Consistency between marker- and genealogy-based heritability estimates in an experimental stand of Prosopis alba (Leguminosae).

Cecilia Bessega, Beatriz O Saidman, María R Darquier, Mauricio Ewens, Léopoldo Sanchez, Philippe P. Rozenberg, Juan C Vilardi

American Journal of Botany 96(2): 458–465. 2009.

CONSISTENCY BETWEEN MARKER- AND GENEALOGY-BASED HERITABILITY ESTIMATES IN AN EXPERIMENTAL STAND OF *PROSOPIS ALBA* (LEGUMINOSAE)[1]

CECILIA BESSEGA,[2,3,6] BEATRIZ O. SAIDMAN,[2,3] MARÍA R. DARQUIER,[2] MAURICIO EWENS,[4] LEOPOLDO SÁNCHEZ,[5] PHILLIPE ROZENBERG,[5] AND JUAN C. VILARDI[2,3]

[2]Departamento de Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón 2, 1428, Buenos Aires, Argentina; [3]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Av. Rivadavia 1917, Buenos Aires, Argentina; [4]Estación Experimental Fernández, Departamento de Robles, Santiago del Estero, Argentina; and [5]Unité Amélioration Génétique et Physiologie Forestières, BP 20619, Ardon, 45166 Olivet Cedex, France

Prosopis represents a valuable forest resource in arid and semiarid regions. Management of promising species requires information about genetic parameters, mainly the heritability (h^2) of quantitative profitable traits. This parameter is traditionally estimated from progeny tests or half-sib analysis conducted in experimental stands. Such an approach estimates h^2 from the ratio of between-family/total phenotypic variance. These analyses are difficult to apply to natural populations of species with a long life cycle, overlapping generations, and a mixed mating system, without genealogical information. A promising alternative is the use of molecular marker information to infer relatedness between individuals and to estimate h^2 from the regression of phenotypic similarity on inferred relatedness. In the current study we compared h^2 of 13 quantitative traits estimated by these two methods in an experimental stand of *P. alba*, where genealogical information was available. We inferred pairwise relatedness by Ritland's method using six microsatellite loci. Relatedness and heritability estimates from molecular information were highly correlated to the values obtained from genealogical data. Although Ritland's method yields lower h^2 estimates and tends to overestimate genetic correlations between traits, this approach is useful to predict the expected relative gain of different quantitative traits under selection without genealogical information.

Key words: Fabaceae; heritability; Leguminosae; molecular markers; *Prosopis alba*; quantitative traits.

In many developing countries, arid and semiarid ecosystems suffer from serious environmental degradation and biodiversity impoverishment. Climatic changes and human activities have resulted in deforestation, overgrazing, soil erosion, loss of fertility, and a predisposition to periodic drought and famine. One key genus of these threatened arid ecosystems is *Prosopis* (Leguminosae, Mimosoideae), with several species in many Latin American countries representing, both ecologically and economically ideal, multipurpose trees.

Prosopis planting programs are currently based on phenotypically selected, plus trees from natural stands (Akindele and Olutayo, 2007). However, phenotypic selection might yield nonsignificant results if the additive genetic component (V_A) of phenotypic variance (V_P) is low. Estimating the heritability ($h^2 = V_A/V_P$) of selectable traits is traditionally based on sib analysis or parent–offspring regression. These approaches are difficult to implement when generation intervals are long and genealogical information is absent, as is the case for forest tree natural populations. In addition, in *Prosopis*, controlled crosses are difficult to practice because of the small size and high number of flowers per inflorescence. Consequently, little is known about the genetic components of quantitative variation of desirable traits in these species. For most purposes, selected plus trees of *Prosopis* were sampled from wild populations (Verga, 2000; Verga, et al., 2005) where pedigrees are frequently unknown. Usually plus trees must be healthy, mature, or old trees of "good shape" (straight and nonbranched trunk) and well developed (Verga, et al., 2005). It is assumed that all plus trees are genetically unrelated but, in forest trees, deviations from this assumption lead to greater inbreeding and loss of genetic gain (Thomas et al., 2002).

With information about the relationships between individuals within a population, heritability may be estimated by comparing the phenotypic variation within and between family groups. Traditionally, relationships are calculated from pedigree records (Jacquard, 1974; Cannings and Thompson, 1981). However, recent theoretical advancements based on DNA profiling techniques have been proposed to estimate relatedness (r) among individuals without pedigree information (Morton et al., 1971; Lynch, 1988; Queller and Goodnight, 1989; Ritland, 1996a; Lynch and Ritland, 1999; Wang, 2002; Milligan, 2003). Thus, marker-inferred relatedness provides a valuable resource for further inference of genetic parameters, namely heritability of quantitative traits, in natural populations (Ritland, 1996a, b; Ritland and Ritland, 1996; Mousseau et al., 1998; Thomas et al., 2000, 2002; Klaper et al., 2001). These in situ inferences have proven their utility in simple natural layouts, where family structure was known (e.g., nonoverlapping generations, population consisting of nonrelated, full sib, and half-sib individuals) (Thomas et al., 2002; Wilson et al., 2003).

[1] Manuscript received 28 February 2008; revision accepted 23 July 2008.

This research was supported by funding from Agencia Nacional de Promociones Científicas y Tecnológicas (ANPCyT) BID 1728 OC/AR, PICT 32064 and PICT 00426, Universidad de Buenos Aires (EX 321 and EX 201), CONICET PIP 5122 to B.O.S., C.B. and J.C.V. Travel from Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires to INRA-Orleans of B.O.S. and J.C.V. was supported by project II-0266-FA (GEMA, Genética de la Madera), coordinated by Ph. Rozenberg, and Fundación para Investigaciones Biológicas Aplicadas (FIBA).

[6] Author for correspondence (e-mail: cecib@ege.fcen.uba.ar)

doi:10.3732/ajb.0800074

In certain conditions, marker-based inferences may be the only method available to obtain results within reasonable time and effort, especially when dealing with forest tree species characterized by costly experimental layouts and lengthy maturations. However, these marker-based inference methodologies have several shortcomings: (1) the estimation of relationship is indirect; (2) coefficients of relatedness (*r*) do not provide enough information to fully reconstruct a genealogy (i.e., parent–offspring and full sib relationships are mixed up); (3) precision and accuracy of different relatedness estimators are affected by allele frequencies and departures from linkage equilibrium; (4) variance of *r* tends to be low in naturally outbreeding tree species; and (5) variation of *r* must be uncorrelated with variation at quantitative trait loci (QTL) (Ritland, 1996b).

The availability of some experimental stands of *Prosopis* species with a family layout gives the opportunity to test the quality of the inference provided by marker-based methodology that might be used in future estimations of genetic parameters in natural populations without any genealogical information.

The objectives of the present paper were to compare the estimates of relatedness inferred using molecular markers with those based on family records and the estimates of heritabilities inferred from molecular markers with those obtained from classical quantitative methods based on pedigreed data. To properly test the suitability of the marker-based method to estimate heritability, we chose 13 traits involving leaf morphology, spine length, and tree biomass with the expectation of covering a large range of heritability values. The validity of using genetic markers instead of genealogy information for heritability estimations in natural stands where pedigrees are unknown is discussed.

MATERIALS AND METHODS

Study population—*Prosopis alba* is an important nitrogen-fixing tree adapted to the semiarid regions of northwestern Argentina. The population analyzed was a progeny trial established in 1990, 10 km from Santiago del Estero, Argentina (27°45′S; 64°15′W) (Felker et al., 2001). This trial was established from seeds collected from 57 individual trees (half sib families) from eight northwestern Argentine sites (Añatuya, Castelli, Gato Colorado, Ibarreta, Pinto, Quimili, Rio Dulce Irrigation district, and Sumampa). The experimental design was a randomized complete block comprising 57 families, seven replicates, and four trees per replicate (with a 4 × 4 m spacing). The total planting material was 1596 individual trees, of which 1289 still survived in 1999. For the last 10 years, this stand was affected by natural conditions without any silvicultural care. For this study, we sampled 142 individuals belonging to 32 different families that had kept their original identification label. The number of trees per family varied between 3 and 12.

Morphometric data—Three biomass traits and 10 leaf morphology traits were analyzed. Height (HEI) and trunk diameter (TDI) (basal diameter at 20 cm above the ground) were scored in the field. Biomass (BMS) of each tree was estimated using the regression equation $BMS = \log_{wt} = 2.7027 \cdot \log_{TDI} - 1.1085$ (Felker et al., 1989), where $\log_{wt}$ is the logarithm of fresh weight (kg) and $\log_{TDI}$ is the logarithm of TDI (cm).

These biomass characters were chosen because they are important for selective programs and the measurement is relatively simple and nondestructive. The morphological traits, measured on herbarium specimens, were petiole length (PEL), number of pairs of leaflets per pinna (NLP), pinna length (PIL), spine length (SPL), number of pinnae (NPI), leaflet length (LEL), leaflet length/width (LEL/LEW), leaflet falcate (LEF), leaflet apex (LEX), and leaflet apex/total area (LEX/LEA). Falcate is defined as the ratio *l*/*f*, where *l* is the length of a right segment from the base to the tip of the leaflet, and *f* is the length from the same point but following the curve line that runs along the middle of the leaflet (Fig. 1A). Leaflet apex (LEX) is the ratio *t*/*s*, where *t* is the area of the upper leaflet third and *s* is the area of a rectangle with the same dimensions (width and length) of the upper leaflet (Fig. 1B). These morphology traits were chosen because, although they vary substantially within and among populations, they are important for species identification (Burkart, 1976; Pasiecznik et al., 2004).

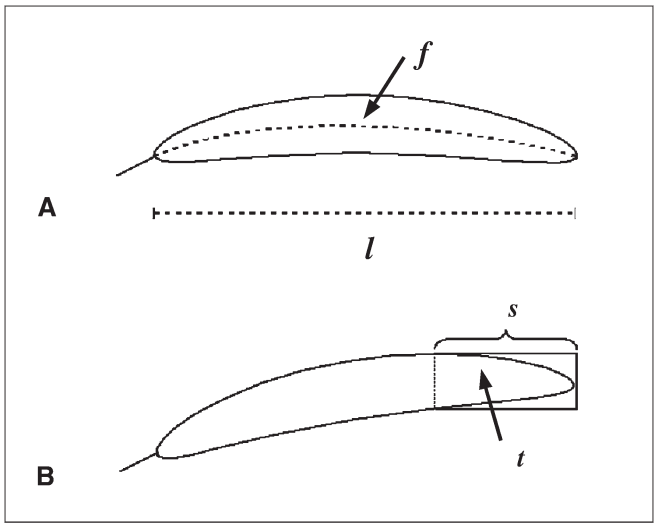

Fig. 1. Description of measurements to estimate (A) leaflet falcate and (B) leaflet apex for *Prosopis alba*. *l*: distance from the base to the tip of the leaflet, *f*: length from the base to the tip of the leaflet following a curved line running along the middle of the leaflet. *t*: area of the upper leaflet third; *s*: area of a rectangle with the same dimensions as *t*.

In each individual, nine repeats of PEL, NLP, PIL, SPL, and NPI were obtained involving three different canopy regions. Ninety repeats were obtained of LEL, LEL/LEW, LEF, LEX, LEX/LEA, involving 10 leaflets from nine pinnae sampled from three different regions of the canopy. All leaflet measures were obtained with the software HOJA1.1 (available from A. Verga, arverga@yahoo.com.ar).

Microsatellite data—Six microsatellites have been developed for *Prosopis chilensis*, and cross-species amplification has been reported (Mottuora et al., 2005). We characterized the genotypes of our sample of 142 individuals using all six microsatellites (Mo05, Mo07; Mo08, Mo09, Mo13 and Mo16).

Leaves were collected from each tree in March 2006 and were silica-gel preserved. DNA was extracted using DNA easy plant mini kit (Qiagen, Valencia, California, USA), and samples were placed in a –20° freezer until analysis. The PCR amplifications were carried out in a 50-μL reaction volume containing 10–30 ng DNA, 0.6 uM each primer, 0.2 mM dNTPs, 0.3 U *Taq* DNA polymerase (Invitrogen, Carlsbad, California, USA), and 1.5 mM $MgCl_2$. A PROGENE Techne thermalcycler (Techne Cambridge Ltd., Duxford Cambridge, UK.) was used for amplifications, where the cycling profile was initial denaturation at 94°C for 5 min followed by 30 cycles at 94° for 45 s denaturation, primer-specific annealing temperature (56°–59°) for 45 s and at 72°C for 45 s extension; and a final extension step at 72° for 10 min. Seven microliters of PCR product were separated by electrophoresis in a Model S2 apparatus (Gibco BRL Sequencing System, Life Technologies (Gaithersburg, Maryland, USA)) through 6% (w/v) polyacrylamide gel containing 5 M urea in 1× TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8). A 10-bp DNA Ladder (Invitrogen) size marker was included twice in each electrophoresis run. Gels were stained with silver nitrate (Bassam et al., 1991).

Genetic variabilty estimates—Individual patterns of microsatelites were converted into population allelic frequencies. As null alleles were suspected in some loci, allelic frequency estimation and comparison between observed and Hardy–Weinberg expected heterozygote frequencies were computed by the maximum likelihood method described in Kalinowski and Taper (2006). The Hardy–Weinberg test for heterozygote deficiency was performed by Monte Carlo randomization as described by Guo and Thompson (1992) and the *U* test statistic described by Rousset and Raymond (1995) with the program ML-NullFreq (available at website http://www.montana.edu/kalinowski). From estimated allelic frequencies, variability was quantified by the unbiased expected heterozygosity (*H*) (Nei, 1978).

Heritability and relatedness—Heritabilities were estimated by three methods. Method 1 was a conventional analysis of variance (ANOVA), assuming that the sampled individuals represent different groups of half sib families. In this case, we used an unbalanced generalized linear model (GLM), $y_{ij} = \mu + f_i + e_j$, where y_{ij} is an observation of the trait for an individual tree of family i in the environment j, μ is the overall mean, f_i represents random family effects, and e_j is the random residual error. Block effects could not be included because in the surviving stand several families were present in single blocks. Variance components were estimated by restricted maximum likelihood (REML). For a half sib design, heritabilities were estimated as $h^2 = 4\sigma_b^2 \div (\sigma_b^2 + \sigma_w^2)$, where σ_b^2 denotes the estimated variance between families and σ_w^2 is the within-family variance component. Heritabilities were considered significant whenever the ANOVA yielded significant differences among families. Confidence intervals for h^2 estimations were obtained following Lynch and Walsh (1998, p. 563).

Methods 2 and 3 were based on a linear model, where narrow-sense heritability (h^2) is estimated by regressing pairwise phenotypic similarity (Z_{ij}) on pairwise relatedness (r_{ij}) (Ritland, 1996a). According to this model,

$$h^2 = \frac{\mathrm{cov}(Z_{ij}, r_{ij})}{2\mathrm{Act\ var}(r_{ij})},$$

where Act var(r_{ij}) is the actual variance of relatedness, as in Ritland (2000), and phenotypic correlation (Z_{ij}) between individuals i and j given by:

$$Z_{ij} = \frac{(Y_i - U)(Y_j - U)}{V},$$

where Y gives the individual trait value, and U and V are, respectively, the corresponding mean and variance of the sample.

In method 2, within-family relatedness (r_{ij}) was assumed to be 0.25, as expected from a cohort of half sib siblings, and between-family relatedness zero as expected from unrelated families. In method 3, however, r_{ij} estimates were based on Ritland's (1996a) estimation method. For individuals i and j with genotypes A_rA_s and A_tA_u, respectively, r_{ij} was estimated as

$$\hat{r}_{ij} = \frac{[(\delta_{rt} + \delta_{ru})/p_r] + [(\delta_{st} + \delta_{su})/p_s] - 1}{4(n-1)},$$

where n is the number of alleles, δ_{rt} is 0 if $A_r \neq A_t$ or δ_{rt} is 1 if $A_r = A_t$, p_r is the frequency of the allele A_r in the population as estimated from the sample and so on.

For multiple locus estimates, r_{ij} corresponds to the sum of locus-specific estimates, each weighted by $(n - 1)$ (Ritland and Travis, 2004).

Pairwise Ritland relatedness estimates were computed using HERINAT, a Visual Fortran program designed ad hoc (available from author, leopoldo.sanchez@orleans.inra.fr).

The significance of heritability estimates based on the regression models 2 and 3 were obtained by a Mantel permutation test between the matrix of phenotypic similarities (Z_{ij}) and the matrix of estimated relatedness (r_{ij}). Confidence intervals for heritability estimates obtained from molecular marker inferred relatedness were obtained empirically from 100 bootstrapped pseudoreplicates of the original data over individuals. Standard errors of h^2 were estimated using eq. 3 in Ritland (1996b).

The reliability of the regression models was tested by Spearman rank correlation and regression analysis of heritability estimates by methods 2 and 3 on the estimates based on the unbalanced ANOVA of method 1.

Genetic correlations between traits were estimated by two methods. The first was based on the correlation between trait family averages as suggested in Lynch and Walsh (1998). The second one is based on a linear model for the covariance between traits (Ritland, 1996b) in which

$$r_{A_{12}} = \frac{\nu_{A_{12}}}{\sqrt{\nu_{A_1}\nu_{A_2}}},$$

where ν_{A12} is the additive covariance between the two traits and ν_{A1} and ν_{A2} the corresponding genetic variances. The sign of the genetic correlation obtained from this method should be the same as the genetic covariance (ν_{A12}), since the denominator is positive by definition. In cases where heritability estimates were negative, the denominator cannot be solved and the genetic correlation was considered as unavailable information.

Analysis of variance, regressions and Mantel tests were conducted using the packages *nlme*, *lmer*, and *ape* of program R (R Development Core Team, 2007).

RESULTS

Microsatellites—The loci analyzed showed between three to seven active alleles. In four of six loci, individuals without any band were observed thus indicating the presence of null alleles. Furthermore, significant heterozygote deficiency was observed, which could also be attributable to the presence of null alleles. Allelic frequencies were estimated taking into account possible bias from the presence of null alleles (Table 1). Expected heterozygosity varied among loci between 0.20 and 0.76. F_{IS} estimates were positive and highly significant in four of the six loci analyzed.

Of 15 pairwise combinations of loci, four showed significant or highly significant gametic disequilibrium (Table 2), two of them were matrixwide significant after applying Bonferroni's correction.

Relatedness estimates—Two pairwise relatedness matrices were obtained. The first one represented genealogical information, and its elements had three alternative values: $r = 1$ for individuals with themselves, $r = 0.25$ for different individuals of the same half sib family, and $r = 0$ for individuals from different families. The second matrix was obtained by Ritland's method from molecular marker data.

The correlation between these matrices ($r = 0.17$, excluding the relatedness of each individual with itself, or $r = 0.56$, including the diagonal with these data) was highly significant according to Mantel test ($P = 0$, based on 10 000 permutations).

Quantitative traits—All genotyped individuals were measured for all quantitative traits (Table 3). In almost all cases the within-family component of variance was higher than the between-family component. The differences among families evaluated through the Fisher (F) statistics as $F = MS_b/MS_w$ (between mean square on within mean square) were significant or highly significant in all cases.

Assuming that families were composed of half sibs, heritability estimates were obtained for each trait by the conventional quantitative approach (Table 4, col. 1). Heritability estimates were significant or highly significant for all traits. The confidence limits are rather symmetrical (Table 4); the standard errors of estimated h^2 (calculated according to Lynch and Walsh, 1998, p. 568) were similar for all traits (0.35 ± 0.01) and were not correlated ($r = -0.27$, $P = 0.38$) with h^2 values. In several cases, h^2 was higher than 1, which could be attributable to actual within-family relatedness higher than expected, with the presence in the cohort of some full sibs and selfs and/or to geographical association within fraternal groups.

For spines and leaf traits, where we had several measurements for each individual, we estimated the components of phenotypic variance between families, between individuals within families, between different canopy regions within individuals, and the residual. The percentage of variance between families and between individuals within families was similar, averaging ≈20% each of the total variance. The highest component was the residual (≈50%), while the minimum variance component corresponded to different canopy regions/individuals (≈10%).

Heritabilities estimated by the regression method from pedigree-based relatedness yielded results consistent with the ANOVA-based estimates (Table 4), although the former estimates were smaller. According to the permutation test for these regression estimates, heritability was nonsignificant for SPL, significant for HEI and BMS, and highly significant for all the

TABLE 1. Allele frequencies, range of allele size (bp), number of individuals analyzed (N), expected (H_e) and observed (H_o) heterozygosities, and fixation index (F_{IS}) for *Prosopis alba*.

	Locus					
Alleles	*Mo 08*	*Mo 13*	*Mo 16*	*Mo 09*	*Mo 07*	*Mo 05*
Null	—	—	0.047	0.029	0.165	0.112
1	0.047	0.270	0.120	0.027	0.140	0.023
2	0.178	0.102	0.402	0.055	0.061	0.818
3	0.395	0.062	0.196	0.890	0.007	0.007
4	0.351	0.551	0.094	—	0.007	0.040
5	0.022	0.015	0.112	—	0.500	—
6	0.007	—	0.029	—	0.018	—
7	—	—	—	—	0.101	—
Range	200–220	210–240	140–170	200–250	180–210	210–220
N	138	137	138	137	136	135
He	0.686	0.609	0.761	0.203	0.689	0.316
Ho	0.384	0.620	0.710	0.153	0.390	0.111
F_{IS}	0.440***	−0.019	0.067	0.246**	0.434***	0.649***

Notes: ** $P < 0.01$; *** $P < 10^{-6}$

other characters. As with the ANOVA approach, this method also yielded h^2 estimates higher than unity.

Marker-based relatedness yielded estimates of h^2 that were lower than the values obtained from any of the two former methods (Table 4). Only six marker-based estimates were significant or highly significant, and three were of borderline significance, according to the permutation test. The confidence intervals obtained from bootstrap resampling were rather narrow and consistent with the other estimation methods, while several h^2 estimates were higher than 1. Standard error was 0.29 for all traits because this depends only on the variance of relatedness and sampling size. This value is slightly lower than those estimated from ANOVA as indicated.

In spite of the differences in h^2 estimates among the three methods, there was consistency in the ranking of estimates obtained from the different methods (Fig. 2). To test this trend, we performed Spearman rank correlation tests that showed a highly significant correlation of marker based h^2 estimates with both ANOVA (rho = 0.615, $P = 0.028$) and regression (rho = 0.857, $P = 0.0002$) methods. A linear regression analysis comparing regression and ANOVA h^2 estimates (Fig. 2) indicated that when relatedness was assigned from family records, 88% of the variance could be explained by the regression ($P = 0$), and the slope was $b = 0.91$ (CI = 0.69–1.13). When relatedness was estimated from molecular markers, 55% of the variance was explained by the regression ($P = 0.004$), with slope $b = 0.68$ (CI = 0.28–1.10). Although the accuracy of marker-based h^2 estimate is lower than that obtained from pedigree-based relatedness, the confidence intervals overlap, and in both cases they include the expected slope value of $b = 1$.

TABLE 2. Probability of gametic disequilibrium between the loci analyzed for *Prosopis alba*.

Locus	*Mo 08*	*Mo 13*	*Mo 16*	*Mo 09*	*Mo 07*	*Mo 05*
Mo 08	—					
Mo 13	0.385	—				
Mo 16	0.122	0.144	—			
Mo 09	**0.018**	**0.003**[a]	0.589	—		
Mo 07	**0.010**	0.093	**0.003**[a]	0.195	—	
Mo 05	0.474	0.066	0.115	0.099	0.864	—

Note: Boldface denotes significant disequilibrium at individual level.
[a] Matrixwide significant applying Bonferroni's correction.

Genetic correlations—Of 78 pairwise estimates of genetic correlations obtained from family trait averages, 26 were significant or highly significant (Table 5). After a sequential Bonferroni's test, six of them were matrixwide significant and eight were matrixwide highly significant. More than half of the correlations between leaf traits (19/36) were significant or highly significant. The linear model based on estimated relatedness allowed estimation of the genetic correlations for 36 trait pairs, whereas the remaining 42 cases involved traits with negative heritability estimates. The estimates of genetic correlations obtained from this method are in general higher (in absolute values) than those obtained from the former (Student's test for paired samples: $t = 4.91$; $P = 2 \times 10^{-5}$), and in many cases (13/36) Ritland's estimates were outside the range 1 to −1.

DISCUSSION

Different molecular marker-based methods for quantitative genetic analyses have been developed to provide a means for examining genetic variation in natural populations (Lynch, 1988; Queller and Goodnight, 1989; Ritland, 1996a; Mousseau et al., 1998; Lynch and Ritland, 1999; Wang, 2002; Milligan, 2003; Hardy, 2003; Garant and Kuruuk, 2005; Ritland, 2005). A relatively simple approximation is based on the estimation of pairwise relatedness using genetic markers (Lynch, 1988; Queller and Goodnight, 1989; Li et al., 1993; Ritland, 1996a; Lynch and Ritland, 1999; Wang, 2002; Hardy, 2003; Milligan, 2003). In long-lived plants with mixed mating systems such as forest tree species, where pedigrees are complex, pairwise-based analyses have the advantage over other approaches in that they can incorporate variable levels of relatedness (Ritland and Ritland, 1996; Andrew et al., 2005).

Prosopis alba, a promising forest species native to central Argentina, is a valuable natural resource in arid and semiarid regions. The success of selection programs to improve quantitative traits of economic importance depends on the extent of additive genetic variance. However, information about the genetic basis of quantitative variation in natural populations of species of *Prosopis* is still scarce (Cony, 1996; Felker et al., 2001). Previous studies of mating system and genetic structure have shown a mixed mating system in a natural population of *P. alba*, with about 28% of selfing (Bessega et al., 2000). Accordingly, most natural populations of *P. alba* have significant homozygote excess (Ferreyra et al., 2007).

TABLE 3. Basic statistics, components of phenotypic variance, and Fisher (F) test of significance of between family differences for the quantitative traits analyzed.

Trait (unit)	Acronym	Mean (SD)	Variance component		F
			Between families	Within families	
Height (m)	HEI	5.00 (0.09)	0.17	1.00	1.71*
Basal Diameter (cm)	TDI	19.17 (0.60)	97.02	410.60	1.98**
Biomass (Kg)	BMS	310.52 (41.03)	$1.6\ 10^{7}$	$1\ 10^{8}$	1.65*
Petiole length (mm)	PEL	26.71 (0.71)	30.33	42.10	4.17***
Pair of leaflets/Pinna	NLP	35.03 (0.45)	7.34	21.47	2.41***
Pinna length (mm)	PIL	95.86 (1.57)	160.72	204.95	4.19***
Spine length (mm)	SPL	2.77 (0.42)	7.17	18.30	2.68***
Number of pinnae	NPI	2.41 (0.05)	0.16	0.18	4.47***
Leaflet length (mm)	LEL	7.97 (0.18)	0.02	0.02	5.10***
Leaflet length/width	LEL/LEW	4.99 (0.07)	0.50	0.23	10.02***
Leaflet falcate	LEF	0.934 (0.003)	$4\ 10^{-4}$	$7\ 10^{-4}$	4.02***
Leaflet apex	LEX	0.960 (0.005)	$1.5\ 10^{-3}$	$1.9\ 10^{-3}$	4.20***
Leaflet apex/total area	LEX/LEA	0.209 (0.002)	$1\ 10^{-4}$	$2\ 10^{-4}$	2.96***

Notes: F = MSb/MSw, where MSb = between-family mean square, MSw = within-family mean square;* $0.01 \le P < 0.05$; ** $0.001 \le P < 0.01$; *** $P < 0.001$

The suitability of marker-based method to estimate heritability in a forest species has been tested for wood density in radiata pine by Kumar and Richardson (2005). Our paper evaluates this methodology and compares the results of heritability estimates by Ritland's (1996b) methods using many quantitative traits in a natural provenance of *P. alba*. This study involves the only available experimental stand with known pedigree and potential source of breeding material of *P. alba*. The most relevant result of this study was the general consistency between classical estimates of heritability and those obtained from molecular measures. All significant estimates obtained from the latter method, corresponding to leaf morphology traits, also yielded significant estimates under the other two alternative classical methods. No negative estimates under Ritland's method were found to be significant. However, biomass traits like height and diameter yielded nonsignificant estimates under the new method. Although classical methods revealed significant genetic variability for these traits, they are known to have low to intermediate heritabilities for an ample range of forest species (Rweyongeza et al., 2005; Zas and Fernández-López, 2005).

The estimates of genetic correlations between traits obtained from family information and marker-inferred relatedness indicated that the second method produces an upward bias, with higher estimates of correlations wherever they can be estimated. The possible cause of this overestimation of genetic correlations may be attributed to the underestimation of heritabilities and genetic variances. In fact, for marker-based estimations, genetic correlation r_{A12} (see Materials and Methods) is defined as the ratio between genetic covariance and the geometric mean between genetic variances of each trait. In this expression, the denominator (square root of $\nu_{A1} \times \nu_{A2}$) is closer to the smallest value between ν_{A1} and $\nu_{A2.}$ Consequently, the smaller the denominator, the higher the r_{A12} estimated.

This study should be considered as a preliminary approach given the fact that some assumptions inherent to this marker-based method might not have been entirely fulfilled. First, there was a homozygote excess for four of the six microsatellite loci under study. This could be partially explained by the presence of null alleles in the corresponding loci. However, this hypothesis cannot be fully tested with this pedigree, given that pollinators are unknown. Furthermore, gametic phase equilibrium could not be assumed for all the loci combinations. Finally, the absence of genetic mapping resources for this pedigree and the set of traits under study did not allow inferences to be made on the eventual linkage between markers and quantitative traits.

According to Ritland (1996b), 2–10 loci, each with 10 alleles, would be needed for adequate estimation of pairwise relatedness. In our case, the number of alleles fell below the desired threshold of 10, with 3–7 alleles per locus (if null alleles are not considered). Concerning the number of loci, though still insufficient, it was within previous recommendations. Undoubtedly, a larger number of highly polymorphic markers would be

TABLE 4. Genealogy- and marker-based heritability estimates for *Prosopis alba*. Confidence intervals (95%) are indicated in parentheses.

Trait	ANOVA	Regression	Molecular markers
Height (m)	0.57 (0.36–0.81) *	0.42*	0.15 (0.08–0.23)
Basal Diameter (cm)	0.76 (0.52–1.00)**	0.65**	−0.29 (−0.32 – −0.20)
Biomass (Kg)	0.5 (0.32–0.77) *	0.35*	−0.22 (−0.24 – −0.17)
Petiole length (mm)	1.76 (1.45–1.95)**	1.29**	0.34 (0.19–0.43)
Pair of leaflets/Pinna	1.02 (0.76–1.25)**	0.76**	0.61 (0.50–0.69)$^{+}$
Pinna length (mm)	1.76 (1.46–1.96)**	1.33**	0.81 (0.71–0.91)**
Spine length (cm)	1.13 (0.89–1.39)**	1.01	0.42 (0.38–0.53)$^{+}$
Number of Pinnae	1.84 (1.54–2.04)**	1.47**	1.36 (1.29–1.44)**
Leaflet length (cm)	1.97 (1.71–2.20) **	1.71**	0.35 (0.26–0.57)$^{+}$
Leaflet length/width	2.74 (2.51–2.89)**	2.58**	1.69 (1.59–1.92)**
Leaflet falcate	1.58 (1.41–1.91)**	1.79**	1.06 (1.02–1.16)**
Leaflet apex	1.79 (1.46–1.96)**	1.24**	0.48 (0.40–0.63)*
Leaflet apex/total area	1.23 (1.01–1.52)**	1.33**	1.03 (0.91–1.14)**

Notes: $^{+}$ $0.05 \le P < 0.10$; * $0.01 \le P < 0.05$; ** $P < 0.01$. Methods for significance and confidence interval estimation described in the text.

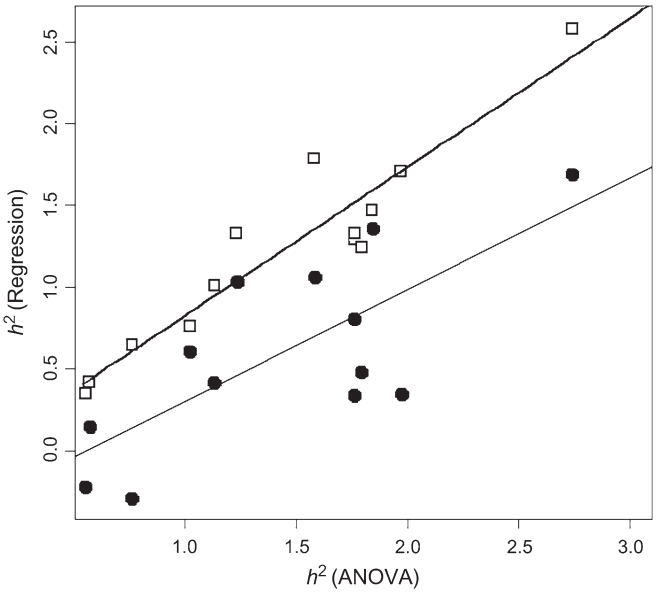

Fig. 2. Comparison of heritability estimates for *Prosopis alba* obtained by the regression method with the classical ANOVA method. Open squares correspond to estimates based on pairwise relationships from family records and full circles to pairwise relationships estimated by Ritland method.

desirable, making this study a first attempt to validate this in situ inference methodology.

As a general feature, heritability estimates obtained from marker-based estimated relatedness were lower than those obtained form genealogical data. Ritland (1996b) has already pointed out the risk of underestimation due to larger than expected sampling variation of molecular relationships as the regression variable. The author corrects this bias, at least partially, by using the actual variance of relatedness, which gives less-biased estimates of the population variance for relatedness. In our case, the actual variance of relatedness estimated from molecular markers was similar to the sampling variance of relatedness based on family information (0.002), but there might be other reasons behind the difference between marker-based and pedigree-based h^2 estimates.

Biased heritability estimates may be attributed to sampling sizes. Small number of families and individuals may result in increased estimate errors. However, it is not expected that this effect will produce the same bias for all analyzed traits as seems to be the case for ANOVA estimates, which in most cases produced h^2 estimates higher than one. Moreover, confidence intervals and standard errors for h^2 are acceptable and are quite similar for ANOVA and marker-based estimates.

A second factor is connected to the assumed relatedness between family members. As crosses are not controlled, pedigree-based h^2 estimates may be upwardly biased as a consequence of the presence of full sibs and selfs within fraternal groups. As stated, selfing was already detected at least in one natural population of this species (Bessega et al., 2000) and is as high as 28%. Moreover, in the present paper, we detected homozygote excess (positive and significant F_{IS}) in four of six microsatellite loci, which may be at least partially explained by inbreeding. An additional error source in the traditional approach is the assumption that individuals from different families are unrelated (i.e., $r = 0$). However, the consequence would have been an underestimation (rather than overestimation) of h^2 because similarities between related individuals alleged to different families would be attributed to nongenetic causes. In the case of marker-inferred relatedness, none of these issues would affect the estimates of h^2 because no prior assumptions of relatedness were made.

TABLE 5. Significant pairwise genetic correlations between quantitative traits estimated from the correlation between trait family means with their confidence intervals (CI) and the corresponding estimate from the linear model (Ritland, 1996b) based on molecular marker information for *Prosopis alba*.

Trait pair		Family mean	CI	Ritland
HEI	BMS	0.527*	0.220–0.740	1.003
HEI	TDI	0.630***	0.360–0.800	1.44
BMS	TDI	0.907***	0.820–0.950	0.994
BMS	PEL	0.35	0.005–0.625	0.842
TDI	PEL	0.41	0.076–0.666	1.281
PEL	PIL	0.573**	0.280–0.768	1.097
PEL	NPI	–0.568**	–0.765 – –0.273	0.361
PEL	LEL	0.654***	0.396–0.817	NA
PEL	LEL/LEW	0.609**	0.331–0.790	–0.655
PEL	LEX	–0.606**	–0.788 – –0.326	NA
NLP	PIL	0.41	0.069–0.662	NA
NLP	NPI	0.456*	0.128–0.694	NA
NLP	SPL	–0.458*	–0.695 – –0.130	NA
PIL	LEL	0.737***	0.523–0.864	NA
PIL	LEL/LEW	0.652***	0.392–0.815	–0.772
PIL	LEX	–0.497*	–0.721 – –0.180	NA
PIL	SPL	–0.41	–0.661 – –0.067	0.228
NPI	LEL	–0.694***	–0.839 – –0.455	NA
NPI	LEL/LEW	–0.513*	–0.731 – –0.200	0.128
NPI	LEX	0.577**	0.286–0.771	NA
LEL	LEL/LEW	0.772***	0.579–0.883	NA
LEL	LEX	–0.713***	–0.850 – –0.484	NA
LEL/LEW	LEX/LEA	–0.4	–0.659 – –0.064	NA
LEL/LEW	LEX	–0.497*	–0.721 – –0.179	NA
LEF	LEX/LEA	–0.491*	–0.717 – –0.173	NA
LEX/LEA	LEX	0.604**	0.323–0.787	NA

Notes: * Highly significant at individual level; ** matrix wide significant; *** matrix wide highly significant.

There might be a third cause for an upward bias that would affect both marker-based and pedigree-based h^2 estimates: the occurrence of geographical association within fraternal groups. When the experimental orchard was planted, a randomized complete block design had been applied (Felker et al., 2001). Each block comprised four contributions per family, planted together with 4×4 m spacing, thus sharing a common environment. Differences in environmental factors between family sets could have been somehow exacerbated in the current experimental layout because several blocks and families were lost, with resulting gaps increasing environmental differences among the surviving family sets. Because of that, low but highly significant correlation occurred between relatedness and geographical distances demonstrated by Mantel tests ($r = -0.10$, $P = 0$ and $r = -0.08$, $P = 0.001$, respectively, for genealogical or marker inferred relatedness, with 1000 permutations). Therefore, if relatives shared environments, some of the phenotypic resemblance between them could have been caused by common growing conditions. It should be noted, however, that in natural conditions this situation would be rare for outbreeding species, and relatives may not be found in such compact clusters. In that sense, the microsatellite analysis needed for marker-based inferences could serve as well to visualize the spatial distribution of relatives in the population under study and, therefore, indicate

whether inheritance inferences are pertinent or not, and if there would be a risk of overestimation.

In sum, the two most plausible causes for upward bias in ANOVA h^2 estimates for most traits may be the underestimation of relatedness and spatial covariance within family groups. The lower marker-based h^2 estimates with respect to those obtained from ANOVA may be due to two main causes. The first is that assumptions of relatedness within and between family groups are not expected to produce overestimations of h^2 by this method. The second is that relatedness is inferred indirectly, and a bias may occur as a consequence of a limited number of available molecular loci and alleles.

However, although the marker-based h^2 estimates are lower than those obtained from genealogical data, there is a highly significant correlation between estimates from the different approaches applied here; and the confidence interval of the regression slope of marker-inferred on ANOVA-estimated h^2 contains the expected value of one. This result suggests that, although the absolute values of marker-based h^2 may not be quantitatively accurate, these estimates are useful to rank the traits according to their actual differences in the proportion of additive variance. As stated by Ritland and Ritland (1996) for microstructured populations, with higher number of marker loci and proper sampling strategies, the precision and accuracy of this method might be greatly increased.

According to our results, marker-based estimates are more accurate for traits with high h^2 values. With low or moderate h^2 traits, the risk of retrieving nonsignificant results must be evaluated with higher number of markers. Our results are therefore promising given the outlined limitations of this preliminary study and may encourage the development of more molecular markers for this methodology, not only for *P. alba* but also for other profitable, related species of *Prosopis* to provide useful information to screen natural populations for their valuable genetic diversity in conservation and breeding programs.

LITERATURE CITED

Akindele, A., and O. Olutayo. 2007. Quantitative variations in the growth of progeny seedlings of *Prosopis africana* (Guill., Perrott. And Rich.) plus trees in Nigeria. *African Journal of Biotechnology* 6: 359–363.

Andrew, R. L., R. Peakall, I. R. Wallis, J. T. Wood, E. J. Knight, and W. J. Foley. 2005. Marker-based quantitative genetics in the wild?: The heritability and genetic correlation of chemical defenses in *Eucalyptus. Genetics* 171: 1989–1998.

Bassam, B. J., G. Caetano-Anollés, and P. M. Greshoff. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry* 196: 80–84.

Bessega, C., L. Ferreyra, N. Julio, S. Montoya, B. O. Saidman, and J. C. Vilardi. 2000. Mating system parameters in species of genus *Prosopis* (Leguminosae). *Hereditas* 132: 19–27.

Burkart, A. 1976. A monograph of the genus *Prosopis* (Leguminosae subfam. Mimosoidae). *Journal of the Arnold Arboretum* 57: 219–525.

Cannings, C., and A. Thompson. 1981. Genealogical and genetic structure. Cambridge University Press, Cambridge, Massachusetts, USA.

Cony, M. 1996. Genetic variability in *Prosopis flexuosa* D.C., a native tree of the Monte phytogeographic province, Argentina. *Forest Ecology and Management* 87: 41–49.

Felker, P., C. Lopez, C. Soulier, J. Ochoa, R. Abdala, and M. Ewens. 2001. Genetic evaluation of *Prosopis alba* (algarrobo) in Argentina for cloning elite trees. *Agroforestry Systems* 53: 65–76.

Felker, P., D. Smith, C. Wiesman, and R. L. Bingham. 1989. Biomass production of *Prosopis alba* clones at 2 non-irrigated field sites in semiarid south Texas. *Forest Ecology and Management* 29: 135–150.

Ferreyra, L. I., C. Bessega, J. C. Vilardi, and B. O. Saidman. 2007. Consistency of population genetics parameters estimated from isozyme and RAPDs dataset in species of genus *Prosopis* (Leguminosae, Mimosoidae). *Genetica* 131: 217–230.

Garant, D., and L. E. B. Kuruuk. 2005. How to use molecular data to measure evolutionary parameters in wild populations. *Molecular Ecology* 14: 1843–1859.

Guo, S. W., and E. A. Thompson. 1992. Performing the exact test of Hardy–Weinberg proportions for multiple alleles. *Biometrics* 48: 361–372.

Hardy, O. J. 2003. Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Molecular Ecology* 12: 1577–1588.

Jacquard, A. 1974. The genetic structure of populations, Springer, Berlin, Germany.

Kalinowski, S. T., and M. L. Taper. 2006. Maximum likelihood estimation of the frequency of null alleles at microsatellite loci. *Conservation Genetics* 7: 991–995.

Klaper, R., K. Ritland, T. A. Mousseau, and M. D. Hunter. 2001. Heritability of phenolic in *Quercus laevis* inferred using molecular markers. *Journal of Heredity* 92: 421–426.

Kumar, S., and T. E. Richardson. 2005. Inferring relatedness and heritability using molecular markers in radiata pine. *Molecular Breeding* 15: 55–64.

Li, C. C., D. E. Weeks, and A. Chakravarti. 1993. Similarity of DNA fingerprints due to chance and relatedness. *Human Heredity* 43: 45–52.

Lynch, M. 1988. Estimation of relatedness by DNA fingerprinting. *Molecular Biology and Evolution* 5: 584–599.

Lynch, M., and K. Ritland. 1999. Estimation of pairwise relatedness with molecular markers. *Genetics* 152: 1753–1766.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits, Sinauer, Sunderland, Massachusetts, USA.

Milligan, B. G. 2003. Maximum-likelihood estimation of relatedness. *Genetics* 163: 1153–1167.

Morton, N. E., S. Yee, D. E. Harris, and R. Lew. 1971. Bioassay of kinship. *Theoretical Population Biology* 2: 507–524.

Mottuora, M. C., R. Finkeldey, A. R. Verga, and O. Gailing. 2005. Development and characterization of microsatellite markers for *Prosopis chilensis* and *Prosopis flexuosa* and cross-species amplification. *Molecular Ecology Notes* 5: 487–489.

Mousseau, T. A., K. Ritland, and D. D. Heath. 1998. A novel method for estimating heritability using molecular markers. *Heredity* 80: 218–224.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583–590.

Pasiecznik, N. M., P. J. C. Harris, and S. J. Smith. 2004. Identifying tropical *Prosopis* species. A field guide. HDRA, Coventry, UK.

Queller, D. C., and K. F. Goodnight. 1989. Estimating relatedness using genetic markers. *Evolution* 43: 258–275.

R Development Core Team. 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Website http://www.R-project.org.

Ritland, K. 1996a. Estimators of pairwise relatedness and individual inbreeding coefficients. *Genetical Research* 67: 175–185.

Ritland, K. 1996b. A marker-based method for inferences about quantitative inheritance in natural populations. *Evolution* 50: 1062–1073.

Ritland, K. 2000. Marker-inferred relatedness as a tool for detecting heritability in nature. *Molecular Ecology* 9: 1195–1204.

Ritland, K. 2005. Multilocus estimation of pairwise relatedness with dominant markers. *Molecular Ecology* 14: 3157–3165.

Ritland, K., and C. Ritland. 1996. Inferences about quantitative inheritance based on natural population structure in the yellow monkeyflower, *Mimulus guttatus. Evolution* 50: 1074–1082.

Ritland, K., and S. Travis. 2004. Inferences involving individual coefficients of relatedness and inbreeding in natural populations of *Abies. Forest Ecology and Management* 197: 171–180.

Rousset, F., and M. Raymond. 1995. Testing heterozygote excess and deficiency. *Genetics* 140: 1413–1419.

RWEYONGEZA, D. M., F. C. YEH, AND N. K. DHIR. 2005. Heritability and correlations for biomass production and allocation in white spruce seedlings. *Silvae Genetica* 54: 4–5.

THOMAS, S. C., D. W. COLTMAN, AND M. PEMBERTON. 2002. The use of marker based relationship information to estimate the heritability of body weight in a natural population: A cautionary tale. *Journal of Evolutionary Biology* 15: 92–99.

THOMAS, S. C., J. M. PEMBERTON, AND W. G. HILL. 2000. Estimating variance components in natural populations using inferred relationships. *Heredity* 84: 427–436.

VERGA, A. 2000. Recursos genéticos, mejoramiento y conservación de especies del género *Prosopis*. *In* Carlos A. Norverto [ed.], Mejores árboles para más forestadores, 205–221. Secretaría de Agricultura, Ganadería, Pesca y Alimentos (SAGPyA), Buenos Aires, Argentina.

VERGA, A., M. MOTTUORA, M. MELCHIORE, J. JOUSEAU, C. CARRANZA, M. LEDESMA, D. RECALDE, AND L. TOMALINO. 2005. El proyecto algarrobo del INTA. Revista de información sobre investigación y desarrollo agropecuario. *Revista Idia XXI* 8: 195–200.

WANG, J. 2002. An estimator for pairwise relatedness using molecular markers. *Genetics* 160: 1203–1215.

WILSON, A. J., G. MCDONALD, H. K. MOGHADAM, C. M. HERBINGER, AND M. M. FERGUSON. 2003. Marker-assisted estimation of quantitative genetic parameters in rainbow trout, *Oncorhynchus mykiss. Genetics Research* 81: 145–156.

ZAS, R., AND J. FERNÁNDEZ-LÓPEZ. 2005. Juvenile genetic parameters and genotypic stability of *Pinus pinaster* Ait. open-pollinated families under different water and nutrient regimes. *Forest Science* 51: 165–174.