing and selfing in the samples of *C. crus-galli* were relatively low ($t_m = 0.3$; $s = 0.17$; Table 3) compared with that of apomixis ($a = 0.47$). These results evidently indicate a preponderance of apomixis and less frequent outcrossing among *C. crus-galli* parent trees. Such a reproductive system is consistent with the lower genetic variation detected within seed families (Table 1).

Per-fruit seed set—In *C. crus-galli*, although 118 of 192 fruits contained one seed and 19 contained two seeds, 55 of the fruits were parthenocarpic (i.e., their pyrenes contained no mature seeds). In the case of *C. punctata*, however, 52 of the 264 fruits contained one seed and 184 contained two or more seeds. Only 28 of the fruits were parthenocarpic. The line representing an average of 2.5 seeds per fruit is an approximate limit for the *C. punctata* fecundity data (Fig. 3). For *C. crus-galli*, however, the corresponding limit represents an average of one seed per fruit (Fig. 3). These results reflect the difference in floral architecture between the two species described above. The maximum per-fruit seed set is two seeds per fruit of *C. crus-galli* (mode = 1) and five seeds per fruit of *C. punctata* (mode = 2 to 3). When the difference in floral architecture between these two species is taken into account, both species can be seen to be converting nearly 50% of their gynoecial units into seed-containing pyrenes under conditions of open pollination.

DISCUSSION

At the level of seed families, we have shown that in hawthorns the consequence of asexual reproduction is reduced genotypic diversity that parallels the reduced morphological variability observed earlier (Dickinson and Phipps, 1985; Dickinson, 1986). Although genotypic diversity is lower in the embryos of the apomictic tetraploid than in the sexual diploid, our data suggest that individuals of the two species are equally fecund on a per-fruit basis.

Genotypic diversity and gene flow via pollen—In the seed families included in our comparisons, genetic diversity differs significantly between *C. punctata* and *C. crus-galli* (Table 1). We attribute this difference to the frequency of outcrossing in *C. punctata* on the one hand, and to that of apomixis and selfing in *C. crus-galli* on the other (Table 3). The greater diversity of the *C. punctata* embryos is unlikely to be influenced by introgression from other species, because this species was the dominant hawthorn species at site ON46 (Table 1). The only other hawthorn found at this site, *C. calpodendron*, flowers a minimum of 1 week after the sympatric *C. punctata* (Phipps and Muniyamma, 1980). In this way, we confirm the predominance of sexual outcrossing in this species, deduced earlier on the basis of embryology and self-incompatibility (Muniyamma and Phipps, 1985; Dickinson and Phipps, 1986). Our results also corroborate those of Talent and Dickinson (2007b), which were based on flow-cytometric analyses of individual seeds.

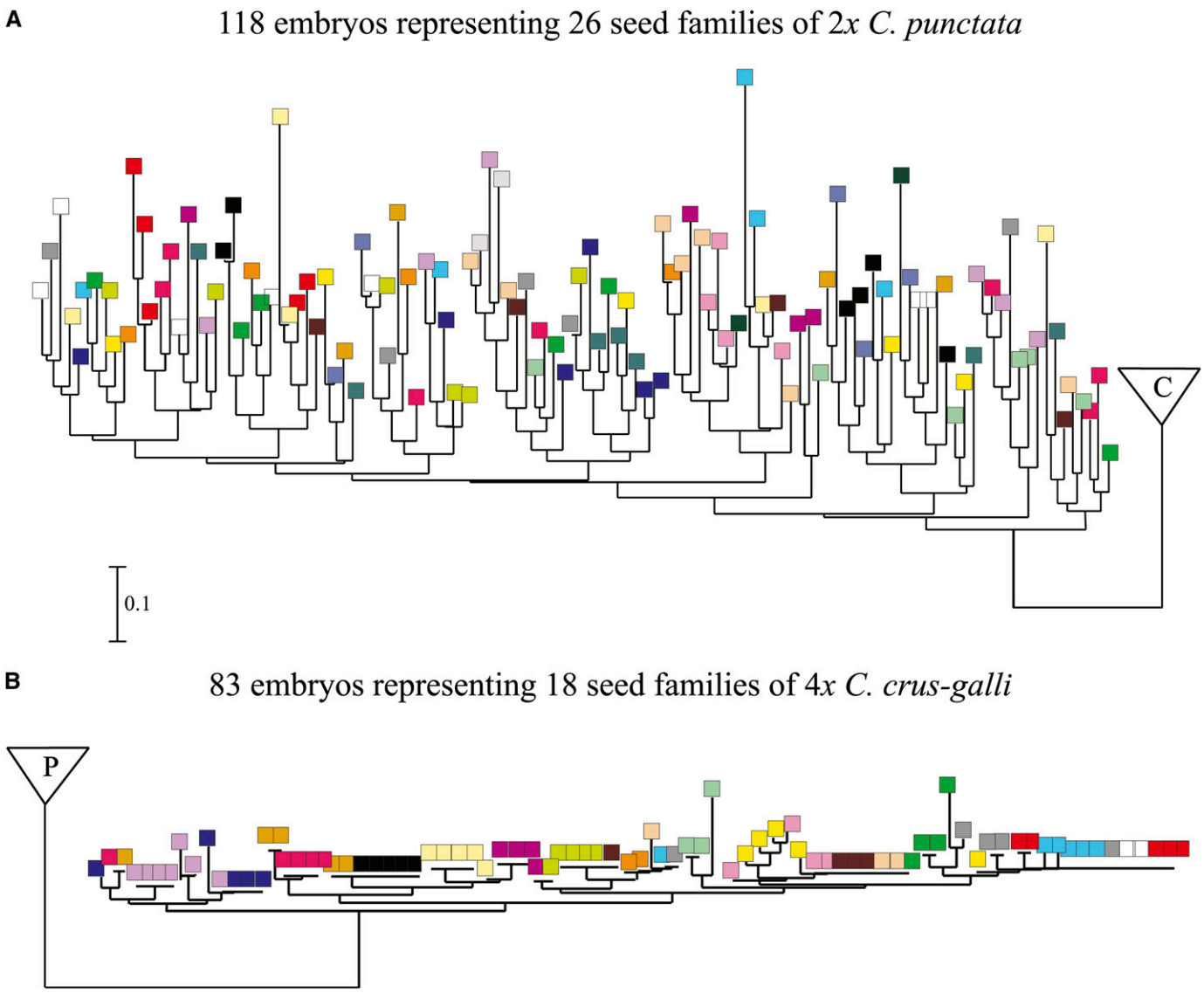

Fig. 1. Bruvo et al.'s (2004) distances for microsatellite data, showing the genetic relatedness among embryos of (A) *Crataegus punctata* and (B) *C. crus-galli* in the form of a neighbor-joining tree. Each embryo is represented by a square, and all those belonging to the same family are the same color. Branches are shown with the same scale for the two taxa. P = *C. punctata*, P = *C. crus-galli*.

Because members of a seed family are expected to share the same maternal genotype, the high genotypic diversity within families of *C. punctata* (Table 1) reflects the self-incompatibility seen in this diploid species, whereas the high genotypic identity within families of *C. crus-galli* (Table 1) reflects the predominance of apomixis and selfing in the polyploid species. These results support earlier inferences from embryology and pollination experiments (Dickinson and Phipps, 1986) and corroborate results from flow cytometry (Talent and Dickinson, 2007b). The slightly greater between-seed-family component of genetic variation detected in *C. crus-galli* (Fig. 2) and the t_m estimate of 0.3 in relation to the t_s of 0.17 (Table 3) suggest that about 36% of pollen flow was within, rather than between, individuals. These findings collectively point to a lack of pollen flow between conspecific individuals. Broadly, scant pollen flow could be attributed to factors such as pollen limitation, a lack of suitable pollinators in the colonized region (Rambuda and Johnson, 2004; Field et al., 2005), or a small founder population (e.g., Widén and Widén, 1990; Goodell et al., 1997; Scobie and Wilcock, 2009). In *C. crus-galli*, earlier findings that include data from the same local population as well as the others suggest that seed set in this species is limited by the availability of compatible pollen, regardless of whether the fates of open-pollinated flowers are compared with those of geitonogamously or xenogamously pollinated ones (Dickinson and Phipps, 1986). Also, on a per-flower basis, seed set was higher in open-pollinated flowers of *C. punctata* than in those of *C. crus-galli* (Dickinson and Phipps, 1986). Taken together, these results suggest that the differences in genetic variability between *C. crus-galli* and *C. punctata* are likely attributable to the contrast in compatibility relationships between them, combined with the occurrence of gametophytic apomixis in *C. crus-galli* but not *C. punctata*. In most of these respects, *C. crus-galli* resembles apomictic species in other genera such as *Rubus* (Kraft

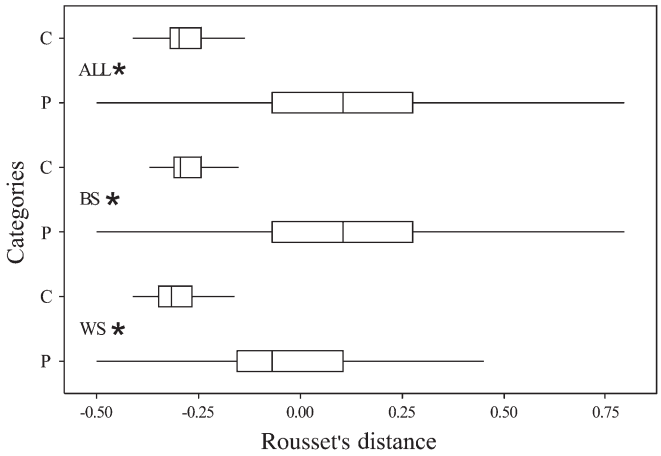

Fig. 2. Box plots of pairwise Rousset's distance ($\hat{a}$; Rousset, 2000) calculated among all embryos (ALL), between seed families (BS), and within seed families (WS) from *Crataegus crus-galli* (C) and *C. punctata* (P). Asterisks indicate significant differences ($P < 0.001$) between species for each category based on the ANOVA comparisons.

et al., 1996; Nybom, 1998), *Aronia* (Persson Hovmalm et al., 2004), *Hieracium* (Shi et al., 1996; Štorchová et al., 2002), *Ranunculus* (Hörandl et al., 2001; Paun et al., 2006), and *Taraxacum* (Lyman and Ellstrand, 1998). These examples also exhibit limited genetic variability in local populations because of selfing or an almost complete lack of sexuality.

Breeding system, floral structure, and fecundity in *Crataegus*—Trees in both the *C. crus-galli* and the *C. punctata* populations produce fruits in which upwards of half the gynoecial units present in the flowers are converted into pyrenes that contain filled seeds (Fig. 3). Thus, at the per-fruit level, there is no indication of inbreeding depression associated with apomixis and self-compatibility having reduced seed set in *C. crus-galli*. Rather, we interpret our results in general as demonstrating the capacity for apomixis and self-fertility to provide reproductive assurance for individual *C. crus-galli* surrounded apparently only by their close relatives. Reproductive assurance, in turn, may explain the switch to a reduced stamen number (10-stamen flowers in tetraploids vs. 20-stamen flowers in diploids) and ovule number that are seen in most North American tetraploid hawthorns (Dickinson et al., 1996). Such a phenomenon warrants further investigations in the other species.

In diploids, self-incompatibility may have selected for increased ovule number per flower and, hence, increased numbers of gynoecial units per flower, as an adaptation to the unpredictable transfer of compatible pollen to stigmas (Burd et al., 2009). In tetraploids like *C. crus-galli*, self-compatibility could compensate for reduced numbers of gynoecial units per flower as well as for reductions in stamen number per flower if optimal pollen transfer to stigmas depends on limiting the amount of pollen available to pollinators at any one time. This would depend on tetraploids, by virtue of their self-compatibility, being effectively less spatially dispersed than self-sterile diploids (Harder and Thomson, 1989; Dickinson et al., 1996).

Some tests of the suggestion that sexual individuals should have twice the costs of meiosis (Maynard Smith, 1978) and lower total progeny production than asexual individuals have failed to demonstrate such a result (e.g., Kearney and Shine, 2005; Houliston et al., 2006). Similarly, there are empirical data that support the idea that, compared with sexual reproduction, apomictic reproduction may result in lower fecundity (e.g., Voigt et al., 2007; Hörandl, 2008). Nevertheless, our findings provide only limited evidence of such effects. Instead, self-compatibility (as distinct from selfing) is an integral component of reproductive success in *C. crus-galli*, in that it enables seed-set, given pseudogamous apomixis (Noirot et al., 1997). The occurrence of diploid *C. crus-galli* well to the south of the Pleistocene glacial margin (Talent and Dickinson, 2005) suggests that the tetraploids could be as much as 10 000 yr old and so have had abundant time in which to purge deleterious alleles. The postglacial migration northward would have likewise provided ample opportunity to refine reproductive and colonizing abilities.

Comparisons between the tetraploid *C. crus-galli* and diploid *C. punctata* studied here presently lack a fully resolved phylogenetic context. Both species are strongly supported as belonging to the same eastern North American clade on the basis of phylogenetic analysis of four chloroplast intergenic spacers and five nuclear loci (clade B; Lo et al., 2009a). However, the exact relationships within this morphologically diverse group remain poorly resolved. Although this may render some of the commentary above more speculative, the lack of precise phylogenetic context does not affect our conclusions about the extent to which apomixis and self-compatibility reduce the genetic variability of seed families in *C. crus-galli* as compared with those of a sexual diploid when both taxa are sampled in southern Ontario.

***Caveats for estimating genetic variation in polyploids with codominant markers*—**As described in the Appendix, maximum-likelihood methods can be developed with which to estimate mating-system parameters in polyploids (see also Murawski et al., 1994; Thompson and Ritland, 2006). Alternatively, if the data are summarized by means of an appropriate resemblance function, there are descriptive multivariate methods that can be employed to depict the pattern of variation present in the sample. Comparisons of genetic variation across ploidy levels using existing methods have been problematic because of the different number of alleles involved and the underlying assumption of disomic inheritance. Thus, microsatellite loci are usually treated as dominant markers that are either present ("1") or absent ("0") in many studies of polyploids (e.g., Cai et al., 2005; Brown et al., 2007; Kloda et al., 2008; Andreakis

TABLE 3. Estimates of outcrossing, selfing, and apomixis rates among *Crataegus punctata* and *C. crus-galli* progenies, including both multilocus and single-locus estimates from microsatellite data (loci as given in Table 2). Standard deviations are obtained from 1000 boostraps with families as resampling units.

	C. punctata	*C. crus-galli*
Multilocus outcrossing rate (t_m)	0.96 ± 0.03	0.3 ± 0.051
Average single locus outcrossing rate (t_s)	0.85 ± 0.035	0.18 ± 0.053
CH01F02-based outcrossing rate	0.67 ± 0.076	0.13 ± 0.054
CH03A02-based outcrossing rate	0.90 ± 0.094	0.25 ± 0.067
CH04F06-based outcrossing rate	0.69 ± 0.069	0.16 ± 0.064
CH04G04-based outcrossing rate	0.71 ± 0.13	0.31 ± 0.016
CH05D04-based outcrossing rate	0.96 ± 0.075	0.03 ± 0.02
Biparental inbreeding ($t_m - t_s$)	0.12 ± 0.035	0.11 ± 0.028
Rate of selfing (s)	0.036	0.17
Rate of apomixis (a)	—	0.47 ± 0.099
Fixation coefficient (F)	0.07 ± 0.061	0.3 ± 0.006

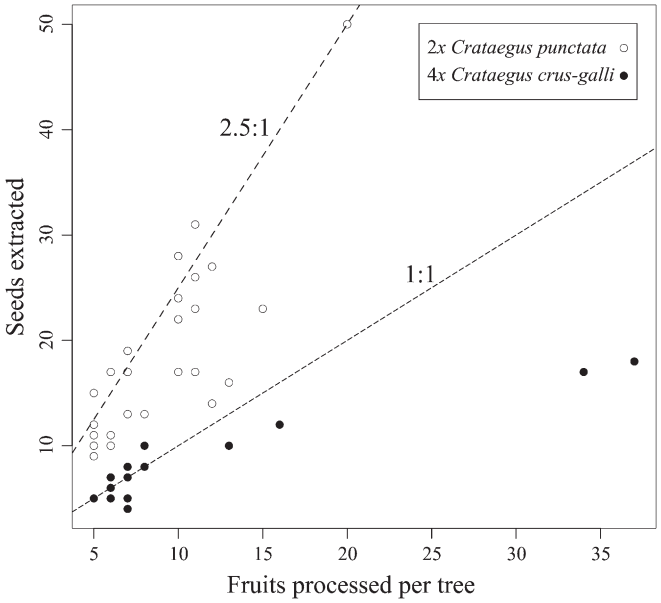

Fig. 3. Scatterplot showing the number of fully developed seeds (i.e., embryo covered with endosperm adhering to the brown seed coat) extracted from samples of mature fruits of *Crataegus crus-galli* and *C. punctata*. Each dot represents an individual tree from which a sample of *x* fruits was examined so as to obtain *y* fully developed seeds. Dashed lines depict 1:1 and 2.5:1 ratios of examined seeds to fruits.

et al., 2009; Garcia-Verdugo et al., 2009). Converting genotypes into binary data can lose valuable information concerning the manner in which loci are inherited and the composition of individual genotypes. Nevertheless, multivariate analysis of converted data appears to be useful in revealing genetic variation among species of different ploidy levels (e.g., Cai et al., 2005; Karlin et al., 2008; Kloda et al., 2008) but may be less suitable for population samples, as in the current study. Here, we have employed three methods to compare genetic variation between diploid and polyploid samples that take advantage of the codominant property of microsatellite markers. First, we use the stepwise mutation model to estimate genotypic diversity indices based on single or multiple loci in GENODIVE (Meirmans and van Tienderen, 2004). Lasso (2008) has discussed the sensitivity of diversity estimates to threshold setting, which could vary among species. Although most studies have applied threshold limits of 2–4%, we recommend that researchers estimate and compare indices across a range of thresholds, as we did here (see Materials and Methods; Table 1) to examine the level of differences. Second, we used distance functions to build phylogenetic trees for visualizing relationships and to evaluate the degree of dispersion among samples (Fig. 1; Weller and Sakai, 1999; Persson Hovmalm et al., 2004). Here, caution is needed in choosing the distance function. Polyploids with more alleles than diploids may appear to be closer together as an artifact of higher allele numbers and, thus, a higher proportion of shared alleles. Bruvo et al. (2004) suggested distance measures that are not only scaled by ploidy level but also take the mutation processes of microsatellites into account. This provides a fair comparison of relatedness between ploidy levels. Third, we compared the global and partitioned Rousset (2000) distances calculated, under the assumption that shared alleles are identical by descent between two samples (Fig. 2). Unlike the preceding methods, this approach compares variation at different hierarchical levels, including between populations, between seed families of a population, and between members of a seed family, to obtain statistical support for such comparisons.

Conclusions—Our results are pertinent for the systematics of long-lived apomicts in general, and for *Crataegus* systematics in particular. They lend molecular support to previous inferences based on analyses of morphometric data (e.g., Dickinson and Phipps, 1985; Dickinson, 1986) about the way in which limited variation within populations of apomicts, and substantially greater variation between their populations, has contributed to the recognition of species that are only slightly differentiated genetically (Dickinson et al., 2008). In addition, we suggest ways in which increased asexual reproduction and selfing in polyploids may be related to floral evolution in these groups. Whether changes in reproductive strategy affect genetic variation and fitness of a species is an important question whose answer may determine the evolutionary potential of a species. Our study demonstrates that agamospermous plants can differ from sexual ones in their levels of genetic variation, because of the differential influences of pollen flow and sexuality, and can achieve short-term reproductive assurance and reproduce as successfully as sexual ones. However, their long-term evolutionary fate remains in question. Similar fine-scale studies on other groups are worthwhile for testing whether such phenomena operate widely in apomictic angiosperms or are restricted to groups like *Crataegus* in which the reproductive success of apomicts is intimately linked with self-incompatibility and polyploidy.

LITERATURE CITED

ANDREAKIS, N., W. H. C. F. KOOISTRA, AND G. PROCACCINI. 2009. High genetic diversity and connectivity in the polyploid invasive seaweed *Asparagopsis taxiformis* (Bonnemaisoniales) in the Mediterranean, explored with microsatellite alleles and multilocus genotypes. *Molecular Ecology* 18: 212–226.

ARENS, P., H. COOPS, J. JANSEN, AND B. VOSMAN. 1998. Molecular genetic analysis of black poplar (*Populus nigra* L.) along Dutch rivers. *Molecular Ecology* 7: 11–18.

ASHMAN, T. L., T. M. KNIGHT, J. A. STEETS, P. AMARASEKARE, M. BURD, D. R. CAMPBELL, M. R. DUDASH, ET AL. 2004. Pollen limitation of plant reproduction: ecological and evolutionary causes and consequences. *Ecology* 85: 2408–2421.

BAYER, R. J., K. RITLAND, AND B. G. PURDY. 1990. Evidence of partial apomixis in *Antennaria media* (Asteraceae: Inuleae) detected by the segregation of genetic markers. *American Journal of Botany* 77: 1078–1083.

BROWN, J. S., R. J. SCHNELL, E. J. POWER, S. L. DOULGAS, AND D. N. KUHN. 2007. Analysis of clonal germplasm from five *Saccharum* species: *S. barberi*, *S. robustum*, *S. officinarum*, *S. sinense* and *S. spontaneum*. A study of inter- and intra species relationships using microsatellite markers. *Plant Genetic Resources* 7: 127–134.

BRUVO, R., N. K. MICHIELS, T. G. D'SOUZA, AND H. SCHULENBURG. 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology* 13: 2101–2106.

BURD, M., T. L. ASHMAN, D. R. CAMPBELL, M. R. DUDASH, M. O. JOHNSTON, T. M. KNIGHT, S. J. MAZER, ET AL. 2009. Ovule number per flower in a world of unpredictable pollination. *American Journal of Botany* 96: 1159–1167.

CAI, A. Q., K. S. AITKENB, Y. H. FANA, G. PIPERIDISC, P. JACKSOND, AND C. L. MCINTYRE. 2005. A preliminary assessment of the genetic relationship between *Erianthus rockii* and the *Saccharum* complex using microsatellite (SSR) and AFLP markers. *Plant Science* 169: 976–984.

CARINO, D. A., AND C. C. DAEHLER. 1999. Genetic variation in an apomictic grass, *Heteropogon contortus*, in the Hawaiian Islands. *Molecular Ecology* 8: 2127–2132.

CLAUSEN, J. 1954. Partial apomixis as an equilibrium system in evolution. *Carylogia* 6: 469–479.

DARLINGTON, C. D. 1958. Evolution of genetic systems. Oliver and Boyd, Edinburgh, UK.

DICKINSON, T. A. 1983. *Crataegus crus-galli* L. sensu lato in southern Ontario: phenotypic variation and variability in relation to reproductive behavior. Ph.D. dissertation, Department of Plant Sciences, University of Western Ontario, London, Ontario, Canada.

DICKINSON, T. A. 1986. Topodeme differentiation in Ontario taxa of *Crataegus* (Rosaceae: Maloideae): leaf morphometric evidence. *Canadian Journal of Botany* 64: 2738–2747.

DICKINSON, T. A. 1999. Species concepts in agamic complexes. *In* L. W. D. van Raamsdonk and J. C. M. Den Nijs [eds.], Evolution in man-made habitats, 319–339. Institute for Systematics and Ecology, Amsterdam, The Netherlands.

DICKINSON, T. A., S. BELAOUSSOFF, R. M. LOVE, AND M. MUNIYAMMA. 1996. North American black-fruited hawthorns: I. Variation in floral construction, breeding system correlates, and their possible evolutionary significance in *Crataegus* sect. *Douglasii* Loudon. *Folia Geobotanica et Phytotaxonomica* 31: 355–371.

DICKINSON, T. A., E. Y. Y. LO, AND N. TALENT. 2007. Polyploidy, reproductive biology, and Rosaceae: understanding evolution and making classification. *Plant Systematics and Evolution* 266: 59–78.

DICKINSON, T. A., E. Y. Y. LO, N. TALENT, AND R. M. LOVE. 2008. Black-fruited hawthorns of Western North America—one or more agamic complexes? *Botany* 86: 846–865.

DICKINSON, T. A., AND J. B. PHIPPS. 1985. Degree and pattern of variation in *Crataegus* section *Crus-galli* in Ontario. *Systematic Botany* 10: 322–337.

DICKINSON, T. A., AND J. B. PHIPPS. 1986. The breeding system of *Crataegus crus-galli* sensu lato in Ontario (Canada). *American Journal of Botany* 73: 116–130.

DIJK, P. J. VAN 2003. Ecological and evolutionary opportunities of apomixis: insights from *Taraxacum* and *Chondrilla*. *Philosophical Transactions of the Royal Society B* 358:1113 –1121.

DOUHOVNIKOFF, V., AND R. S. DODD. 2003. Intra-clonal variation and a similarity threshold for identification of clones: application to *Salix exigua* using AFLP molecular markers. *Theoretical and Applied Genetics* 106: 1307–1315.

DURAND, J., L. GARNIER, I. DAJOZ, S. MOUSSET, AND M. VEUILLE. 2000. Gene flow in a facultative apomictic Poacea, the savanna grass *Hyparrhenia diplandra*. *Genetics* 156: 823–831.

ELLSTRAND, N. C., AND M. L. ROOSE. 1987. Patterns of genotypic diversity in clonal plant species. *American Journal of Botany* 74: 123–131.

ESSELINK, G. D., H. NYBOM, AND B. VOSMAN. 2004. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting-peak ratios) method. *Theoretical and Applied Genetics* 109: 402–408.

ESSELMAN, E. J., Q. L. JIAN, D. J. CRAWFORD, J. L. WINDUS, AND A. D. WOLFE. 1999. Clonal diversity in the rare *Calamagrostis porteri* ssp. *insperata* (Poaceae): comparative results for allozymes and random amplified polymorphic DNA (RAPD) and intersimple sequence repeat (ISSR) markers. *Molecular Ecology* 8: 443–451.

EXCOFFIER, L., P. E. SMOUSE, AND J. QUATTRO. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.

FELSENSTEIN, J. 2006. PHYLIP (Phylogeny Inference Package), version 3.66c. Zoology Department, University of Washington, Seattle, Washington, USA.

FIELD, D. L., D. J. AYRE, AND R. J. WHELAN. 2005. The effect of local plant density on pollinator behavior and the breeding system of *Persoonia bargoensis* (Proteaceae). *International Journal of Plant Sciences* 166: 969–977.

FU, Y. X., AND R. CHAKRABORTY. 1998. Simultaneous estimation of all parameters of a stepwise mutation model. *Genetics* 150: 487–497.

GARCIA-VERDUGO, C., M. F. FAY, C. GRANADO-YELA, R. R. DE CASAS, L. BALAGUER, G. BESNARD, AND P. VARGAS. 2009. Genetic diversity and differentiation processes in the ploidy series of *Olea europaea* L.: a multiscale approach from subspecies to insular populations. *Molecular Ecology* 18: 454–467.

GARNIER, L. K. M., J. DURAND, AND I. DAJOZ. 2002. Limited seed dispersal and microspatial population structure of an agamospermous grass of West African savannahs, *Hyparrhenia diplandra* (Poaceae). *American Journal of Botany* 89: 1785–1791.

GOODELL, K., D. R. ELAM, AND J. D. NASON. 1997. Gene flow among small populations of a self-incompatible plant: an interaction between demography and genetics. *American Journal of Botany* 84: 1362–1371.

GRANT, V. 1981. Plant speciation. Columbia University Press, New York, New York, USA.

GUSTAFSSON, Å. 1942. The origin and properties of the European blackberry flora. *Hereditas* 28: 249–277.

HARDER, L. D., AND J. D. THOMSON. 1989. Evolutionary options for maximizing pollen dispersal of animal-pollinated plants. *American Naturalist* 133: 323–344.

HARDY, O. J., N. CHARBONNEL, H. FREVILLE, AND M. HEUERTZ. 2003. Microsatellite allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics* 163: 1467–1482.

HARDY, O. J., AND X. VEKEMANS. 2002. SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2: 618–620.

HÖRANDL, E. 2008. Evolutionary implications of self-compatibility and reproductive fitness in the apomictic *Ranunculus auricomus* polyploid complex (Ranunculaceae). *International Journal of Plant Sciences* 169: 1219–1228.

HÖRANDL, E., G. JAKUBOWSKY, AND C. DOBEŠ. 2001. Isozyme and morphological diversity within apomictic and sexual taxa of the *Ranunculus auricomus* complex. *Plant Systematics and Evolution* 226: 165–185.

HÖRANDL, E., AND O. PAUN. 2007. Patterns and sources of genetic diversity in apomictic plants: implications for evolutionary potentials. *In* E. Hörandl, U. Grossniklaus, P. J. Van Dijk, and T. F. Sharbel [eds.], Apomixis: evolution, mechanisms and perspectives, 169–194. ARG-Gantner, Ruggell, Liechtenstein.

HOULISTON, G. J., H. M. CHAPMAN, AND R. A. BICKNELL. 2006. The influence of genotype and environment on the fecundity and facultative expression of apomixis in *Hieracium pilosella*. *Folia Geobotanica et Phytotaxonomica* 41: 165–181.

KARLIN, E. F., S. B. BOLES, AND A. J. SHA. 2008. Resolving boundaries between species in *Sphagnum* section *Subsecunda* using microsatellite markers. *Taxon* 57: 1189–1200.

KEARNEY, M., AND R. SHINE. 2005. Lower fecundity in parthenogenetic geckos than sexual relatives in the Australian arid zone. *Journal of Evolutionary Biology* 18: 609–618.

KLODA, J. M., P. D. G. DEAN, C. MADDREN, D. W. MACDONALD, AND S. MAYES. 2008. Using principle component analysis to compare genetic diversity across polyploidy levels within plant complexes: an example form British Restharrows (*Ononis spinosa* and *Ononis repens*). *Heredity* 100: 253–260.

KONDRASHOV, A. S. 1993. Classification of hypotheses on the advantage of amphimixis. *Journal of Heredity* 84: 372–387.

KRAFT, T., H. NYBOM, AND G. WERLEMARK. 1996. DNA fingerprint variation in some blackberry species (*Rubus* subg.*Rubus*, Rosaceae). *Plant Systematics and Evolution* 199: 93–108.

LASSO, E. 2008. The importance of setting the right genetic distance threshold for identification of clones using amplified fragment length polymorphism: a case study with five species in the tropical plant genus *Piper*. *Molecular Ecology Resources* 8: 74–82.

LEGENDRE, L., AND P. LEGENDRE. 1983. Numerical ecology. Elsevier, Amsterdam, The Netherlands.

LIEBHARD, R., L. GIANFRANCESCHI, B. KOLLER, C. D. RYDER, R. TARCHINI, E. VAN DE WEG, AND C. GESSLER. 2002. Development and characterisation of 140 new microsatellites in apple (*Malus* × *domestica* Borkh). *Molecular Breeding* 10: 217–241.

LO, E. Y. Y., S. STEFANOVIĆ, K. I. CHRISTENSEN, AND T. A. DICKINSON. 2009a. Evidence for genetic association between East Asian and western North American *Crataegus* L. (Rosaceae) and rapid divergence of the eastern North American lineages based on multiple DNA sequences. *Molecular Phylogenetics and Evolution* 51: 157–168.

LO, E. Y. Y., S. STEFANOVIĆ, AND T. A. DICKINSON. 2007. Molecular reappraisal of relationships between *Crataegus* and *Mespilus* (Rosaceae, Pyreae): Two genera or one? *Systematic Botany* 32: 596–616.

LO, E. Y. Y., S. STEFANOVIĆ, AND T. A. DICKINSON. 2009b. Population genetic structure of diploid sexuals and polyploid apomicts of *Crataegus* (hawthorns; Rosaceae) in the Pacific Northwest. *Molecular Ecology* 18: 1145–1160.

LYMAN, J. C., AND N. C. ELLSTRAND. 1998. Relative contribution of breeding system and endemism to genotypic diversity: the outcrossing endemic *Taraxacum californicum* vs. the widespread apomict *T. officinale* (sensu lato). *Madrono* 45: 283–289.

LYNCH, M., AND W. GABRIEL. 1983. Phenotypic evolution and parthenogenesis. *American Naturalist* 122: 745–764.

MABLE, B. K. 2008. Genetic causes and consequences of the breakdown of self-incompatibility: case studies in the Brassicaceae. *Genetical Research* 90: 47–60.

MAYNARD SMITH, J. 1978. The evolution of sex. Cambridge University Press, London, UK.

MEIRMANS, P. G., H. C. M. DEN NIJS, AND P. H. VAN TIENDEREN. 2006. Male sterility in triploid dandelions: asexual females vs asexual hermaphrodites. *Heredity* 96: 45–52.

MEIRMANS, P. G., AND P. H. VAN TIENDEREN. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4: 792–794.

MENKEN, S. B. J., E. SMIT, AND J. C. M. HANS DEN NIJS. 1995. Genetical population structure in plants: gene flow between diploid sexual and triploid asexual dandelions (*Taraxacum* section *Ruderalia*). *Evolution* 49: 1108–1118.

MICHAELS, H. J., AND F. A. BAZZAZ. 1989. Individual and population responses of sexual and apomictic plants to environmental gradients. *American Naturalist* 134: 190–207.

MOGIE, M., N. F. BRITTON, AND J. A. STEWART-COX. 2007. Asexuality, polyploidy and the male function. *In* E. Hörandl, U. Grossniklaus, P. J. Van Dijk,, and T. Sharbel [eds.], Apomixis: evolution, mechanisms and perspectives, 195–214. ARG-Gantner, Ruggell, Liechtenstein.

MORRIS, A., N. E. MORTON, A. COLLINS, J. MACPHERSON, D. NELSON, AND S. SHERMAN. 1995. An n-allele model for progressive amplification in the FMR1 locus. *Proceedings of the National Academy of Sciences, USA* 92: 4833–4837.

MUNIYAMMA, M., AND J. B. PHIPPS. 1979. Cytological proof of apomixis in *Crataegus* (Rosaceae). *American Journal of Botany* 66: 149–155.

MUNIYAMMA, M., AND J. B. PHIPPS. 1984. Further cytological evidence for the occurrence of apomixis in North American hawthorns. *Canadian Journal of Botany* 62: 2316–2324.

MUNIYAMMA, M., AND J. B. PHIPPS. 1985. Cytological evidence for sexuality in some diploid and tetraploid species of North American hawthorns. *Canadian Journal of Botany* 63: 1319–1324.

MURAWSKI, D. A., T. H. FLEMING, K. RITLAND, AND J. L. HAMRICK. 1994. Mating system of *Pachycereus pringlei* an autotetraploid cactus. *Heredity* 72: 86–94.

NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, New York, USA.

NOGLER, G. A. 1984. Gametophytic apomixis. *In* Embryology of angiosperms,475–518. Springer-Verlag, Berlin, Germany.

NOIROT, M., D. COUVET, AND S. HAMON. 1997. Main role of self-pollination rate on reproductive allocations in pseudogamous apomicts. *Theoretical and Applied Genetics* 95: 479–483.

NYBOM, H. 1998. Biometry and DNA fingerprinting detected limited genetic differentiation among populations of the blackberry *Rubus nesssensis* (Rosaceae). *Nordic Journal of Botany* 18: 323–333.

NYBOM, H. 2007. Unique reproduction in dogroses (*Rosa* sect. *Caninae*) maintains successful and highly heterozygous genotypes. *In* E. Hörandl, U. Grossniklaus, P. J. Van Dijk, and T. F. Sharbel [eds.], Apomixis: evolution, mechanisms and perspectives, 281–298. ARG-Gantner, Ruggell, Liechtenstein.

ODDOU-MURATORIO, S., R. J. PETIT, B. LE GUERROUE, D. GUESNET, AND B. DEMESURE. 2001. Pollen- versus seed-mediated gene flow in a scattered forest tree species. *Evolution* 55: 1123–1135.

PAGE, R. D. M. 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 12: 357–358.

PAUN, O., J. GREILHUBER, E. M. TEMSCH, AND E. HÖRANDL. 2006. Patterns, sources and ecological implications of clonal diversity in apomictic *Ranunculus carpaticola* (*Ranunculus auricomus* complex, Ranunculaceae). *Molecular Ecology* 15: 897–910.

PAUN, O., AND E. HÖRANDL. 2006. Evolution of hypervariable microsatellites in apomictic polyploid lineages of *Ranunculus carpaticola*: directional bias at dinucleotide loci. *Genetics* 174: 387–398.

PEAKALL, R., AND P. E. SMOUSE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

PERSSON HOVMALM, H. A., N. JEPPSSON, I. BARTISH, AND H. NYBOM. 2004. RAPD analysis of diploid and tetraploid populations of *Aronia* points to different reproductive strategies within the genus. *Hereditas* 141: 301–312.

PHIPPS, J. B., AND M. MUNIYAMMA. 1980. Studies in *Crataegus* (Rosaceae: Maloideae) III. A taxonomic revision of *Crataegus* (Rosaceae) in Ontario. *Canadian Journal of Botany* 58: 1621–1699.

PHIPPS, J. B., K. R. ROBERTSON, P. G. SMITH, AND J. R. ROHRER. 1990. A checklist of the subfamily Maloideae (Rosaceae). *Canadian Journal of Botany* 68: 2209–2269.

RAMBUDA, T. D., AND S. D. JOHNSON. 2004. Breeding systems of invasive alien plants in South Africa: does Baker's rule apply? *Diversity & Distributions* 10: 409–416.

RAYMOND, A. E., AND F. ROUSSET. 1995. GENEPOP version 3.3: population genetics software for exact test and ecumenism. *Journal of Heredity* 86: 248–249.

RICE, W. R. 1989. Analysing tables of statistical tests. *Evolution* 43: 223–225.

RITLAND, K. 2002. Extensions of models for the estimation of mating systems using *n* independent loci. *Heredity* 88: 221–228.

ROGSTAD, S. H., B. KEANE, AND J. BERESH. 2002. Genetic variation across VNTR loci in central North American *Taraxacum* surveyed at different spatial scales. *Plant Ecology* 161: 111–121.

ROUSSET, F. 2000. Genetic differentiation between individuals. *Journal of Evolutionary Biology* 13: 58–62.

SCOBIE, A. R., AND C. C. WILCOCK. 2009. Limited mate availability decreases reproductive success of fragmented populations of *Linnaea borealis*, a rare, clonal self-incompatible plant. *Annals of Botany* 103: 835–846.

SHI, Y., R. J. GORNALL, J. DRAPER, AND C. A. STACE. 1996. Intraspecific molecular variation in *Hieracium* sect. *Alpina* (Asteraceae), an apomictic group. *Folia Geobotanica et Phytotaxonomica* 31: 305–313.

STEBBINS, G. L. 1950. Variation and evolution in plants. Columbia University Press, New York, New York, USA.

ŠTORCHOVÁ, H., J. CHRTEK JR., I. V. BARTISH, M. TETERA, J. KIRSCHNER, AND J. ŠTEPÁNEK. 2002. Genetic variation in agamospermous taxa of *Hieracium* sect. *Alpina* (Compositae) in the Tatry Mts. (Slovakia). *Plant Systematics and Evolution* 235: 1–17.

TALENT, N. 2009. Evolution of gametophytic apomixis in flowering plants: an alternative model from Maloid Rosaceae. *Theory in Biosciences* 128: 121–138.

TALENT, N., AND T. A. DICKINSON. 2005. Polyploidy in *Crataegus* and *Mespilus* (Rosaceae, Maloideae): evolutionary inferences from flow cytometry of nuclear DNA amounts. *Canadian Journal of Botany* 83: 1268–1304.

TALENT, N., AND T. A. DICKINSON. 2007a. Endosperm formation in aposporous in *Crataegus* L. (Rosaceae, Spiraeoideae, Pyreae): Parallel to Ranunculaceae and Poaceae. *New Phytologist* 173: 231–249.

TALENT, N., AND T. A. DICKINSON. 2007b. Ploidy level increase and decrease in seeds from crosses between sexual diploids and asexual

triploids and tetraploids in *Crataegus* L. (Rosaceae, Spiraeoideae, Pyreae). *Canadian Journal of Botany* 85: 570–584.

THOMPSON, S. L., G. CHOE, K. RITLAND, AND J. WHITTON. 2008. Cryptic sex with male-sterile polyploid populations of the Easter daisy, *Townsendia hookeri*. *International Journal of Plant Sciences* 169: 183–193.

THOMPSON, S., AND K. RITLAND. 2006. A novel mating system analysis for modes of self-oriented mating applied to diploid and polyploid Arctic Easter daisies (*Townsendia hookeri*). *Heredity* 97: 119–126.

THOMPSON, S. L., AND J. WHITTON. 2006. Patterns of recurrent evolution and geographic parthenogenesis within apomictic polyploid Easter daisies (*Townsendia hookeri*). *Molecular Ecology* 15: 3389–3400.

VAN DER HULST, R. G. M., T. H. M. MES, J. C. M. DEN NIJS, AND K. BACHMANN. 2000. Amplified fragment length polymorphism (AFLP) markers reveal that population structure of triploid dandelions (*Taraxacum officinale*) exhibits both clonality and recombination. *Molecular Ecology* 9: 1–8.

VEKEMANS, X., AND O. J. HARDY. 2004. New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* 13: 921–935.

VOIGT, M. L., M. MELZER, T. RUTTEN, T. MITCHELL-OLDS, AND T. F. SHARBEL. 2007. Gametogenesis in the apomictic *Boechera holboellii* complex: the male perspective. *In* E. Hörandl, U. Grossniklaus, P. J. Van Dijk, and T. Sharbel [eds.], Apomixis: evolution, mechanisms and perspectives, 235–258. ARG-Gantner, Ruggell, Liechtenstein.

WARD, D. B. 1974. The "ignorant man" technique of sampling plant populations. *Taxon* 23: 325–330.

WELLER, S. G., AND A. K. SAKAI. 1999. Using phylogenetic approaches for the analysis of plant breeding system evolution. *Annual Review of Ecology and Systematics* 30: 167–199.

WHITTAKER, J. C., R. M. HARBORD, N. BOXALL, I. MACKAY, G. DAWSON, AND R. M. SIBLY. 2003. Likelihood-based estimation of microsatellite mutation rates. *Genetics* 164: 781–787.

WHITTON, J., C. J. SEARS, E. J. BACCK, AND S. P. OTTO. 2008. The dynamic nature of apomixis in the angiosperms. *International Journal of Plant Sciences* 169: 169–182.

WIDÉN, B., AND M. WIDÉN. 1990. Pollen limitation and distance-dependent fecundity in females of the clonal gynodioecious herb *Glechoma hederacea* (Lamiaceae). *Oecologia* 83: 191–196.

WINFIELD, M. O., G. M. ARNOLD, AND F. COOPER. 1998. A study of genetic diversity in *Populus nigra* subsp. *betulifolia* in the Upper Severn area of the UK using AFLP markers. *Molecular Ecology* 7: 3–10.

APPENDIX 1.

A. Joint estimation of apomixis and selfing for using multiallelic loci in tetraploids

At multiallelic tetraploid loci, there are a large number of possible genotypes, and listing each progeny probability is impractical. Following the appendix of Murawski et al. (1994), we use a general formula for progeny, conditioned on parentage and type of mating (outcross, self, apomixis). We denote the maternal parent as $A_iA_jA_kA_l$ and its offspring as $A_mA_nA_oA_p$. The subscripts can take on any integer value, for example $A_1A_1A_2A_3$, in which case $i = j = 1$, $k = 2$, and $l = 3$. However, for the purpose of incorporating apomixis, the alleles are ranked in numerical order in both parent and progeny data.

Define the indicator variable δ_{im} where $\delta_{im} = 1$ if $i = m$; otherwise $\delta_{im} = 0$. These pairwise δ values are defined for all ways that a pair of alleles are sampled, one from the parent and the other from the progeny, and indicate the identity-by-state of alleles. Likewise, for triplets of alleles, let $\delta_{mno} = 1$ if $m = n = o$; $\delta_{mno} = 0$ otherwise, etc., and for quadruplets of alleles, let $\delta_{mnop} = 1$ if $m = n = o = p$; otherwise $\delta_{mnop} = 0$, etc. The probability of observing $A_mA_nA_oA_p$, given its parentage and that it was a self, is

$$\text{prob}(A_mA_nA_oA_p \mid A_iA_jA_kA_l, \text{self}) = \left(\frac{c}{3}\right)\left(\left(\frac{\Delta_{mn}}{12}\right)\left(\frac{\Delta_{op}}{12}\right) + \left(\frac{\Delta_{mo}}{12}\right)\left(\frac{\Delta_{np}}{12}\right) + \left(\frac{\Delta_{mp}}{12}\right)\left(\frac{\Delta_{no}}{12}\right)\right)$$

where

$$c = 24 - 12\left(\delta_{mn}+\delta_{mo}+\delta_{mp}+\delta_{no}+\delta_{np}+\delta_{op}\right) + 6\left(\delta_{mn}\delta_{op}+\delta_{mo}\delta_{np}+\delta_{mp}\delta_{no}\right) + 16\left(\delta_{mno}+\delta_{mnp}+\delta_{mop}+\delta_{nop}\right) - 33\delta_{mnop}$$

and

$$\Delta_{mn} = (\delta_{im}+\delta_{jm}+\delta_{km}+\delta_{lm})(\delta_{in}+\delta_{jn}+\delta_{kn}+\delta_{ln}) - (\delta_{im}\delta_{in}+\delta_{jm}\delta_{jn}+\delta_{km}\delta_{kn}+\delta_{lm}\delta_{ln}), \quad \text{etc.}$$

The probability of observing $A_mA_nA_oA_p$ given its parentage and that it was outcrossed is

$$\text{prob}(A_mA_nA_oA_p \mid A_iA_jA_kA_l, \text{outcross}) = (c/6)[(\tfrac{\Delta_{mn}}{12})p_{op} + (\tfrac{\Delta_{mo}}{12})p_{np} + (\tfrac{\Delta_{mp}}{12})p_{no} + (\tfrac{\Delta_{no}}{12})p_{mp} + (\tfrac{\Delta_{np}}{12})p_{mo} + (\tfrac{\Delta_{op}}{12})p_{mn})]$$

where p denotes pollen gene frequencies (e.g. p_{mn} is the frequency of alleles m and n in the paternal gamete pollen pool). Finally, the probability of observing $A_mA_nA_oA_p$ given its parentage and that it was produced by apomixis is simply

$$\text{prob}(A_mA_nA_oA_p \mid A_iA_jA_kA_l, \text{apomixis}) = \delta_{im}\,\delta_{jn}\,\delta_{ko}\,\delta_{lp}$$

(all alleles must match between parent and progeny). The selfing, outcrossing and apomixis probabilities combine these segregational probabilities to give the overall probability of an offspring genotype as

$$\text{prob}(A_mA_nA_oA_p \mid A_iA_jA_kA_l) = s\ \text{prob}(A_mA_nA_oA_p \mid A_iA_jA_kA_l, \text{self}) + t\ \text{prob}(A_mA_nA_oA_p \mid A_iA_jA_kA_l, \text{outcross}) + a\ \text{prob}(A_mA_nA_oA_p \mid A_iA_jA_kA_l, \text{apomixis})$$

where s is the prior selfing rate, t the prior outcrossing rate, and a the prior rate of apomixis ($s + t + a = 1$). The multilocus extension of this formula is of the form

$$\text{prob}(G_{off}|G_{par}) = s\prod_i \text{prob}(G_{off,i}|G_{par,i}, \text{self}) + t\prod_i \text{prob}(G_{off,i}|G_{par,i}, \text{outcross}) + a\prod_i \text{prob}(G_{off,i}|G_{par,i}, \text{apomixis})$$

where $G_{off,i}$ and $G_{par,i}$ denote offspring and parent genotypes, respectively, at locus i; the probs are the single-locus probabilities derived above. At this point, the standard procedure and assumptions for outcrossing estimation are applied (see Ritland, 2002).

To ensure that the above likelihood analysis has sufficient information for joint estimation of these three rates of mating, the information matrix was calculated by first calculating Fisher information values (second derivatives of the log-likelihood) across 10^6 simulated parent–offspring pairs, then inverting the matrix to find the variance–covariance structure of estimates. Lack of invertability means that there is not sufficient information to jointly estimate s, t, and a. True values of s, t, and a were each one third, and gene frequencies were assumed to be equal for cases of 2, 4, and 6 alleles per locus. The results are summarized in Appendix 1B.

Appendix 1B shows, first of all, that selfing and apomixis are separable statistical entities and can be jointly estimated (given that the information matrix was invertable). For both diploid and tetraploid models, apomixis can be estimated more accurately than selfing. As expected, additional marker loci reduce the variance (increase the information) about all parameters in both the diploid and tetraploid cases. However, asymptotic variances are nearly attained by

about 2–4 loci (4–6 alleles per locus) or 4–8 loci (2 alleles per locus); at this point, improvement can occur only by adding more individuals rather than loci. Appendix 1B also shows that both selfing and apomixis were more accurately estimated with tetraploid data, given that the number of markers is doubled in tetraploids. This is partly because dosage was ascertained; if dosage cannot be ascertained, the difference will be less (a program for such a model has yet to be written). Finally, the statistical correlation between the joint estimates was strongly negative; this is expected because with more loci they are close to binomial random variables whose covariance is $-sa$, or in this case $-(1/3)(1/3) = -1/9$ which corresponds to a correlation of –0.5.

Appendix 1B. Statistical variance (per progeny) of joint estimates of the rates of selfing (*s*) and apomixis (*a*) as a function of the number of loci, number of alleles per loci, and mode of inheritance (diploid vs. tetraploid).