Pollen contamination and mating patterns in a *Prosopis alba* clonal orchard: impact on seed orchards establishment

Ivan D'Amico [1],
Juan Cesar Vilardi [1-2],
Beatriz Ofelia Saidman [1-2],
Mauricio Ewens [3],
Cecilia Bessega [1-2]

Prosopis alba (Leguminosae) is an important species from ecologic and economical points of view in arid and semi-arid regions of Argentina. In several open-pollinated species, pollen contamination from off-orchard parents and selfing have been proven to reduce orchard seed quality. In 2002, the first clonal orchard of *Prosopis alba* was established in Fernández (Santiago del Estero, Argentina) with 12 trees phenotypically selected from a progeny trial, based on height, pod production per year and pod sweetness. The aim of this study was to evaluate the mating patterns and pollen contamination rate in the orchard using ten SSR markers and paternity analysis. All the clones together with the progeny of a single clone (open-pollinated seeds) were genotyped. Data was processed by two different methods based on likelihood and Bayesian approaches, respectively. A high consistency (89%) of results was observed between the two methods, and pollen contamination rate was estimated between 27% and 37%. The minimum number of different pollen donors per mother plant varied from three to five and selfing occurrence was low (<1.6%). Based on the estimated status number (N_s = 4.4), the expected coancestry in the seed crop is equal to a Mendelian population with an effective size of 4-5 individuals. Genetic analyses are encouraged during the establishment and monitoring of trials in forest breeding and management programmes. It is strongly recommended to establish seed orchards in isolated areas and to guarantee equal representation of parental genotypes in the orchards.

Keywords: Cervus, MasterBayes, Microsatellites, Paternity Analysis, Mesquite, *Prosopis alba*, Seed Orchard

Introduction

Prosopis alba is an important ecological and economical species distributed in arid and semiarid areas of South America, belonging to genus *Prosopis* (Leguminosae – Burkart 1976). It is a native forest tree widely found in Argentina, Uruguay, Paraguay and from south of Bolivia to Peru. Nowadays, overexploitation and land conversion into crop plantations in the biogeographical Chaco region have led to important loss of its forest habitats. The large genetic differentiation among *P. alba* populations and its distribution in a variety of different environments may be exploited in forest tree breeding programs of this species. Previous studies indicated that *Prosopis* species are mainly outcrossers (Bessega et al. 2011) with endozoic seed dispersion mediated by small and medium mammals (Reynolds 1954), while pollen is mainly dispersed by insects (Genisse et al. 1990). In a natural population of *P. alba* the pollen was estimated to disperse from 5 to 30 m around each tree (Bessega et al. 2011). Consistently, studies on the spatial genetic structure indicated that the dispersal distance in *P. alba* ranges from 66 to 250 m (Bessega et al. 2016).

Seed orchards are efficient tools to produce seeds easily, abundantly and economically from trees with desirable genetic properties (Di Giovanni & Kevan 1991). They can be constituted from seeds (sexual orchards) or from vegetative propagules from selected individuals based on phenotypic traits (clonal orchard – Chmura et al. 2012). Their efficiency depends on the effective size of the orchard and the ability to exclude non-selected individuals (Wang et al. 1991). Ideally, seed orchards are designed to avoid selfing and outer pollen contamination, assuming that all crosses are equally available and compatible (Buiteveld et al. 2001). Pollen contamination by non-selected trees from outside of the orchards and selfing may have similar negative effects on the progeny by reducing the genetic gain (Lowe & Wheeler 1993). Pollen contamination level depends on the amount of pollen produced by the selected trees and their geographic isolation from external non-desired sources.

An accurate estimation of seed orchard reproductive patterns are necessary. Parental analyses have become an important method for providing a detail description of genetic diversity and pollination patterns inside the seed orchards mainly in anemophilous tree species (Hansen 2008, Torimaru et al. 2009, Feng et al. 2010, Bilgen & Kaya 2014, Ebrahimi et al. 2018).

✉ (1) Departamento de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires (Argentina); (2) Instituto de Ecología, Genética y Evolución (IEGEBA) CONICET-UBA, Buenos Aires (Argentina); (3) Estación experimental Fernández, Universidad Católica de Santiago del Estero (UCSE), Santiago del Estero (Argentina)

@ Cecilia Bessega (cecib@ege.fcen.uba.ar)
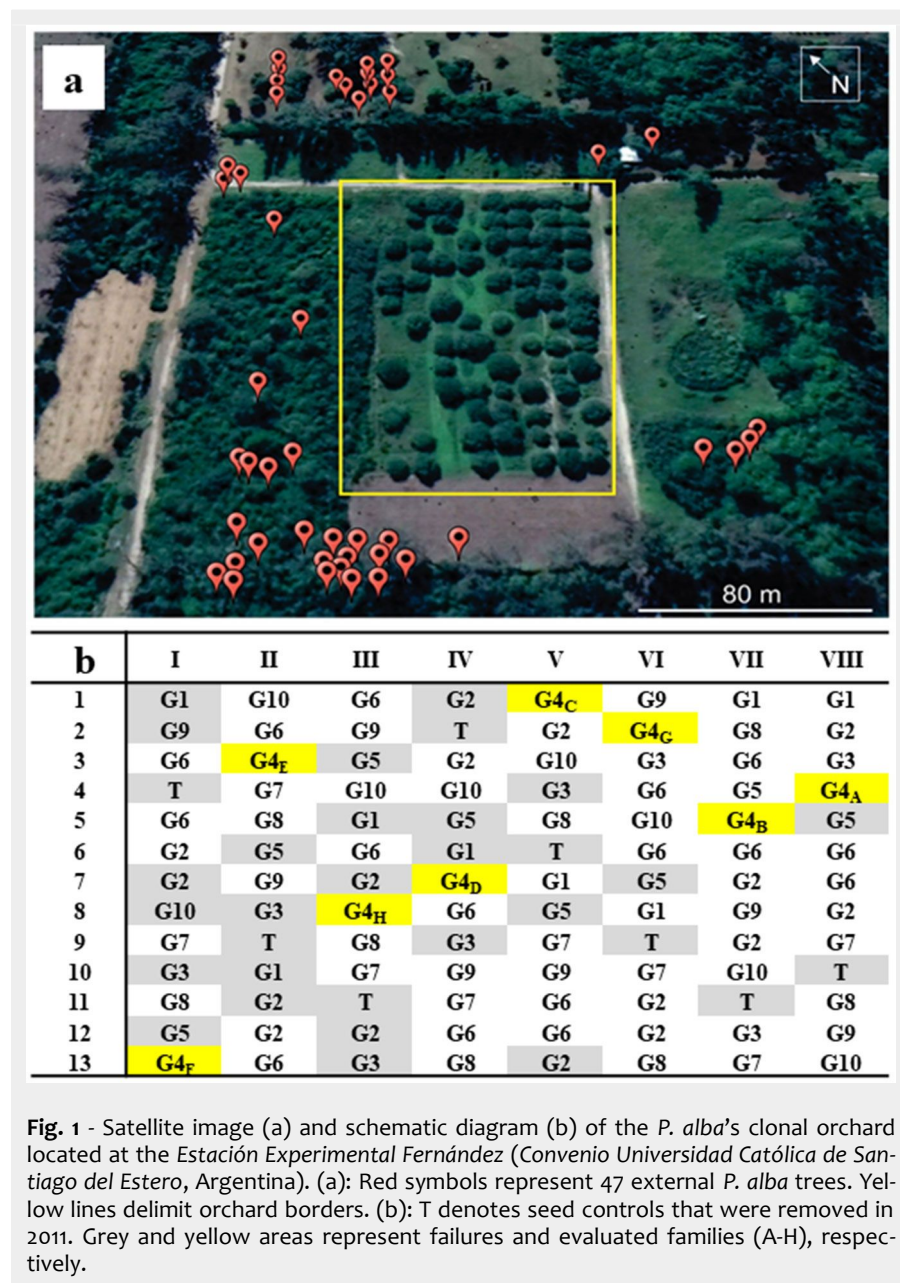
Fig. 1 - Satellite image (a) and schematic diagram (b) of the *P. alba*'s clonal orchard located at the *Estación Experimental Fernández* (*Convenio Universidad Católica de Santiago del Estero*, Argentina). (a): Red symbols represent 47 external *P. alba* trees. Yellow lines delimit orchard borders. (b): T denotes seed controls that were removed in 2011. Grey and yellow areas represent failures and evaluated families (A-H), respectively.

and 3 trees). From these trees a clonal orchard was established in 2002 at the *Estación Experimental Fernández* (*Convenio Universidad Católica de Santiago del Estero*, Santiago del Estero, Argentina – Felker et al. 2001).

The purpose of the present study was to estimate pollen contamination rate and pollination patterns in this first-generation clonal orchard of *P. alba* using SSR markers, in order to provide significant information for developing clonal seed orchard management strategies and evaluating the impact of pollen contamination on the expected gain.

Materials and methods

Seed orchard and plant material

A *P. alba* progeny trial was established in 1990, 10 km East of Santiago del Estero City, Argentina (27 ° 45′ S; 64 ° 12′ W). The experimental design was a randomized complete block design with 57 families, seven replications and four trees per replication. From them, 12 trees were selected based on height, pod production and pod sweetness and cloned (Felker et al. 2001). Shoots from these clones were multiplied by rooting of cuttings and implanted in 2002 in a new clonal orchard situated at the *Estación Experimental Fernández* (*Convenio Universidad Católica de Santiago del Estero*, Santiago del Estero, Argentina – 27° 56′ S, 65° 52.5′ W) using 10 × 10 m spacing between trees. The orchard was constituted by 8 completely randomized blocks, including ramets of the 12 selected clones and one reference tree (non selected) in each block, making a total of 104 plantlets (Fig. 1). However, the design became unbalanced because at the implantation time the number of cuttings that were successfully rooted and developed varied among the clones reducing the effective number of installed trees (M. Ewens, personal communication). Naturally regenerated trees in the clonal orchard site were not allowed. In the surrounding area, 47 *P. alba* plants were recognized at distance ranging from 15 to 150 m from the orchard edge.

At sampling, all trees were 13 years old. Adult fresh leaflets were taken from each tree and stored in a bag with silica gel until DNA extraction. One of the clones that was represented in all blocks was chosen randomly for seed analysis (clone 4). For adult genotyping, fresh leaflets were sampled from ramets at block VIII for all the clones but clone 5 that was sampled from block VII (Fig. 1) as it was missing in the former. Pods samples were collected from all ramets of the clone 4 located at all the different blocks (I-VIII). The progeny sample analyzed from of each mother plant (ramet) included eight seeds randomly taken from different pods, yielding 64 individuals.

DNA isolation and SSR analysis

Total genomic DNA was extracted using the DNeasy Plant® kit (Quiagen, Valencia,

However, for insect pollinated trees only few studies are available (Telfer et al. 2015, Yang et al. 2017). Simple sequence repeats (SSRs) are the most used markers for parental analysis. Both likelihood and Bayesian approaches (Jones et al. 2010, Telfer et al. 2015) showed to be useful for parental assignments in forest tree species.

Currently, *P. alba* breeding programmes imply the establishment of several provenance trials in different sites of Argentina together with a few clonal orchards (Felker et al. 2001, Lopez 2005, Ewens & Felker 2010, Ewens et al. 2012). Although origin certified seeds are available from INASE (*Instituto Nacional de Semillas* – National Institute of Seeds), seeds from provenance trials have not quality certification yet. In 1990, a progeny trial was established in San Carlos, Santiago del Estero (Argentina), in a randomized complete block design with 57 families, 7 replication and 4 trees per replication. This progeny trial was established

from seeds representing wild open-pollinated families from mother plants with good phenotypic characteristics, considering vigor, crown size, unbranched condition, long and straight trunk, and healthy aspect (Felker et al. 2001). The original stand included samples from eight unrelated provenances representing different areas and *P. alba* morphotypes. The *Santiagueño* type was represented by the populations of Rio Dulce Irrigation Zone (RDU), Sumampa (SUM), Pinto (PIN) and Añatuya (ANA), the *Chaco Norte* by Castelli (CAS) and Ibarreta (IBA) and the *Chaco Sur* type by Quimili (QUI) and Gato Colorado (GCO – Bessega et al. 2015b). In 1999, 12 trees were selected from this trial, based on height, pod production per year and pod sweetness. The selected trees belong to different families and derived from the 4 provenances RDU and PIN from Santiago del Estero, CAS from Chaco, and IBA from Formosa (represented respectively by 4, 1, 4,

CA, USA) following the manufacturer's instructions from leaves and 2- to 5-day-old seedling cotyledons. A total of 64 seeds and the 12 adult samples were characterized. Samples were genotyped at 10 SSR markers (M008, M009, GL6, GL8, GL9, GL12, GL15, GL18, GL21 and GL24) previously described for *Prosopis* (Mottura et al. 2005, Bessega et al. 2013). PCR amplïfications are described in Bessega et al. (2009, 2015a) using FAM- or HEX fluorescent dye-label primers in a VERITI® termocycler (Applied Biosystems, Foster City, CA, USA). PCR products were electrophoresed by MACROGEN service (https://dna.macrogen.com/eng/index.jsp).

Data analysis

PCR products were sized using GeneMarker v. 1.91 (SoftGenetics 2018). Genetic diversity was quantified through the number of alleles (A), observed (H$_O$) and expected Nei's diversity (H$_E$), and polymorphic information content (PIC) indices, using the software Cervus ver. 3.0.7 (Marshall et al. 1998).

In order to determine the power of the loci analyzed to discriminate between unique individuals we conducted two analyses. First, we obtained the genotype accumulation curve using the command "genotype_curve" of the "poppr" package (Kamvar et al. 2014, Kamvar et al. 2015) in R ver. 3.5 (R Core Team 2018), setting the number of resampling to 1000. Then, the accumulated probability of identity (PI) and the probability of identity when related individuals are included in the sample (PIsibs) were estimated using the software package GenAlex ver. 6.5 (Peakall & Smouse 2012).

Parental analysis was performed following two approaches: (i) maximum likelihood, using the software Cervus ver. 3.0.7; and (I) Bayesian inference, using the MasterBayes package of R. Both methods are capable of incorporating missing data, genotyping errors and null alleles in parentage inference (Hadfield et al. 2006, Kalinowski et al. 2007).

Cervus

Paternity inference with Cervus ver. 3.0 assigns paternity using a statistics (delta), which is defined as the difference in the log-likelihood values of the two most likely pollen donors. In this sense, an individual will be assigned paternity if its likelihood is sufficiently higher than that of the second most likely male parent. Simulations (100,000) were run in order to estimate a critical value of delta above which a stated proportion of assignments in the population would be correct (Marshall et al. 1998, Kalinowski et al. 2007). For this purpose, allele frequencies at each locus were estimated taking into account the unbalanced state of the orchard, and the parameters for simulating genotypes were: simulated offspring = 100,000; candidate parents = 12; proportion of candidate parents sampled =

0.6; proportion of loci typed = 0.91; proportion of loci mistyped = 0.01; and minimum typed loci = 8. Relaxed (80%) and strict (95%) confidence criteria were considered to assign paternity, following the manual recommendations (Marshall et al. 1998).

MasterBayes

Based on genotypic information, the marginal posterior distributions of father-offspring pairs were used to calculate individual confidence levels in each assignment. This model was equivalent to the model run in Cervus except that confidence in relationships was assessed at the individual as opposed to the population level. First, the allelic frequencies were estimated considering the unbalanced design using the function "extractA". Then, the offspring specific design matrices were constructed with the "varPed" function, assuming outcrossing = 1 based on Bessega et al. (2011) and considering that G4 was the corresponding mother tree. Based on the Bayesian approach (using the "MCMCped" function), a categorical pedigree was generated in which only assignments with at least 80% confidence were accepted. This threshold was used for all MasterBayes assignments to allow comparison with relaxed confidence criteria used in Cervus analysis (see above).

Based on the resulting assignments, pollen contamination rate (*m*) in the clonal orchard, explained by the pollen contribution from the surrounding natural population, was estimated as the proportion of seeds which a father from the orchard could not be assigned. The individuals were considered as selfed when the assigned father genotype was the same as the known

mother genotype. When the orchard seeds resulted from fertilization by pollen from outer trees, the prediction genetic gain is considered to be one half of that expected under no contamination (Squillace & Long 1981). The expected reduction in genetic gain under contamination (G$_R$) is therefore (eqn. 1):

$$G_R = \frac{m \cdot G}{2} \tag{1}$$

where *G* is the gain expected under no contamination (taken as unity), assuming that external pollen sourced originated only from completely unimproved trees.

The paternity patterns were analyzed in different ways. First, the relationship among paternity and ramet abundance was evaluated using a generalized linear model with negative binomial distribution, with the "MASS" package (Venables & Ripley 2002) in R. This analysis was performed excluding G4, because according to Bessega et al. (2011) selfing in *P. alba* is absent. Second, in order to quantify the effect of unequal male contribution, the status number (N$_S$) was estimated by the following formula (Lindgren & Mullin 1998 – eqn. 2)

$$N_s = \frac{1}{\sum p_i^2} \tag{2}$$

where $\sum p_i^2$ is the sum of the squared male contribution (*p$_i$*). Finally, for each family (mother plant G4, individuals G4$_A$ to G4$_H$), the minimal number of different fathers retrieved were counted and compared.

Results

The SSR loci analyzed were highly variables, showing between 2 and 12 alleles each (mean A value = 6.7 – Tab. 1). Aver-

Tab. 1 - Genetic diversity estimates of *P. alba* clonal orchard and details of the SSR loci analyzed. (*T$_a$*): Annealing primer temperature; (*N*): sample size; (*A*): number of alleles; (*H$_O$*, *H$_E$*): observed and expected heterozygosity, respectively; (*PIC*): polymorphic information content; (*f$_{null}$*): null allele estimated frequency.

Locus	Sequences (5' -3')	T$_a$ (°C)	N	A	H$_O$	H$_E$	PIC	f$_{null}$
GL6	F: CTGGTTGCTGTGATTGGAGG R: CTCCAGGGATCACAAGACAAAC	62	136	4	0.36	0.32	0.29	-0.08
GL8	F: CAGGTGGGCATGAAGTTTCC R: CCAAGAACAACCTGCCGAAG	58	131	12	0.80	0.74	0.72	-0.04
GL9	F: ACTCTGCGGGTTAGGTAAGC R: ACCTGGAGCTGACATGGATC	58	125	2	0.17	0.19	0.17	0.07
GL12	F: GAGTGAAGGTCGGGAAGAGG R: CCATTGGACCAAGGCAGAAC	58	134	9	0.66	0.74	0.71	0.06
GL15	F: GTGTTATGGTCCCAACAGCC R: TGAAGAGGGAGGAATCGCAG	58	127	8	0.80	0.78	0.75	-0.01
GL18	F: GAGAATCTGGAGCAGCAACG R: AAGGTAGCGTCCCAGGTATG	58	133	10	0.96	0.81	0.78	-0.10
GL21	F: ATCTCCGTCACAACTTGCAC R: ACCCTCACTCCCGAATGATG	58	136	4	0.49	0.55	0.47	0.05
GL24	F: CCTTAATCTCCCTCTCGGCC R: AACCAGGCTCTGCAGAAATG	58	130	11	0.92	0.78	0.75	-0.08
MO08	F: TATCCTAAACGCCGGGCTAC R: TCCCATTCATGCATACTTAAACC	59	136	4	0.38	0.34	0.32	-0.08
MO09	F: ATTCCTCCCTCACATTTTGC R: CATTATGCCAGCCTTTGTTG	59	134	3	0.29	0.26	0.24	-0.07
Mean	-	-	132.2	6.7	0.58	0.55	0.52	-0.03
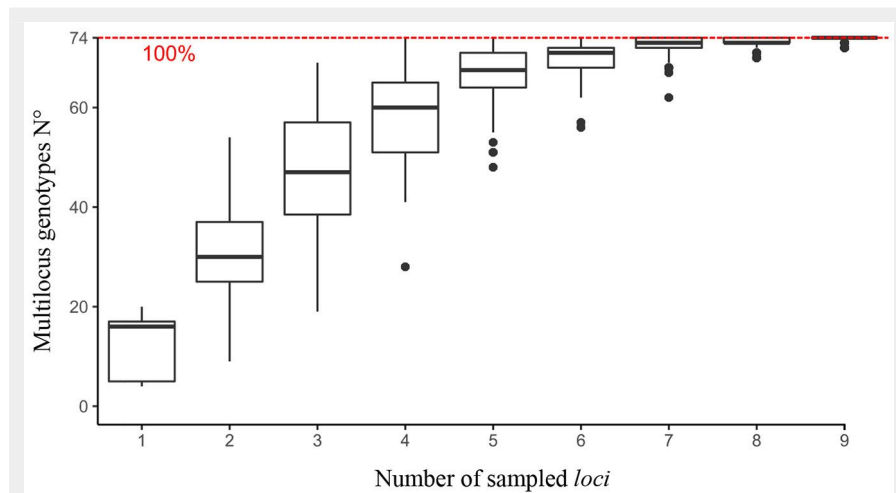
Fig. 2 - Genotype accumulation curve. Box plots were constructed from the number of observed genotypes considering a randomized bootstrap without resampling from the original data.

aged observed heterozygosity was higher than the expected heterozygosity. All the estimated genetic parameters indicated a rather high variability for the loci studied, and only 2 out of 10 loci (GL9 and MO09) showed a relatively low polymorphism (PIC < 0.25). The information produced by the 10 loci allowed the identification of 74 distinct multilocus genotypes in the clonal orchard, 10 corresponding to the adult trees (G1-G10) and 64 from the seeds analyzed.

The analysis of the adult plants allowed to distinguish only 10 different multilocus genotypes, although the clonal orchard was established using 12 clones. Indeed, using the set of loci analyzed here clone 2 and clone 8 cannot be differentiated from each other and the same holds for clone 6 and clone 7. Each genotype can be cross referenced with the clonal selection made by Felker et al. (2001) as follows. G1 corresponds to clone 1 (B2F17T2, block 2 family 17 tree 2 from R. Dulce), G2 to clone 2

(B6F12T1 – Rio Dulce) and clone 8 (B6F5T2 – Ibarreta), G3 to clone 3 (B1F8T4 – Castelli), G4 to clone 4 (B6F1T4 – R. Dulce), G5 to clone 5 (B2F4T3 – Ibarreta), G6 to clone 6 (B1F5T3 – Pinto) and clone 7 (B1F6T3 – Castelli), G7 to clone 9 (B6F5T2 – Ibarreta), G8 to clone 10 (B5F9T2 – Castelli), G9 to clone 11 (B1F5T4 – Ibarreta) and G10 to clone 12 (B7F6T4 – Castelli). The genotype accumulation curve (Fig. 2) suggests that the full recognition (100%) of all the studied multilocus genotypes can be achieved using only 9 loci. The accuracy of the individual identification was high considering that the accumulated probability of identity (PI) and probability of identity when related individuals are included in the sample (PI-sibs) were $7.2 \cdot 10^{-7}$ and 0.003, respectively.

Paternity analysis

Null allelic frequencies were considered negligible as they were close to zero for most of the analyzed loci, being between

0.05 and 0.1 only in three cases (Tab. 1). Based on Cervus analysis, the average non-exclusion probability for one candidate parent given the genotype of the known mother was low (NE-2P= 0.006) and no incongruence was detected among progeny and the known mother genotypes. From the total offspring analyzed, 47 individuals (seeds) were successfully assigned to one parent (pollen donor) inside the clonal orchard, of which 44 using the strict criterion and three with the relaxed one (Tab. S1 in Supplementary material). These values indicate that 69-73% of the analyzed seeds have been pollinated by a donor from the orchard, according to the different confidence criteria. The maximum pollen contamination rate (*m*) was 31%. G4 was assigned as male parent in only one case, yielding a selfing rate of 0.016.

Based on the posterior probability of each parental assignment obtained using MasterBayes (Tab. S1 in Supplementary material), 63-72% of the offspring exhibited multilocus genotypes compatibles with pollen donors from the clonal orchard, considering strict and relaxed criteria, respectively. Consequently, the external pollen contamination rate was 28-37% and no selfing was detected. The analysis of seeds originated by external pollen donors indicated that the number of external trees acting as male parents was only 2.6 (CI$_{95\%}$= 1.2-5.0).

A high consistency was found (89% of cases based on relaxed criteria) between the parental assignments using both maximum likelihood and Bayesian approaches (see Tab. S1 in Supplementary material). In 67% of cases the consistent assignments were from internal pollen donors, while 22% were from external male parents. The remaining 11% of cases showed differences among algorithms: 6% were assigned only by Cervus and 5% only by MasterBayes.

The maximum pollen contamination rates estimated by both methods allowed to as-

Tab. 2 - Number of seed assigned to each genotype, considering the different evaluated families (mother tree G4$_A$ to mother tree G4$_H$). *dN* represents the number of descendants corresponding to each adult genotype; p_i represents the relative contribution of each genotype over the total evaluated progeny; r_i represents the proportion of clone repeats in the orchard at sampling.

Assigned Male Parent	Families								dN	p_i	r_i
	G4$_A$	G4$_B$	G4$_C$	G4$_D$	G4$_E$	G4$_F$	G4$_G$	G4$_H$			
G1	0	0	0	0	0	0	0	0	0	0.00	0.06
G2	1	0	0	1	3	2	3	0	10	0.16	0.14
G3	1	1	0	0	0	0	1	1	4	0.06	0.04
G4	0	0	0	0	0	0	0	0	0	0.00	0.11
G5	0	1	0	0	0	0	0	0	1	0.02	0.01
G6	2	4	1	3	2	1	0	5	18	0.28	0.22
G7	0	1	0	0	0	0	1	0	2	0.03	0.11
G8	0	0	1	1	0	1	0	1	4	0.06	0.11
G9	0	0	1	0	0	1	2	0	4	0.06	0.10
G10	0	1	2	0	0	0	0	0	3	0.05	0.10
External	4	0	3	3	3	3	1	1	18	-	-
Pollen donors	4	5	5	4	3	5	5	4	-	-	-
Mean				4.4					-	-	-

sess a decrease in the expected clonal orchard genetic gain from 15.5 to 18.5%. All the genotypes, with the exception of G1, were able to fertilize the G4 ramet egg cells. G6 was the clone that most contributed to G4 offspring (Tab. S1 in Supplementary material).

The number of offspring assigned to each clone varied among 0 to 18 (Tab. 2). The different contribution of different clones to G4 offspring may be partially explained by the variation in the number of surviving ramets. Despite the orchard had originally been designed with a balanced design, the rooting of the cuttings was not equally effective. Therefore, 32 trees were not present at the time of sampling, which caused the design to be treated as unbalanced (Fig. 1b, Fig. 3a). At sampling, only one out of 8 ramets corresponding to G5 was present, but G6 exhibited 16 repeats as it represents clones 6 and 7. The regression of paternity success on the number of surviving ramets of each clone (Fig. 3b) is positive and highly significant (P= 0.0004). The status number (N_s), based on the relative contribution of each genotype in the gene pool of male gametes (Tab. 2), was 4.4 and the minimum number of different male parents that pollinated each mother plant (family G4, individual G4A to individual G4H) varied from 3 to 5, with an average of 4.4 (Tab. 2).

Discussion

The establishment of seed clonal orchards is an important step of the *Prosopis alba* breeding programmes in Argentina. In order to provide valuable information for the development of these orchards and suggest future management strategies, we estimated the pollen contamination rate and pollination patterns in a clonal *P. alba* stand located in Santiago del Estero, Argentina.

SSR markers are largely used tools for parental analyses (Torimaru et al. 2009, Feng et al. 2010, Yang et al. 2017). The SSR loci used here were highly polymorphic and null allele frequencies were negligible for all loci. The high genetic variability detected allowed to discriminate individual genotypes and to exclude unrelated candidate trees from parentage of any arbitrary offspring. Moreover, the genotype accumulation curve suggests that the combination of markers used were adequate to discriminate all individuals of the progeny. The lack of differentiation between some pairs of clones showing the same multilocus genotype could be solved by analyzing a larger number of molecular markers both in the clones and the corresponding plants they originated from.

Pollen contamination is an important factor that reduces the expected genetic gain in a clonal orchard. We demonstrated that between 27% and 37% of the sampled seeds are not genetically compatible with male parents located inside the clonal orchard, according to the different estimation
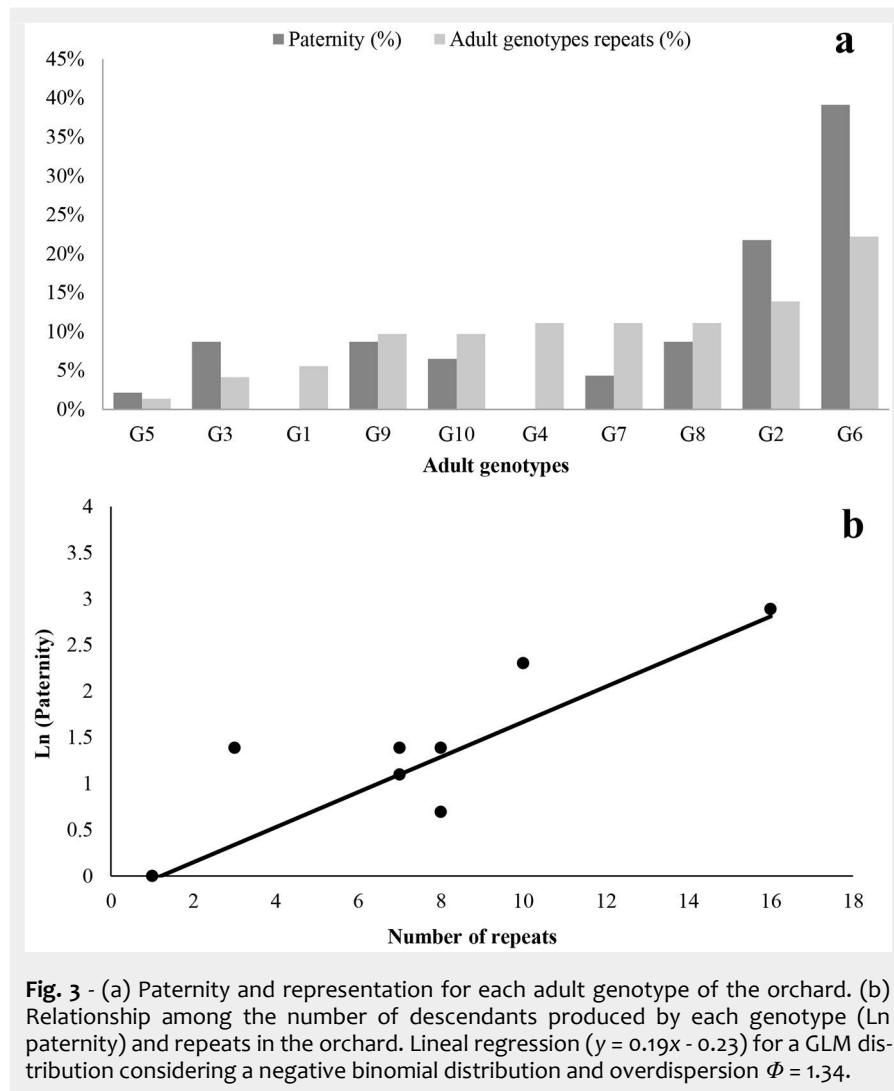
Fig. 3 - (a) Paternity and representation for each adult genotype of the orchard. (b) Relationship among the number of descendants produced by each genotype (Ln paternity) and repeats in the orchard. Lineal regression (*y* = 0.19*x* - 0.23) for a GLM distribution considering a negative binomial distribution and overdispersion *Φ* = 1.34.

methods and confidence criteria. These contamination values are within the range recorded in the literature for both wind and insect pollinated orchards. Bilgen & Kaya (2014) reported that 29% of the embryos of *Pinus brutia* were sired by pollen sources outside the orchard. Feng et al. (2010), Torimaru et al. (2009) and Buiteveld et al. (2001) estimated contamination rates of 25%, 51% and 70% for *Pinus koraiensis*, *Pinus silvestris*, and *Quercus robus,* respectively. In *Juglans nigra*, Ebrahimi et al. (2018) proved that the within-orchard pollination was significantly higher in an isolated orchard than in a non-isolated one. For the insect pollinated *Schima superba,* Yang et al. (2017) detected a lower, but important, level of pollen contamination (7.01%) originated from outside the orchard. In all these cases, individuals of the same species or a close related one could be found near the clonal orchard.

Failures in the recognition of parental trees within the studied orchard can be explained by: (i) natural regeneration of individuals from outside (*e.g.*, seeds carrying unrelated genotypes) establishing within the clonal orchard; (ii) mutational events (such as inversions) in the offspring, whose rate is fairly high in SSR regions of the

genome; (iii) scoring mistakes; and (iv) pollen flow from outside the orchard (Plomion et al. 2001, Bilgen & Kaya 2014). Some of these causes can be easily ruled out in our case. First, natural regeneration from outside could be excluded as the orchard is controlled by constant thinning in order to preserve the genotypes of interest. Secondly, mutational events can be considered as unlikely because no mismatch was found between seed multilocus genotypes and the known mother genotype. Genotyping errors might skew the estimation of pollen contamination, but the error quantification estimated by MasterBayes based on the 10 SSR loci analyzed was very low (being the allelic drop out and stochastic error equal to 0.002 and 0.02, respectively).

Considering that other *P. alba* trees were found near around the orchard, external pollen contamination could be the most likely reason to explain the occurrence of multilocus genotypes not attributable to male parents among the clones that constitute the trial. Pollen contamination rates here obtained can be explained by the contribution of 1 to 5 different trees from outside the orchard that were acting as pollen donors. A low number of parents partici-

pating in pollination was also observed in natural populations of insect-pollinated tree species (Bessega et al. 2011, Ottewell et al. 2012). The level of pollen contamination depends on factors like the amount of pollen produced inside the orchard, flowering synchronization among the clones, and annual weather variation (Harju & Muona 1989, Burczyk et al. 2004). It is not possible to determine which factor mostly affected the results of this study. Asynchronous reproductive phenology has been reported to cause temporal isolation between the orchard's breeding subpopulations of Douglas fir, leading to increased pollen contamination and increased selfing (due to the decreased number of pollen donors and to the reduction of gametic competition between orchard and outside parents) increasing the chance of gene flow (El-Kassaby & Ritland 1986). The non-synchronous flowering among families was reported as the major cause for differences among the pollen clouds of each family in *Schima superba* (Yang et al. 2017), where precocity was described for some families. Additionally, the variation of climatic factors, such as rain, frost and prevalent winds, may represent a gene flow barrier. Indeed, rainfall may strongly affect flowering and/or pollination, and intensive raining episodes may cause different clones to be differentially affected, yielding differential amount of pollen. Furthermore, insects like *Prosopidopsylla flava* and fungal pathogens like *Pestalotiopsis* spp. are known to attack mature and immature leaves and flowers of *Prosopis,* and the Argentine parrots (*Myiopsitta monachus*) to feed on clone pods. As some clones in the studied orchard are more precocious than others, it may be hypothesized that they could represent traps or escapes to pests' invasions. Finally, the entomophilous pollination may be also affected by rainy periods, when the insects' visiting rate of flowers is reduced, thereby differently influencing the pollination of clones with different blooming time. The lack of studies on phenology and pollen production in the studied orchard is a limitation of the present investigation that can help clarifying the aforementioned issues. The importance of gaining greater empirical knowledge of factors affecting pollen-mediated gene flow from *Prosopis* species is needed for avoiding gain loss. To maximize the genetic gain, this kind of studies must be repeated at different times to quantify possible variations in pollen contamination rate or, alternatively, establish seed orchards in areas well isolated from other pollen sources.

Pollen contamination from the wild to the orchard is relevant in terms of the possible reduction of the expected gains. On the other hand, pollen from the orchard (especially in case of exotic material) could represent a harmful contamination of natural populations in the surrounding environment. The risk of gene escape from the orchard to the wild may be evaluated by fine

scale genetic analyses, flowering synchrony, analysis of barrier at pre- and post-pollination and post-seed dispersal, as proposed by Barbour et al. (2006) in *Eucalyptus* species. The studied orchard has been established not far from other *Prosopis* stands near the experimental station. Avoiding contamination from any other sources is crucial for setting up new areas of seed collection to be cataloged as selected material. The results of this study are relevant for future implementations of *Prosopis* orchards for seed production, since the current stands include materials from different distant sources.

Parental inference was carried out using both likelihood and Bayesian approaches. Here, such methodologies mainly differ in the algorithm used to estimate the confidence of each assignment. While Cervus uses population-level confidence, Master-Bayes relies on individual-level confidence thresholds. Nonetheless, the results obtained by both approaches were highly consistent. Similar results have been reported by Walling et al. (2010), who found only 7 cases of parental assignment mismatches between the two methods. However, in all cases the main difference was attributable to different confidence levels, as both approaches indicated the same parental multilocus genotypes as the most probable parents.

The unequal contribution of different parental clones in the orchard to the studied offspring yielded a status number (N_s) of 4.4, *i.e.*, the expected coancestry in the seed crop is equal to a Mendelian population with an effective size of 4-5 individuals. Hansen & Kjaer (2006) found a N_s= 4.2 in a *Abies nordmanniana* orchard including 13 different clones, while Hansen (2008) estimated N_s= 4.6 for a *A. alba* orchard with 12 different clones. It was reported that the number of successful pollinations by each clone is highly related to its number of ramets in the orchard (Ebrahimi et al. 2018). Populations with low effective sizes may increase the level of relatedness between seedlings in the orchard, leading to genetic drift and inbreeding depression in future crosses (Williams & Savolainen 1996). In addition, the estimation of a status number lower than expected may indirectly increase the selective process that is done during the constitution of the orchard.

In this study the estimated number of pollen donors per mother plant was near four. This result is fairly similar to the findings of Bessega et al. (2011) in a *P. alba* natural population, where six pollen donors per mother tree were assessed by indirect methods. Moreover, the low selfing rate obtained in this study is in agreement with previous studies on this species in the wild (Bessega et al. 2011). It may be argued that the presence of multiple ramets of the same clone within the orchard led to an increased level of selfing. However, the low proportion of selfed offspring in the orchard is not a relevant problem from the

productive point of view, as reported for other forest species (Kaya et al. 2006, Hansen 2008, Torimaru et al. 2009). Our results were also consistent with previous studies on the insect-pollinated *Schima superba* (Yang et al. 2017), where a small number of effective pollen donors per mother tree (2.3) and a high overall crossover rate (98.5%) were also reported. According to Yang et al. (2017), it can be suggested that pollen distance and dispersion in trees pollinated by bees (as is the case of *P. alba*) depends on several biotic and abiotic factors such as the distance between individuals, phenology of flowering, behavior of the dispersing agent and weather conditions.

Conclusion

The *Prosopis alba* orchard analyzed in this study is the only one established by vegetative propagation for species of the genus *Prosopis*. The lack of similar trials is due to technical difficulties to obtain rooted cuttings able to successfully develop in the field. Despite the limitations due to the unbalanced design and the relatively low number of clones currently available, our results support the usefulness of genetic analyses in forest breeding and management programmes, particularly for the establishment and monitoring of trials. A recommended management strategy to reduce pollen contamination should consider the establishment of seed orchards in areas well-isolated from putative sources of contamination. Moreover, the removal of related trees acting as non-selected pollen donors or the use of gene flow barriers is encouraged. Finally, the equal representation of genotypes within the orchard is crucial to guarantee similar contributions of each genotype in the seed production. The lack of descendants for some genotypes observed in this investigation highlights the importance of conducting preliminary studies on synchrony and compatibility among the selected genotypes.

References

Barbour RC, Potts BM, Vaillancourt RE, Tibbits WN (2006). Gene flow between introduced and native *Eucalyptus* species: flowering asynchrony as a barrier to F1 hybridization between exotic *E. nitens* and native Tasmanian *Symphomyrtus* species. Forest Ecology and Management 226: 9-21. - doi: 10.1016/j.foreco.2006.01.017

Bessega C, Saidman BO, Darquier MR, Ewens M, Sánchez L, Rozenberg P, Vilardi JC (2009). Consistency between markerand genealogy-based heritability estimates in an experimental stand of *Prosopis alba* (Leguminosae). American Jour-

nal of Botany 96: 458-465. - doi: 10.3732/ajb.080 0074

Bessega C, Pometti CL, Ewens M, Saidman BO, Vilardi JC (2011). Strategies for conservation for disturbed *Prosopis alba* (Leguminosae, Mimosoidae) forests based on mating system and pollen dispersal parameters. Tree Genetics and Genomes 8: 277-288. - doi: 10.1007/s11295-011-0439-6

Bessega C, Pometti CL, Miller JT, Watts R, Saidman BO, Vilardi JC (2013). New microsatellite loci for *Prosopis alba* and *P. chilensis* (Fabaceae). Applied Plant Science 1 (5): 1200324. - doi: 10.3732/apps.1200324

Bessega C, Pometti CL, Ewens M, Saidman BO, Vilardi JC (2015a). Improving initial trials in tree breeding using kinship and breeding values estimated in the wild: the case of *Prosopis alba* in Argentina. New Forests 46: 427-448. - doi: 10.1007/s11056-015-9469-5

Bessega C, Pometti CL, Ewens M, Saidman BO, Vilardi JC (2015b). Evidences of local adaptation in quantitative traits in *Prosopis alba* (Leguminosae). Genetica 143: 31-44. - doi: 10.1007/s10709-014-9810-5

Bessega C, Pometti CL, Ewens M, Saidman BO, Vilardi JC (2016). Fine-scale spatial genetic structure analysis in two Argentine populations of *Prosopis alba* (Mimosoideae) with different levels of ecological disturbance. European Journal Forest Research 135: 495-505. - doi: 10.1007/s10342-016-0948-9

Bilgen BB, Kaya N (2014). Chloroplast DNA variation and pollen contamination in a *Pinus brutia* ten clonal seed orchard: implication for progeny performance in plantations. Turkish Journal of Agriculture and Forestry 38: 540-549. - doi: 10.3906/tar-1307-108

Buiteveld J, Bakker EG, Bovenschen J, Vries De SMG (2001). Paternity analysis in a seed orchard of *Quercus robur* L. and estimation of the amount of background pollination using microsatellite markers. Forest Genetics 8: 331-337.

Burczyk J, Lewandowski A, Chalupka W (2004). Local pollen dispersal and distant gene flow in Norway spruce (*Picea abies* [L.] Karst.). Forest Ecology and Management 197: 39-48. - doi: 10.1016/j.foreco.2004.05.003

Burkart A (1976). A monograph of the genus *Prosopis* (Leguminosae subfam. Mimosoideae) (Part 1 and 2). Catalogue of the recognized species of *Prosopis*. Journal of Arnold Arboretum 57 219 (249): 450-525. [online] URL: http://www.jstor.org/stable/43781999

Chmura DJ, Rozkowski R, Chalupka W (2012). Growth and phenology variation in progeny of Scots pine seed orchards and commercial seed stands. European Journal of Forest Research 131: 1229-1243. - doi: 10.1007/s10342-012-0594-9

Di Giovanni P, Kevan PG (1991). Factors affecting pollen dynamics and its importance to pollen competition: a review. Canadian Journal of Forest Research 21: 1151-1170. - doi: 10.1139/x91-163

Ebrahimi A, Lawson SS, Frank GS, Coggeshhall MV, Woeste KE, McKenna JR (2018). Pollen flow and paternity in an isolated and non isolated black walnut (*Juglans nigra* L.) timber seed orchard. PLoS ONE 13 (12): e0207861. - doi: 10.1371/journal.pone.0207861

El-Kassaby YA, Ritland K (1986). The relation of outcrossing and contamination to reproductive

phenology and supplemental mass pollination in a Douglas-fir seed orchard. Silvae Genetica 35: 240-244. [online] URL: http://www.research gate.net/publication/235763592

Ewens M, Felker P (2010). A comparison of pod production and insect ratings of 12 elite *Prosopis alba* clones in a 5-year semi-arid Argentine field trial. Forest Ecology and Management 260: 378-383. - doi: 10.1016/j.foreco.2010.04.030

Ewens M, Gezan S, Felker P (2012). Five year field evaluation of *Prosopis alba* clones on pH 9-10 soils in Argentina selected for growth in the greenhouse at seawater salinities (45 dS m⁻¹). Forests 3: 95-113. - doi: 10.3390/f3010095

Felker P, Lopez C, Soulier C, Ochoa J, Abdala R, Ewens M (2001). Genetic evaluation of *Prosopis alba* (algarrobo) in Argentina for cloning elite trees. Agroforestry Systems 53: 65. - doi: 10.1023/A:1012016319629

Feng FJ, Sui X, Chen MM, Zhao D, Han HJ, Li MH (2010). Mode of pollen spread in clonal seed orchard of *Pinus koraiensis*. Journal of Biophysical Chemistry 1 (1): 33-39. - doi: 10.4236/jbpc.2010.11004

Genisse J, Palacios RA, Hoc PS, Carrizo R, Moffat L, Mom MP, Agullo MA (1990). Observaciones sobre la biologia floral de *Prosopis* (Leguminosae, Mimosoidae). II Fases florales y visitantes en el distrito Chaquenio Serrano. [Observations on the floral biology of *Prosopis* (Leguminosae, Mimosoidae). II Floral phases and visitors in the Chaquenio Serrano district]. Darwiniana 30: 71-85. [in Spanish]

Hadfield JD, Richardson DS, Burke T (2006). Towards unbiased parentage assignment: combining genetic, behavioral and spatial data in a Bayesian framework. Molecular Ecology 15: 3715-3730. - doi: 10.1111/j.1365-294X.2006.03050.x

Hansen OK, Kjaer ED (2006). Paternity analysis with microsatellites in a Danish *Abies nordmanniana* clonal seed orchard reveals dysfunctions. Canadian Journal of Forest Research 36: 1054-1058. - doi: 10.1139/x05-299

Hansen OK (2008). Mating patterns, genetic composition and diversity levels in two seed orchards with few clones - Impact on planting crop. Forest Ecology and Management 256: 1167-1177. - doi: 10.1016/j.foreco.2008.06.032

Harju A, Muona O (1989). Background pollination in *Pinus sylvestris* seed orchards. Scandinavian Journal of Forest Research 4: 513-520. - doi: 10.1080/02827588909382584

Jones AG, Small CM, Paczolt KA, Ratterman NL (2010). A practical guide to methods of parentage analysis. Molecular Ecology Resources 10: 6-30. - doi: 10.1111/j.1755-0998.2009.02778.x

Kalinowski ST, Taper ML, Marshall TC (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16: 1099-1106. - doi: 10.1111/j.1365-294X.2007.03089.x

Kamvar ZN, Tabima JF, Grünwald NJ (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ 2: e281. - doi: 10.7717/peerj.281

Kamvar ZN, Brosís JC, Grünwald NJ (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Fron-

tiers in Genetics 6: 208. - doi: 10.3389/fgene.2015.00208

Kaya N, Isik K, Adams WT (2006). Mating system and pollen contamination in a *Pinus brutia* seed orchard. New Forest 31: 409-416. - doi: 10.1007/s11056-005-0876-x

Lindgren D, Mullin TJ (1998). Relatedness and status number in seed orchard crops. Canadian Journal of Forest Research 28: 276-28. - doi: 10.1139/x97-217

Lopez C (2005). Evaluación de la variación genética de especies del género *Prosopis* de la Región Chaqueña Argentina para su conservación y mejoramiento [Genetic variation evaluation in species of genus *Prosopis* from the Chaquenian Region for conservation and Management]. In: "Mejores árboles para más forestadores" [Better trees for foresters] (Norberto C ed). Secretaria de Agricultura, Ganadería y Pesca, Buenos Aires, Argentina, pp. 195-204. [in Spanish]

Lowe WJ, Wheeler NC (1993). Pollen contamination in seed orchards. In: "Advances in Pollen Management" (Bramlen DL, Askew GR, Blush TD, Bridgwater FE, Jett JB eds). Agriculture Handbook 698, USDA Forest Service, Washington, DC, USA, pp 49-53.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998). Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7: 639-655. - doi: 10.1046/j.1365-294x.1998.00374.x

Mottura MC, Finkeldey R, Verga AR, Gailing O (2005). Development and characterization of microsatellite markers for *Prosopis chilensis* and *Prosopis flexuosa* and cross-species amplification. Molecular Ecology Notes 5: 487-489. - doi: 10.1111/j.1471-8286.2005.00965.x

Ottewell K, Grey E, Castill F, Karubian J (2012). The pollen dispersal kernel and mating system of an insect pollinated tropical palm, *Oenocarpus bataua*. Heredity 109: 332-339. - doi: 10.1038/hdy.2012.40

Peakall R, Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - An update. Bioinformatics 28: 2537-2539. - doi: 10.1093/bioinformatics/bts460

Plomion C, LeProvost G, Pot D, Vendramin G, Gerber S, Decroocq S, Brach J, Raffin A, Pastuszka P (2001). Pollen contamination in a maritime pine polycross seed orchard and certification of improved seeds using chloroplast microsatellites. Canadian Journal of Forest Research 31: 1816-1825. - doi: 10.1139/x01-115

R Core Team (2018). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [online] URL: http://www.R-project.org/

Reynolds HG (1954). Some interrelations of the Merriam kangaroo rat to velvet mesquite. Journal of Range Management 7: 176-180. - doi: 10.2307/3894450

SoftGenetics (2018). GeneMarker, the biologist friendship software. SoftGenetics LLCTM, web site. [online] URL: http://softgenetics.com/

Squillace AE, Long EM (1981). Proportion of pollen from non orchard sources. In: "Pollen Management Handbook" (Franklin EC ed). USDA Forest Service, Washington, DC, USA, pp. 15-19.

Telfer EJ, Stovold GT, Li Y, SilvaJunior S, Grattapaglia DG, Dungey HS (2015). Parentage reconstruction in *Eucalyptus nitens* using SNPs and microsatellite markers: a comparative analysis of marker data power and robustness. PLoS ONE 10 (7): e0130601. - doi: 10.1371/journal.pone.0130601

Torimaru T, Wang XR, Fries A, Andersson B, Lindgren D (2009). Evaluation of pollen contamination in an advanced Scots pine seed orchard. Silvae Genetica 58: 262-269. - doi: 10.1515/sg-2009-0033

Venables WN, Ripley BD (2002). Modern applied statistics with S (4th edn). Springer, New York, USA, pp. 495.

Walling CA, Pemberton JM, Hadfield JD, Kruuk LEB (2010). Comparing parentage inference software: reanalysis of a red deer pedigree. Molecular Ecology 19: 1914-1928. - doi: 10.1111/j.1365-294X.2010.04604.x

Wang X, Lindgren D, Szmidt AE, Yazdani R (1991). Pollen migration into a seed orchard of *Pinus sylvestris* L. and the methods of its estimation using allozyme markers. Scandinavian Journal of Forest Research 6: 379-385. - doi: 10.1080/02827589109382675

Williams CG, Savolainen O (1996). Inbreeding depression in conifers: implications for breeding strategy. Forest Science 42: 102-117. [online] URL: http://academic.oup.com/forestscience/article/42/1/102/4627306

Yang H, Zhang R, Zhoiu Z (2017). Pollen dispersal, mating patterns and pollen contamination in an insect pollinated seed orchard of *Schima superba* Gardn. Et Champ. New Forest 48: 431-444. - doi: 10.1007/s11056-017-9568-6

Supplementary Material

Tab. S1 - Parental assignments based on Cervus and MasterBayes.

Link: D_Amico_2936@suppl001.pdf