New Forests
DOI 10.1007/s11056-015-9469-5

Improving initial trials in tree breeding using kinship and breeding values estimated in the wild: the case of *Prosopis alba* in Argentina

C. Bessega · C. Pometti · M. Ewens · B. O. Saidman · J. C. Vilardi

Received: 19 March 2014 / Accepted: 2 February 2015

Abstract Most Argentinean forests have been lost by over exploitation and expansion of agricultural areas. The National breeding program for the native species *Prosopis alba* is still in its initial phase with only a few progeny trials installed from material collected in the wild and the first genetic studies are underway. Breeding value (BV) estimates based on pedigree data from a progeny trial were first compared to those derived by using microsatellite based kinship estimates to confirm the potential accuracy of g-best linear unbiased predictions (G-BLUP) when pedigree information is lacking. Afterwards, the possible genetic effects of alternative selection strategies to collect improved seeds from a wild population were evaluated. To achieve this goal the relationship among average genetic gain (predicted by G-BLUP), inbreeding and sampling size of the collected materials were weighed in a wild population consisting of trees of similar ages. The results obtained suggest that kinship estimates based on molecular data and breeding value predictions BV used for the selection of elite trees in wild populations may contribute to improve the genetic properties of the founder population. Controlling kinship allows reducing sampling size from 20 to 10 individuals per origin with no significant increase of inbreeding or loss of genetic gain. For a selected group of only ten top individuals per origin, the replacement of two strongly reduces the average group coancestry with minimal gain loss. Simultaneous selection for two traits by selection index might produce a gain of near 6 % in height and 2 % in diameter. The use of molecular marker information may contribute to reduce the time needed in a *P. alba* improvement program.

Keywords BLUP · Breeding program · Breeding value · Coancestry · Microsatellites

C. Bessega (✉) · C. Pometti · B. O. Saidman · J. C. Vilardi
Departamento de Ecología, Genética y Evolución, Instituto IEGEBA (CONICET-UBA), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina
e-mail: cecib@ege.fcen.uba.ar

M. Ewens
Estación Experimental Fernández, UCSE, Santiago del Estero, Argentina

Published online: 10 February 2015

Introduction

Traditional selection breeding implies repeated cycles where breeding, testing and selection are successively carried out (Namkoong et al. 1988). When each cycle is completed, the superior individuals are selected and then included in the subsequent breeding cycles where finally some elite genotypes are chosen to constitute seedling seed orchards and/or clonal seed orchards to produce abundant genetically improved seed crop for reforestation purposes (White and Neale 2007). While the improvement of economic traits is the main focus, it is also desirable to maintain a broad genetic diversity (El-Kassaby and Lstiburek 2009). This may be achieved by combining seeds from trees selected in several parallel assays following the multiple population breeding system (MPBS) approach (Namkoong et al. 1988). The whole process is costly and time-demanding especially in the case of forest species.

An alternative novel focus in tree breeding suggests the bypass of the conventional breeding via a simplified testing of either open pollinated Half-sib families or populations (El-Kassaby and Lstiburek 2009; Lstiburek et al. 2012). This method is based on the molecular fingerprinting of individuals to establish paternity or full pedigree reconstruction and the subsequent quantitative analysis to detect elite genotypes (El-Kassaby and Lstiburek 2009; El-Kassaby et al. 2012; Lstiburek et al. 2012). This approach is especially appealing for domestication of promising native species naturally adapted to environments which are not favorable for growing traditional forest species.

The "algarrobo blanco" (white mesquite) *Prosopis alba* Griseb. (Leguminosae, Mimosoidae) is a native forest tree widely distributed in Argentina, Uruguay, Paraguay and from the south of Bolivia to Peru. Its wood is of high quality (Pometti et al. 2009) and is used for carpentry and floor construction. Its fruits are considered a good foraging resource while they can also be used by humans (Roig 1993), rendering the species as very valuable for rural economies in Argentina. It is considered the most important species of this genus from an economic perspective due to its wide distribution in a great variety of environments that resulted in an extended genetic variation that can be used in forest tree breeding programs (Salto 2011).

Unfortunately, in the Chaco ecoregion of Argentina most of the original *Prosopis* forests have already been lost (FAO 2007). The effects of natural causes such as fires and periodical floods are minimal in comparison to man induced disturbances. The most important damage is caused by non-sustainable exploitation of natural forests stimulated by the multiple uses of "algarrobo" wood (Roig 1993) and the expansion of agricultural areas (Adamoli et al. 2011). In the last 35 years 7,630,000 ha of native forests have been converted into cultivated agro systems. About 85 % of the agricultural expansion is concentrated in the provinces of Santiago del Estero, Córdoba, Salta, and Chaco (Adamoli et al. 2011). Argentinean furniture industry consumes 100,000 tons of *P. alba* logs per year from existing wild stands, but only a few plantations from a total area of 20–30 ha exist (Felker and Guevara 2003). The consequences of non-sustainable exploitation are dramatic forest fragmentation and desertification of wide areas, endangering the resources of forest tree species.

The strategy adopted in *Prosopis* breeding programs in Argentina is based on a modification of the method proposed by Namkoong et al. (1988) for multiple populations (Verga 2005). *P. alba* in Argentina follows the multiple population breeding system (MPBS) and the breeding program is still in its initial stage. The most important goals of this program are improving height, diameter, stem shape and growing rate, although insect attack and tolerance to soil salinity are also under consideration. So far only some progeny

trials have been installed in the Provinces of Córdoba, Santiago del Estero, Chaco, Salta, and Formosa. Each progeny trial involves open pollinated half sib families sampled from selected plus trees growing in the wild (Lopez et al. 2001; Delvalle et al. 2003; Verga 2009). Furthermore, in order to capture genetic variation, all progeny trials included several (8–10) origins. The number of origins and families per origin differs among trials. Currently, these trials are under evaluation in order to identify the best trees to provide improved seeds or to be asexually propagated to establish seedling seed orchards or clonal seed orchards. As *P. alba* trials were not established from material genetically improved, the whole process is expected to take many years. Although genetic gains in clonal seed orchards are expected to occur faster than in seedling seed orchards (Felker and Guevara 2003), so far there is only one clonal orchard under evaluation with a narrow genetic basis. A cause for this is that high variability exists in *P. alba* clones for rooting ability (Felker and Guevara 2003; Felker et al. 2008). Only 21 *P. alba* genotypes were asexually propagated by rooting cuttings using bottom heat, water soluble salts of plant hormones and *Agrobacterium rhizogenes* (Ewens et al. 2012).

When the first progeny trials were installed, the number of selected plus trees for the production of Half-sib families was relatively low (6–20), while selection of those plus trees in the wild was not supported by genetic criteria to weigh, to a certain extent, possible side effects of sample size on the level of inbreeding of the selected group. Furthermore, plus trees were defined by taking into account several traits as growth, vigour, proportion between trunk and crown, proportion of stem without branches, length and stem straightness and health condition, without the use of any statistical criteria for multiple trait selection, which might take into account possible correlations between traits that may affect the gain in those considered as key breeding objectives.

A novel approach to improve the initial process of selecting plus trees by applying genetic (rather than only phenotypic) criteria takes advantage of molecular markers. El-Kassaby et al. (2012) proposed using the genomic best linear unbiased predictor (G-BLUP) to obtain individual breeding values. The method uses the animal model (BLUP) replacing the pedigree matrix by a relatedness matrix (G) based on molecular fingerprinting. With this method genetic diversity can be maximized while optimizing the gain of the trait under study. In applying this strategy it is valuable to consider the level of inbreeding present in the selected group because selection over several generations can lead to increase of inbreeding and a narrowing down of the genetic basis (Burdon and Shelbourne 1971). All artificial selection strategies begin by ranking available candidates based upon some estimation of their relative genetic merits. The expected gains may be estimated by BLUP breeding value (BV), but usually the selection of candidate trees involves some constraints on relatedness among individual genotypes. One criterion to cope with the balance between gain and inbreeding is based on optimal effective population size for breeding and production populations (Weir and Lindgren 1996; Lindgren and Mullin 1997; Anderson et al. 1999; Olsson et al. 2001). Brisbane and Gibson (1995) proposed a method that maximizes the genetic gain for a given level of inbreeding based on group merit selection (GMS), which incorporates breeding value and coancestry into a single selection criterion that penalizes the increase of average coancestry in the selected group GMS. Stoehr et al. (2008) evaluated different selection scenarios in Douglas fir and concluded that a method maximizing gain under explicit constraints on coancestry for seed orchard construction results in the production of seedlings with the potential for the highest merchantable volume at rotation time.

Most forest breeding programs are designed for single trait selection (Gill 1965; Namkoong et al. 1988), but importance should also be given to multiple trait selection

(Sanchez et al. 2008; Yanchuk and Sanchez 2011). In the latter case, a constraint in the initial selection step is related with the relative gains for traits that are or can be correlated. The merits should be defined by combining data from different traits into a single index (selection index) (Hazel 1943). The main difficulty that arises when several traits are selected is determining the appropriate weight value of the selection index used for each particular trait (White and Neale 2007). Potential criteria are (1) relative economic importance and (2) desired gain production for each trait.

In this paper, we first compared BV predictions based on BLUP and G-BLUP using data from a progeny trial previously studied by Bessega et al. (2009, 2010) to test the accuracy of G-BLUP. Then we considered the potential advantages of combining phenotypic and molecular information to select plus trees from the wild to establish new initial progeny trials for a breeding program aiming to improve a native forest resource. As a case study, we applied this approach to a wild population of *P. alba* in Santiago del Estero by considering several alternative selection scenarios. We took into account the relation of sample size with the average breeding value and average kinship of the breeding group as well as the strategies of selecting for a single trait (height) or for two (height and basal diameter) correlated traits of economic importance. The potential of this strategy to be applied on other species during initial stages of low-input breeding is discussed below.

Materials and methods

Testing G-BLUP accuracy in a half sib progeny trial

Plantation

The plantation analyzed to validate the G-BLUP estimation is the same as the one studied in Bessega et al. (2009, 2010). It is a *P. alba* progeny trial established in 1990, 10 km from the City of Santiago del Estero in Argentina (27°45′S; 64°15′W) (Felker et al. 2001). This trial was established with seeds collected from eight northwestern populations in Argentina; (Añatuya, Castelli, Gato Colorado, Ibarreta, Pinto, Quimili, Rio Dulce Irrigation District and Sumampa). The experimental design was a randomized complete block comprising 57 open-pollinated families, seven blocks and four trees per block (with a 4 by 4 m spacing). The total planting material was 1,596 individual trees, out of which 1,289 were still alive in 1999. During the last 15 years, this trial was affected by natural conditions without any silvicultural care. For this study, we obtained the BV of the same 142 individuals from 32 different open pollinated families used in Bessega et al. (2009, 2010) to estimate genetic parameters for quantitative traits. Due to the lack of silvicultural care, there was high mortality at the moment of data collection and survivors of each family were not represented in all blocks. Thus, the block effect could not be included and the model was unbalanced with the number of individuals per family varying from 3 to 12. For a thorough description of experimental design and the analysis of variance component design see Bessega et al. (2009, 2010).

Morphometric data

Three biomass traits and ten leaf morphology traits were analyzed. Height (H) and basal trunk diameter (BD) (at 20 cm above ground level) were scored in the field. As the slope in

the studied area is negligible, it did not affect the BD measures. Biomass predictions (BMS) of each tree were estimated by using the regression equation $BMS = \log_{wt} = 2.7027 \times \log_{DB} - 1.1085$ (Felker et al. 1989), where $\log_{wt}$ is the logarithm of fresh weight (kg) and log DB is the logarithm of DB (cm).

To measure leaf morphology traits, samples from each tree were first mounted on specimen boards. Measured traits were petiole length (PEL), number of pairs of leaflets per pinnate (NLP), pinnate 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). In each individual, nine repeats of PEL, NLP, PIL, SPL, and NPI were obtained involving boards from three different canopy regions (lower, medium and higher third). Ninety repeats were obtained of LEL, LEL/LEW, LEF, LEX, LEX/LEA, involving ten 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, "Instituto de Fisiología y Recursos Genéticos Vegetales, INTA, ARGENTINA".

Reconstructed relatedness matrices

To estimate breeding values, genetic analysis was performed with relatedness inferred from either the expected pedigree in the case of an open pollinated progeny trial or molecular data in the case of a wild population. The trial was assumed to be composed of half sib families because outcrossing rate in *P. alba* is $t \approx 1$ (Bessega et al. 2000, 2011). Consequently, the "potentially real" relatedness matrix was constructed with three alternative values: $r = 1$ for individuals with themselves, $r = 0.25$ for different individuals of the same family, and $r = 0$ for individuals from different families.

For the construction of the molecular relatedness matrix, six microsatellites were used following Mottura et al. (2005). Bessega et al. (2009) obtained pairwise Ritland's (1996a) relatedness estimator using the software HERINAT (a Visual Fortran program designed ad hoc, available from author, leopoldo.sanchez@orleans.inra.fr).

Breeding value predictions

Breeding values (BV) for each trait were estimated by the best linear unbiased prediction method (BLUP), which is based on the following mixed model:

$$Y = X\beta + Zg + \varepsilon$$

where Y is a vector of observations of all individuals, β is a vector of fixed effects, X represents a design matrix (of 0 and 1 s) relating the appropriate fixed effects to each individual, g is a vector of random effects (BV), Z is a design matrix relating the appropriate random effects to each individual and ε is a vector of residual errors (Kruuk 2004). In our case, block fixed effects could not be included because in the surviving stand several families were present in single blocks and the random effect is represented by the family (mother plant) only. The general mixed model is thus reduced to a random model as follows:

$$Y = Zg + \varepsilon$$

According to this equation the vector of g (BV) can be obtained as:

$$\hat{g} = AZ'\sigma_A^2 V^{-1} Y$$

In this equation A represents the relatedness matrix, Z' is the transposed Z, σ_A^2 is the additive genetic variance, and V^{-1} is the inverse of the phenotypic variance. The product $\sigma_A^2 V^{-1}$ is by definition the parameter of heritability (h^2) for each trait.

Genomic predictions were obtained by ridge regression using the package rrBLUP (Endelman 2011; Endelman and Jannink 2012) of R (R Core Team 2013), with A matrix obtained using pedigree based (P-BLUP) or molecular based (G-BLUP) relatedness matrices. The function *kin.blup* of rrBLUP requires the phenotype (trait) measured in each individual and the kinship matrix. The package calculates BV and the prediction error variance (PEV) of each estimated BV as the square of the standard error of the BLUPs. The expected accuracy of BLUP predictions (*ri*) were estimated according to (Clark et al. 2012):

$$ri = \sqrt{1 - \frac{PEV}{V_g \times K_{ii}}},$$

where K_{ii} is the diagonal of the relatedness matrix for individual i.

Kinships and breeding value predictions in the wild

Natural population and sampling design

A natural population of *P. alba* was sampled in December 2009 in the region Fernandez-Forres following road 34 near Santiago del Estero, Argentina (from 27°51.31′S, 63°59.88′W to 27°53.80′S 63°54.83′W; Fig. 1). This area originally involved a very dense natural forest where dominant species were hardwood trees (*Schinopsis* sp., *P. alba*, *P. nigra*, *P. kuntzei*). The region has become highly fragmented mainly by over exploitation and the expansion of the agricultural land frontier.

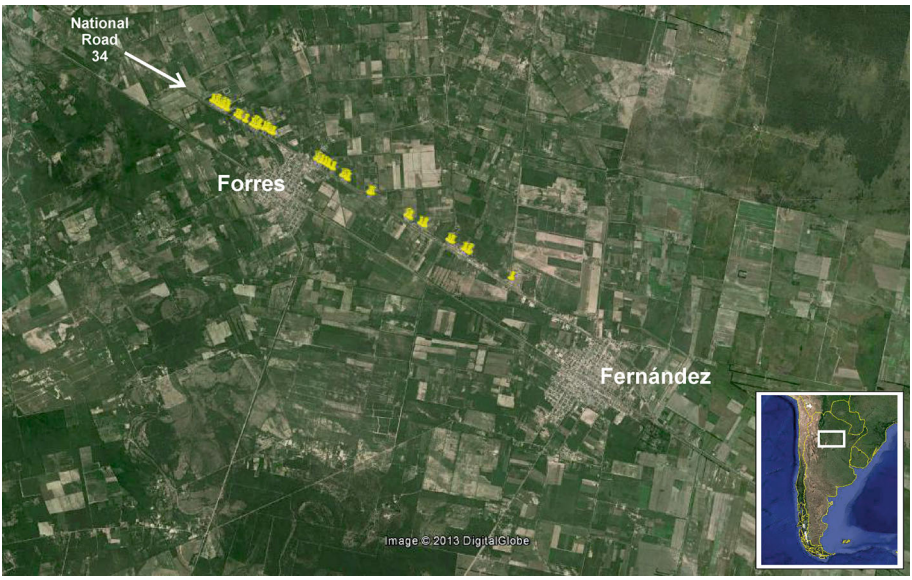

Fig. 1 Spatial distribution of the sampled *Prosopis alba* trees at road 34, Santiago del Estero, Argentina

In the sampled area it is possible to assume that most trees are about the same age. The population extends on both sides of a road constructed in the 60 s. Along both sides of the road all vegetation was felled to a width of about 70 m during the 70 s. After this period the population recovered from preexisting seeds or reintroductions favoured by livestock drive. By the end of the 80 s most of the road sides were colonized by "algarrobos" (mesquites) and other shrubs. Since the 90 s, road sides have been managed and small shrubs and trees have been removed, keeping bigger and healthier trees.

Previous studies (Bessega et al. 2000, 2011) on genetic structure and mating system in *P. alba* and other species from section Algarobia demonstrated that the pollen and seed dispersal is limited, determining that the trees located within 50 m from each other are highly related. Based on this reason, it was recommended for conservation purposes and genetic variation estimation to sample trees separated more than 50 m from each other in order to avoid sampling of genetically related material. The application of the recommended strategy for *Prosopis* collection determined to sample 87 mother trees separated from each other by more than 50 m in the sampling area. Although the genetic variance estimated in a sample depends directly on the number of individuals analyzed, the strategy applied in this study is aiming at reducing the inbreeding within the sample in comparison to a random sampling. Furthermore, additive genetic variance and covariance in metrical traits is expected to be reduced at a rate of $1/(2N_e)$ per generation by random sampling of genes (Lande 1979, 1980). With 87 randomly sampled mother trees, the expected loss of additive variance is 0.57 %. Moreover, the constraint applied to reduce the inbreeding in the sample determines that the effect on additive genetic variance should be even lower than the theoretical value.

Trait measurement and breeding value estimation

Two economically important traits, height (H) and basal diameter (BD) of the 87 sampled trees were scored in the wild. H was measured with a telescopic rod at the nearest decimeter and BD with a tape measure at the nearest centimeter. The same trees were genotyped using a total of 12 microsatellites, six described by Mottura et al. (2005) and six by Bessega et al. (2013), following Bessega et al. (2009). Leaves were collected from each tree and preserved in silica-gel. DNA was extracted using DNA easy plant mini kit (Qiagen, Valencia, California, USA). The PCR amplifications were carried out in a 50 μL reaction volume containing 10–30 ng DNA, 0.6 μM each primer, 0.2 mM dNTPs, 0.3 U Taq DNA Polymerase (Invitrogen, Carlsbad, California, USA), and 1.5 mM $MgCl_2$. A MyCycler thermalcycler (BioRAD Laboratories) 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. PCR products were electrophoresed in an ABI313XL (HITACHI) automated DNA sequencer and automatically sized using GENEMARKER ver 1.91 (SoftGenetics LLC™ www.softGenetics.com).

Classic genetic diversity parameters were estimated using the software SPAGEDI (Hardy and Vekemans 2002). From the multilocus genotypes a pairwise relatedness matrix was obtained using Ritland estimator (1996a) with the software Coancestry (Wang 2011).

In the wild population, G-BLUP was applied to get BV estimates with the package rrBLUP (Endelman 2011; Endelman and Jannink 2012). In this case we used the information of individual phenotypes of the 87 sampled trees and the relatedness matrix among these individuals was inferred from molecular markers.

In order to test possible bias in BV prediction associated to the relatedness estimation methods, we also applied a method of pedigree reconstruction based on simulated annealing (Fernandez and Toro 2006; Cros et al. 2014) implemented in the software MOLCOANC ver 3.0 (Fernandez and Toro 2006). In order to obtain a pedigree that best explains the molecular kinship matrix, several runs were conducted considering one or two previous generations and 50–1,500 ancestors. The best result was obtained with two previous generations involving 150 and 100 parents for generations one and two respectively. The solution coancestry matrix was compared with Ritland's relatedness estimator by Mantel test, using the package ade4 (Dray and Dufour 2007; Dray et al. 2007; Chessel et al. 2004) of R. Pearson's correlation coefficient between BV predictions obtained using Ritland's relatedness matrix and the kinship matrix estimated by Fernandez and Toro (2006) was analyzed with the package *stats* of R.

The heritability of H and DB was estimated by two alternative methods. The first one is based on the results from *kin.blup* function of rrBLUP (Endelman 2011; Endelman and Jannink 2012) package.

Heritability was estimated as $h^2 = Vg/(Ve + Vg)$, where Vg and Ve correspond to the REML estimates of genetic and error variances obtained with this package, using Ritland relatedness matrix.

The significance of this h^2 estimate was tested by two methods. The first one was a Mantel permutation test between the matrix of phenotypic similarities (Z_{ij}) and the matrix of estimated relatedness (r_{ij}). Z_{ij} are defined by (Ritland 1996b) as:

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

where Y_i and Y_j give the individual trait values for individual i and j respectively and U and V are, respectively, the corresponding mean and variance of the sample.

The second method was a regression analysis of BV estimates on phenotypic values (as h^2 is expected to equal the slope of the regression). Regression analysis and Mantel test were conducted with the program R (R Core Team 2013), using the packages *stats* and *ade4* (Dray and Dufour 2007; Dray et al. 2007; Chessel et al. 2004), respectively.

The second approach to estimate h^2 was an analysis of variance component, using a Markov chain Monte Carlo Sampler for Multivariate Generalised Linear Mixed Models implemented in the package MCMCglmm (Hadfield 2010). Basically, this method requires as data the vector of phenotypic values and a pedigree. In this case the pedigree was replaced by a phylogenetic object obtained from the molecular based relatedness matrix, using the package ape (Paradis et al. 2004) of the program R. The conditions used in MCMCglmm runs were 400.000 iterations, thinning = 180, and burnin = 40.000.

Posterior heritability was estimated as the ratio $h^2 = Vg/(Ve + Vg)$ where Vg and Ve represent, respectively, posterior modes of genetic and residual variances. The significance of h^2 was tested by the deviance information criteria (DIC) when comparing the model including both genetic and residual variance with a model considering only the residual component.

Relationships between average breeding value and average kinship of the selected group

The individuals were ranked according to their height breeding values. For the following analysis the top 20 individuals were selected. We obtained all possible combinations of these 20 individuals by taking 10 individuals at a time. In this way we obtained 184,756 groups each consisting of individuals randomly chosen from the 20 top trees.

For each group we estimated the average BV, the percent gain relative to the phenotypic mean (as the average BV over phenotypic mean × 100) and the relative merit (BV_{mod}), defined by penalizing BV criteria by the increase in average kinship of the corresponding group (θ_{sel}), according to the expression (Lindgren and Mullin 1997):

$$BV_{mod} = BVsel - \delta\theta_{sel}$$

where δ is a weight given to θ_{sel}. When $\delta = 0$ the only selection criterion is the average BV whereas positive infinite δ value minimizes kinship. We explored the consequences of varying δ values on gain and average kinship of the selected group. The δ values considered were 0, 0.25, 0.5, 0.75, 1, 2, and 5). The maximum δ considered here was five because, empirically, we determined that average kinship of the group with the maximum merit was virtually zero.

Effect of selected group size on average breeding value and average kinship with different merit criteria

From the 20 best ranked individuals on height BVs, all possible groups composed of 10, 13, 16 and 20 individuals were considered. The numbers of possible combinations of 20 individuals for each sampling size considered (10, 13, 16 and 20 individuals) were respectively 184,756; 77,520; 4,845 and 1. These groups were ranked according to two criteria: (1) conventional selection, using average BV only without parental restrictions and (2) relative merits (BV_{mod}), penalizing kinship with $\delta = 0.5$. In both cases average BV and kinships were related with the selected group size.

Effect of selection by index on the gains of two traits

A selection mixed index (I) was proposed that considers relative contribution or weight (w), of height and basal diameter as follows (White and Neale 2007):

$$I = w \times \text{height-gain} + (1 - w) \times \text{basal diameter-gain},$$

where height-gain and basal diameter-gain are the BLUP predicted breeding values for each trait. Relative weight (w) can take values from 0 to 1.

These gains were expressed as a percentage of the maximum value possible for each trait if the trees were selected solely on the basis of the corresponding trait.

Results

Test of G-BLUP accuracy in the San Carlos progeny trial

The correlations between BV estimates obtained by G-BLUP and P-BLUP were high and highly significant (average $r = 0.765$, $SD = 0.107$) for all the traits ranging from $r = 0.577$ for Biomass to $r = 0.917$ for number of pinnae. These results indicate that BVs estimated by G-BLUP without family information are close to those produced by P-BLUP based on pedigree information. Following the results above and despite the drawbacks that may arise due to the small sample size, there is an indication that G-BLUPs might potentially be a quite accurate estimation of BVs in a wild population and thus, potentially be

used as selection criterion. Results derived from large sample size are needed to verify the above potential use and conclusion.

Natural population

Molecular relatedness estimation

The SSR loci analyzed in the wild population of Fernandez were highly variable showing from 5 to 23 alleles (Table 1). The measures of genetic diversity were similar to estimates based on SSR in other forest trees (Assoumane et al. 2009; Degen et al. 1999; Boys et al. 2005; Rojas et al. 2011) and much higher than isozyme variability estimates recorded in *P. alba* (Ferreyra et al. 2007). A positive average F_{IS} indicates a low excess of homozygotes compatible with population substructure and/or low rates of selfing as previously described for *P. alba* (Bessega et al. 2011).

The program MOLCOANC produces an input molecular kinship matrix from data and a solution kinship matrix calculated from the estimated genealogy. Correlation between input and solution coancestry matrices was high and highly significant ($r = 0.84$, $p < 5 \times 10^{-4}$ based on 2,000 permutations). The correlations of Ritland relatedness matrix and the kinship matrices produced by MOLCOANC were in both cases highly significant ($p < 5 \times 10^{-4}$ based on 2,000 permutations), but Ritland relatedness was more correlated with the input ($r = 0.505$) than with the solution MOLCOANC matrix ($r = 0.28$). These results indicate that pedigree reconstruction is associated with a distortion of the original molecular estimated relatedness matrix.

Breeding value and heritability estimation

In the case of the natural population we assumed that all measures were taken from even aged trees. Although this assumption can produce a bias in the estimates of breeding values and variance components, it was supported by historical information as indicated above. Height (H) and basal diameter (DB) were highly variable with an about sixfold difference between minimum and maximum sizes (Table 2). Scaling the standard deviation relative to

Table 1 Genetic diversity parameters estimated for the 12 microsatellites loci in the wild population (Fernandez-Forres)

Locus	*A*	*ne*	*rg*	*He*	F_{IS}
Mo08	6	1.94	5.91	0.4844	0.145
Mo09	5	1.26	4.99	0.2059	0.17
MO16	10	3.62	9.99	0.7238	0.051
MO05	6	1.96	5.99	0.4908	0.426
Mo13	10	3.14	9.73	0.6814	−0.144
Mo07	11	2.51	11	0.6009	0.624
GL18	11	2.8	10.66	0.6426	0.26
GL8	17	4.84	16.52	0.7936	−0.002
GL6	10	2.69	9.65	0.6283	−0.238
GL12	9	4.15	8.77	0.7589	−0.099
GL15	16	5.4	16	0.815	−0.08
GL24	23	12.15	22.44	0.9177	−0.064
Average	11.17	3.87	10.97	0.6453	0.058

A denotes number of alleles, *ne* the effective number of alleles (Nielsen et al. 2003), *rg* the allelic richness (expected number of alleles among 150 gene copies), *He* the heterozygosity corrected for sample size (Nei 1978) and F_{IS} the individual inbreeding coefficient

Table 2 Phenotypic values (P), Breeding values (BV) and expected accuracy of BLUP (r_i) for the 87 trees sampled in a natural population of *Prosopis alba* of Fernandez-Forres, Santiago del Estero

ID	Height			Basal diameter		
	P (m)	BV	r_i	P (cm)	BV	r_i
115	8.0	−0.093	0.482	133.0	−1.899	0.226
116	5.0	−0.301	0.471	94.0	−1.363	0.220
117	3.2	−0.428	0.411	38.0	−2.764	0.187
118	3.4	−0.483	0.415	38.0	−1.875	0.190
119	3.0	−0.665	0.477	39.0	−2.827	0.225
120	6.0	−0.310	0.425	75.0	−1.685	0.195
121	7.2	−0.042	0.434	140.0	0.574	0.201
122	7.0	−0.194	0.504	90.0	−2.969	0.237
123	4.6	−0.594	0.440	76.0	−3.121	0.204
124	4.5	−0.339	0.438	35.0	−2.271	0.203
125	4.5	−0.396	0.428	73.5	−2.367	0.199
126	7.5	−0.031	0.506	116.5	−1.631	0.240
127	5.7	−0.419	0.464	91.0	−1.520	0.218
128	7.0	−0.065	0.395	183.0	1.648	0.179
129	5.5	−0.177	0.398	41.0	−1.393	0.181
130	8.0	0.102	0.481	144.0	0.493	0.228
131	5.5	−0.271	0.411	110.0	−0.240	0.187
132	9.0	0.138	0.411	150.0	0.573	0.188
133	6.5	0.086	0.402	37.0	−1.099	0.183
134	7.8	0.049	0.424	100.0	−2.174	0.195
135	7.5	0.028	0.494	104.5	0.612	0.233
136	7.3	−0.057	0.422	160.0	0.536	0.194
137	7.5	−0.222	0.418	240.0	1.641	0.193
138	8.0	0.160	0.469	150.0	1.856	0.219
139	6.2	−0.018	0.400	117.0	0.383	0.181
140	7.0	−0.352	0.472	138.0	−2.177	0.222
141	7.5	0.058	0.428	155.0	0.161	0.196
142	7.5	−0.095	0.429	83.0	−0.786	0.197
143	6.3	0.027	0.410	78.0	−0.806	0.187
144	6.5	−0.465	0.534	92.0	−1.777	0.259
145	9.0	0.406	0.401	140.0	0.945	0.183
146	7.2	0.037	0.416	112.0	0.756	0.191
147	7.0	0.149	0.505	160.0	1.920	0.242
148	6.0	−0.213	0.464	100.0	−1.515	0.216
149	6.0	−0.227	0.422	95.0	−1.533	0.194
150	6.0	−0.255	0.430	150.0	0.130	0.199
151	7.8	−0.223	0.472	96.0	−2.342	0.221
152	8.0	0.030	0.452	110.0	1.172	0.211
153	8.0	0.122	0.390	164.0	0.870	0.176
154	8.0	0.273	0.469	155.0	2.724	0.217
155	11.6	0.505	0.396	202.5	2.900	0.179

Table 2 continued

ID	Height			Basal diameter		
	P (m)	BV	r_i	P (cm)	BV	r_i
156	7.5	−0.070	0.414	150.0	0.224	0.188
157	8.5	0.111	0.419	114.0	0.235	0.193
158	10.5	0.259	0.413	212.0	1.792	0.188
159	11.5	0.766	0.420	200.0	3.768	0.194
160	9.0	0.021	0.401	153.0	−0.157	0.183
161	8.0	−0.022	0.415	150.0	−0.392	0.190
162	10.5	0.347	0.441	130.0	0.348	0.205
163	10.0	0.599	0.493	216.0	3.172	0.231
164	9.5	0.240	0.423	214.0	0.197	0.194
165	8.5	0.151	0.421	160.0	0.812	0.193
166	7.5	0.430	0.442	240.0	3.598	0.205
167	5.3	−0.449	0.420	119.0	−0.953	0.193
168	8.0	0.040	0.415	197.0	1.389	0.190
169	6.5	−0.054	0.435	136.0	0.726	0.201
170	8.5	0.308	0.398	93.0	−0.041	0.180
171	8.0	0.242	0.410	198.0	2.538	0.187
172	9.5	0.492	0.403	153.0	1.960	0.184
173	8.5	−0.040	0.477	161.0	0.590	0.225
174	4.0	−0.424	0.450	44.0	−1.113	0.211
175	4.0	−0.262	0.493	44.0	−0.468	0.233
176	9.0	0.301	0.420	190.0	2.763	0.194
177	8.2	−0.196	0.513	147.0	−1.898	0.243
178	10.0	0.180	0.500	185.0	1.631	0.236
179	7.0	0.426	0.497	100.0	0.789	0.234
180	10.5	0.245	0.456	145.5	−1.029	0.216
181	8.5	−0.031	0.431	183.0	−0.495	0.200
182	10.0	0.255	0.506	203.0	0.503	0.242
183	7.5	0.134	0.394	200.0	2.234	0.179
185	7.6	0.386	0.454	111.0	0.509	0.212
186	2.5	−0.476	0.396	91.0	−0.091	0.180
187	9.2	0.262	0.405	175.0	2.226	0.184
188	6.0	−0.146	0.433	46.0	−2.440	0.199
189	9.4	0.619	0.399	136.0	0.876	0.182
190	15.0	1.112	0.394	53.0	−1.148	0.179
191	10.0	0.329	0.397	155.0	0.389	0.181
192	7.0	−0.249	0.429	84.0	−3.258	0.199
193	7.7	−0.050	0.416	152.0	0.584	0.190
194	9.5	0.393	0.419	240.0	2.969	0.193
195	8.0	−0.365	0.487	87.0	−4.525	0.231
196	7.5	0.286	0.426	184.0	2.677	0.194
197	8.0	0.171	0.445	155.0	3.694	0.206
198	8.2	0.219	0.425	196.0	3.063	0.196

Table 2 continued

ID	Height			Basal diameter		
	P (m)	BV	r_i	P (cm)	BV	r_i
199	8.0	0.240	0.504	122.0	1.422	0.241
200	8.5	−0.169	0.441	240.0	1.103	0.206
201	10.0	0.307	0.443	255.0	3.164	0.207
202	6.3	−0.160	0.430	122.0	−0.333	0.200
Average (SE)	7.554 (0.219)	0.022 (0.034)	–	133.879 (5.898)	0.131 (0.199)	–

the mean trait, DB exhibits higher variation than H (CV = 0.41 and 0.27 respectively). In spite of the relatively low correlation between Ritland relatedness matrix and the solution matrix from MOLCOANC, the correlation of BV estimated using both matrices was very high for both H and DB ($r > 0.80$, $p < 2 \times 10^{-16}$). As results attained with different approaches are very consistent, for simplicity hereafter, all analyses are based on BV predictions derived from Ritland's relatedenss estimator.

Individual BVs estimated by rrBLUP varied between −0.665 and 1.112 for H, with 13 out of 87 individuals higher than one SD. For DB, values ranged between −4.525 and 3.768 with 16 individuals over one SD. The expected accuracy (*ri*) of BV predictions ranged between 0.390 and 0.534 for H and between 0.176 and 0.259 for DB (Table 2).

Individual BVs for H and DB tend to show the same sign in 69 out of 87 cases and statistical analyses demonstrated highly significant correlation between these traits both for phenotypic values ($R = 0.59$ $p = 2.10^{-9}$) and BV's ($R = 0.68$ $p = 6\ 10^{-13}$).

According to rrBLUP the variance components for H were Vg = 0.541 and Ve = 3.624, giving an estimated heritability of $h^2 = 0.130$. For DB, variance components were Vg = 75.988 and Ve = 2,948.058, with $h^2 = 0.025$. The regressions of BV estimates on phenotypic values were highly significant for both traits (Fig. 2; $R^2 = 0.709$, $p = 2 \times 10^{-16}$ for H and $R^2 = 0.590$, $p = 2 \times 10^{-16}$ for DB) which suggests that the heritabilities, although low, are significant for both traits.

The analysis of variance components by glmmMCMC also suggested that the heritability would be significant for both traits. *Vg* and *Ve* estimates for H were 0.802 and 2.916 respectively. The inclusion of a genetic component was supported by a lower DIC (349.4) than the model excluding genetic variance (DIC = 374.2). The estimated h^2 was 0.191 (SE = 0.007, IC 0.049–0.799). For DB variance components were $Vg = 840.49$ and $Ve = 1{,}728.31$. In this case the model was also supported by DIC criterion (DIC = 922.07 and 947.68 respectively for the model, considering both components and the model excluding *Vg*). In the case of DB, $h^2 = 0.293$ (SE = 0.007, IC = 0.096–0.844).

Relationships between average breeding value of H and average kinship of the selected group

We evaluated the relationship between average kinship and BV for all (184,756) possible groups of 10 individuals randomly chosen from the top 20 individuals (Fig. 3). The average kinship increases significantly with average BV ($r = 0.326$, $p < 2 \times 10^{-16}$). This result implies that selecting for BV has an unwanted side effect on the inbreeding of the selected group.

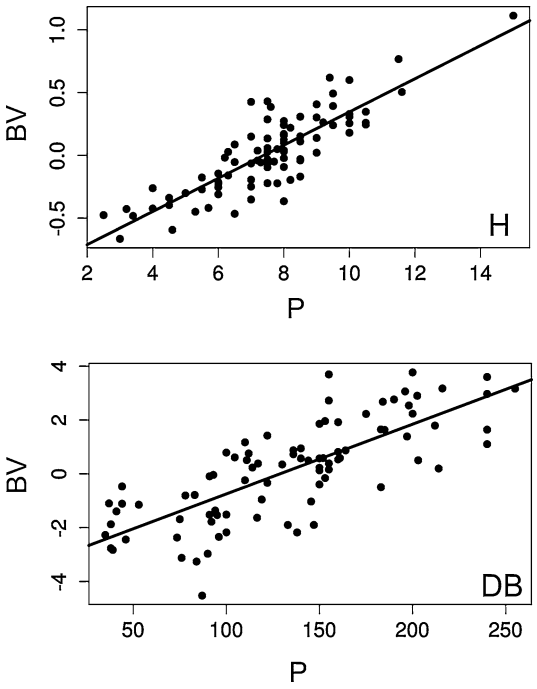

Fig. 2 Regression of breeding values (BV) on phenotypic values (P) for height (H) and basal diameter (DB)

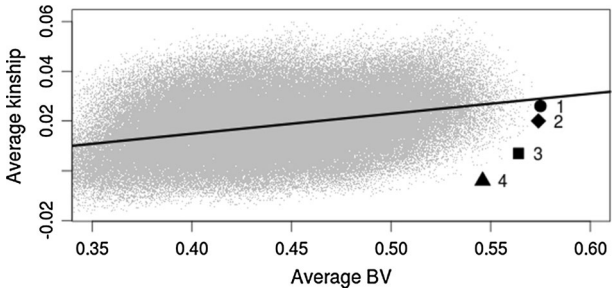

Fig. 3 Average kinship on Average breeding value considering $\delta = 0$. The 4 selected groups from Table 4 are indicated

When the merit of the selected group penalizes the increase of average kinship, four different individual groups may be selected depending on the δ applied (Table 3; Fig. 3): group 1 was the top for $\delta = 0$, group 2 was the top for $\delta = 0.25$, 0.50, and 0.75, group 3 had the highest merit for $\delta = 1$ and 2, and group 4 was the top for $\delta = 5$. Group 1 differed only in one individual with respect to group 2 (individual 194 is replaced by 185). In group 3 individuals 145 and 194 are changed by 185 and 201 with respect to group 1. Group 4 differs from group 1 in three individuals: 172, 189, and 194 which were replaced by 176, 185 and 201.

These four individual groups were compared in terms of average kinship and BV (Table 3). As the weight of kinship (δ) is increased in the selection criteria, the average coancestry decreases from 0.026 to −0.004 by replacing three individuals, whereas the relative gain is reduced by 9 % (from 7.61 to 6.94 %). Based on these empirical results, if

Table 3 Average coancestry based on probability of IBD between genes (θ (*ave*)), merits expressed as average BV_{mod} ($BV_{mod} = BV - \delta\theta$) and gain (expressed as a percentage relative to the phenotypic mean = 7.554 m) of the selected group for height

Group	Tree ID[a]	θ (ave)	Merits[b]							Gain (%)
			δ							
			0	0.25	0.5	0.75	1	2	5	
1	145, 155, 159, 163, 166, 172, 179, 189, 190, 194	0.026	**0.575**	0.568	0.562	0.555	0.548	0.522	0.443	7.61
2	145, 155, 159, 163, 166, 172, 179, 185, 189, 190	0.020	0.574	**0.569**	**0.564**	**0.559**	0.554	0.535	0.476	7.60
3	155, 159, 163, 166, 172, 179, 185, 189, 190, 201	0.007	0.564	0.562	0.561	0.559	0.557	**0.550**	0.529	7.47
4	145, 155, 159, 163, 166, 176, 179, 185, 190, 201	−0.004	0.524	0.525	0.526	0.527	0.528	0.533	**0.546**	6.94

[a] Underlined numbers denote individuals that are not present in the group N° 1

[b] Numbers in bold denote the highest merits for each δ

the selected group is small (10) (a sampling size close to the number of selected trees per origin in the progeny trials established of *P. alba*), replacing one or two individuals is efficient compromise to significantly reduce inbreeding with non-significant (2 %) reduction in gain.

Effect of sample size on average H breeding value and kinship of the selected group

To test how the merit is modified according to sample size of the selected group we used two criteria, $\delta = 0$ (kinship non penalized) and $\delta = 0.5$ (a reasonable penalization according to the results described in the former section). As expected, the average BVs decreases with increasing sample size in both cases (Fig. 4A). Average kinship also decreases with increasing number of individuals for $\delta = 0$, but remains without significant change if $\delta = 0.5$ (Fig. 4B). These results indicate that penalizing merits by average kinship allows selecting low numbers of individuals without significant increase of inbreeding.

Effect of selection by index on the gains of two traits

We evaluated the effect of selecting by two traits at a different contribution level (w) by a mixed index that combines H and DB. The relative gain of each trait selecting the 10 % top trees was estimated for selection indices with relative contributions of H of $w = 0.00$, 0.17, 0.34, 0.5, 0.66, 0.83 and 1.00.

With relative weights of H from 0.00 to 0.80, relative gains are almost constant with higher values for DB and lower values for H (Fig. 5). A steep change occurs for w values between 0.8 and 1. Approximately at $w = 0.83$ the profiles of H and DB intersect, suggesting that this value is the best compromise between the relative gains of H and DB (Fig. 5).

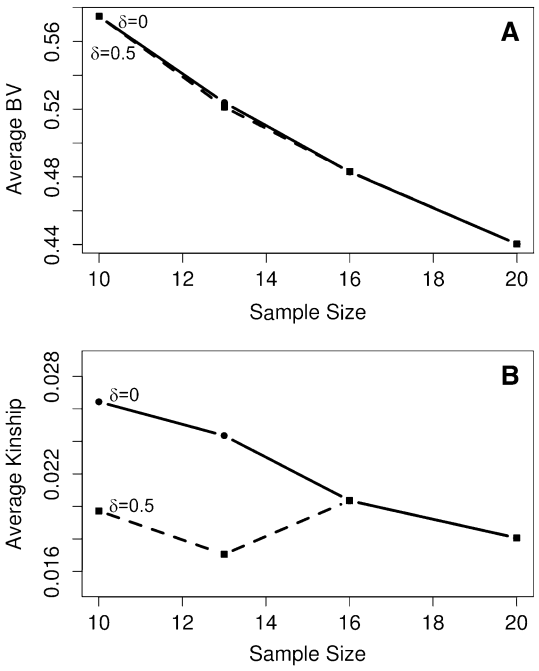

Fig. 4 Effect of sample size on average H breeding value (**A**) and kinship (**B**) of the selected group

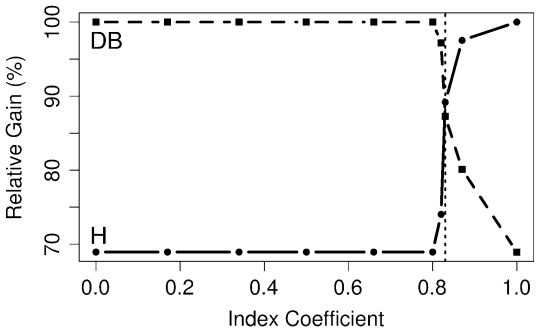

Fig. 5 Effect of selecting by two traits, H and DB, at different contribution levels by a selection mixed index

Hereafter we analyzed the effect of sample size on relative gains for the mixed index ($w = 0.83$). In this case, the same groups were retrieved independently of using $\delta = 0$ or $\delta = 0.5$ for all sample sizes. The average BVs for the mixed index decreases with increasing sample size from 0.88 ($N = 10$) to 0.74 ($N = 20$) (Table 4 column 1).

In terms of relative gain, the response of each trait is quite different. For H relative gain ranges from 6.72 to 5.43 % (with respect to trait average), which implies that gain decrease is about 20 % when comparing the groups of 10 and 20 individuals. DB relative gains in all cases are lower than those observed for H, which is not unexpected taking into account the lower heritability of this trait according to rrBLUP estimates. The effect of increasing N determines a reduction of 13 % (comparing $N = 10$ and 20) in DB gain (Table 4).

Table 4 Average Breeding value (BV) for mixed index, height (H) and basal diameter (DB) considering different sample size

Sample size	Mixed index	*H*	*DB*
10	0.878	0.508 (6.72 %)	2.684 (2.00 %)
13	0.835	0.450 (5.96 %)	2.716 (2.03 %)
16	0.797	0.436 (5.77 %)	2.559 (1.91 %)
20	0.737	0.410 (5.43 %)	2.335 (1.74 %)

In parenthesis BV are expressed as percentage with respect to the phenotypic mean

Discussion

The use of molecular markers allows heritability estimation in wild populations with varying unknown levels of relatedness between sampled individuals by an approach alternative to pedigree based analysis of variance (Frentiu et al. 2008; Wilson et al. 2010; Sillampää 2011). The basis of this approach is applying the animal model, but replacing the pedigree-based by a marker-based relationship matrix (Frentiu et al. 2008). For this reason, an important effort is being made to develop methods of pedigree reconstruction based on molecular markers when genealogical information is incomplete or even absent (Hadfield et al. 2006; Fienieg and Gakusera 2013). Tree breeding selection based on molecular markers is potentially superior to phenotypic selection because of the greater response per unit of time (Wong and Bernardo 2008; Johnson 2004; El-Kassaby et al. 2012).

The progeny trials of *Prosopis* established in Argentina give the opportunity to compare the estimation of genetic parameters based on (partially) known pedigree data with those estimated over predicted relatedness parameters based on molecular data (Bessega et al. 2009, 2010). The highly significant correlations between P-BLUP and G-BLUP predicted breeding values in the progeny trial in San Carlos provide support to the application of G-BLUPs for early selection of genotypes in wild populations where genealogical information is completely lacking.

According to Ritland (1996b), 2–10 loci with ten alleles each would be adequate for estimation of pairwise relatedness. In a recent paper Cros et al. (2014) concluded, from simulation experiments, that the number of loci necessary to get accurate pedigree reconstruction by simulated annealing is about 30. However, in that work the average number of alleles per locus is rather low (from 2.2 to 5.5). As stated by Lynch and Ritland (1999), an increase in the number of alleles per locus should reduce the sampling variance of relatedness estimates because alleles that are identical by state will be more reliable as indicators of identity by descent. In this paper, although we used a number of loci lower than the recommendation made by Cros et al. (2014), the number of alleles per locus is more than twice the average of those used in their work, which is expected to compensate the error attributable to loci number. Furthermore, Wang et al. (2010) recommended a number of loci similar to that used in the present study to state that it may estimate pairwise relatedness coefficients between individuals.

Among the potential risk factors when estimating genetic parameters in the wild are the effects of non-uniformity in tree ages and the high environmental (non-controlled) variation. As stated above, in the wild population studied, it was assumed that most sampled trees are about the same age. In respect to the environmental variation, if genotypes are not randomly distributed throughout a sampled area, genotype-environment covariance (σ_{GE}) might be significant. If $\sigma_{GE} > 0$, the genotype-environment correlation is expected to

inflate the estimates of genetic components of phenotypic variance and hence overestimate heritability. The occurrence of genotype-environment correlation does not seem plausible in our population as the correlation between pairwise spatial distances and Ritland relatedness is non-significant ($r = -0.07$; $p = 0.9995$, based on 2,000 permutations). Another source of bias for the estimation of heritability values under wild conditions is genotype-environment interaction. This variance component cannot be separately measured and may be confused with the environmental variance. If the relative merits of different genotypes change in sign according to varying environments, averaging merits over the different environments would give non significant differences among genotypes. A situation like this would determine an underestimation of genetic variance and heritability in the wild.

If we compare progeny trials and wild populations, we can also expect a lower heritability value in a single wild population in comparison with the plantation based on eight origins (Felker et al. 2001). In the present analysis the estimated heritabilities in the wild population were relatively low ($h^2 = 0.129$ and 0.025 for H and DB respectively). However, the results of the analysis of variance components by the multivariate generalized linear mixed model supports that the genetic component of variance (and hence the heritability) would be significant for both traits.

Sampling strategies of elite trees from wild populations in applied breeding should consider several complex factors (Yanchuk and Sanchez 2011) such as heritabilities and the unequal economic importance of different traits. Different scenarios of selection can be discussed in terms of benefits per unit of time. If selection is based only on phenotypic characteristics, the genetic information is lacking and if the best group of trees includes related individuals, it may result in an increased level of inbreeding. So the breeder must select individuals with minimal coancestry and thus achieve minimal genetic loss. In this work, we showed that the replacement of only a few individuals in the top selected group represents a good strategy to cope with the compromise between gain and coancestry increase.

Based on this conclusion our recommendation for the initiation of progeny trials for *P. alba* conservation and breeding is to collect genetic material from non related mother trees based on the information derived from molecular markers that contribute towards reduction of coancestry and maximization of the average BV estimates. With this prior step an improved genetic basis with low kinship among selected genotypes can be obtained thus increasing gain per time unit.

Another important consideration is the size of the sample to be taken in order to initiate a plantation. The initiation with a great sample size is recommended when studies of inbreeding are not possible. If the economic aspect is not a limitation, a high number of individuals can be included and the effect of coancestry can be compensated, but in the case the number of individuals is a limiting factor, then the coancestry estimation should be considered. This also holds in the case of many populations of *P. alba* that are highly fragmented and individuals within patches may be highly related. According to our results, controlling kinship by penalizing BVs allows reduction of sample size from 20 to 10 individuals, retaining a significant genetic gain without increasing the inbreeding of the group.

The *P. alba* progeny trials in Argentina were founded with 7–20 open-pollinated families per origin. For historical reasons the interaction among the number of selected trees (N), inbreeding and gain of the founder group were not evaluated. The availability of molecular markers to control these parameters would contribute to improve samplings for future progeny trial installation.

Finally, when the best phenotype is defined for more than one trait, correlations between traits should be considered. Negative correlations may imply serious constraints which involve a compromise between gains at an individual trait level (Sanchez et al. 2008). In the case analyzed here H and DB proved to be phenotypically and genetically positively correlated, hence when we select for one trait, we are indirectly selecting for the other. However, unless the correlation between traits is one, the relative weight put in each trait in order to rank individuals by applying a selection index based on BVs may affect the economic success of the breeding program.

Following the historical approach to evaluate the relative contribution of each trait through an index (Hazel 1943), we determined that the optimum condition in the analyzed population implies a higher weight for H (83 %) than for basal diameter (17 %), which is explained by the higher heritability of H. An interesting fact when applying selection index for H and DB is that the top groups do not depend on the inbreeding based penalization (δ) of BVs. This means that selecting for two traits involves the inclusion of fewer related individuals in the sampled group. This effect is expected to be more evident as the genetic correlation between traits is reduced. With the applied criteria for selecting H and DB, a gain near 6 % in height and 2 % in basal diameter might be obtained in the collection of wild material to initiate the progeny trial. If we consider these values in relation to the trunk volume production, nearly 10 % gain is expected. Considering that the average time to harvest *P. alba* is about 22 years, the estimated gain in terms of time is equivalent to saving 2 years from sowing to harvest.

Although the power of this analysis is limited by the low progeny sample size and few SSRs, it is a common problem with low-input breeding species or with high conservation value species. The strategy of collecting elite trees from wild populations based on breeding values predicted by G-BLUP, together with weighting possible effects of inbreeding and sampling sizes may improve the material at the stand foundation phase with some genetic gain at the first generation. The recommendations made here for the first steps of the breeding would help promote a more efficient and cost-saving way to rationalize management of *Prosopis* and other endangered native forest species.

Acknowledgments This research was supported by funding from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) PIP 11220090100147, Universidad de Buenos Aires (20020100100008) and Agencia Nacional de Promociones Científicas y Tecnológicas (ANPCyT) PICT-2013-0478 given to JCV and BOS. We extend our sincere appreciation to all the anonymous reviewers.

References

Adamoli JR, Ginzburg, Torrilla S (2011) Escenarios productivos y ambientales del Chaco Argentino: 1977–2010. http://agrolinux.agrositio.com/producir_conservando/documentos/escenarios_productivos_ambientales_chaco.pdf

Anderson WE, Sanchez L, Andersson B (1999) Group coancestry controlled selection in a Pinus sylvestris L breeding program. Theor Appl Genet 99:73–80

Assoumane AA, Vaillant A, Mayaki Z, Verhaegen D (2009) Isolation and characterization of microsatellite markers for Acacia Senegal (L.) Willd., a multipurpose arid and semi-arid tree. Mol Ecol Resour 9(5):1380–1383

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

Bessega C, Saidman BO, Darquier MR, Ewens M, Sánchez L, Rozenberg P, Vilardi JC (2009) Consistency between marker and genealogy-based heritability estimates in an experimental stand of *Prosopis alba* (Leguminosae). Am J Bot 96:458–465

Bessega C, Saidman BO, Darquier MR, Ewens M, Felker P, Vilardi JC (2010) Accuracy of dominant markers for estimation of relatedness and heritability in an experimental stand of *Prosopis alba* (Leguminosae). Tree Genet Genomes 7:103–115

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 Genet 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). Appl Plant Sci 1(5): 1200324. doi:10.3732/apps.1200324; http://www.bioone.org/loi/apps

Boys A, Cherry M, Dayansdan S (2005) Microsatellite analysis reveals genetically distinct populations of red pine (*Pinus resinosa*, Pinaceae). Am J Bot 92(5):833–841

Brisbane JR, Gibson JP (1995) Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. Theor Appl Genet 91:421–431

Burdon RD, Shelbourne CJA (1971) Breeding populations for recurrent selection: conflicts and possible solutions. N Z J For Sci 1:174–193

Chessel D, Dufour AB, Thioulouse J (2004) The ade4 package-I-One-table methods. R News 4:5–10

Clark SA, Hickey JM, Daetwyler HD, van der Werf JHJ (2012) The importance of information on relatives for the genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schenes. Genet Sel Evol 44:4. doi:10.1186/1297-9686-44-4

Cros D, Sánchez L, Cochard B, Samper P, Denis M, Bouvet JM, Fernández J (2014) Estimation of genealogical coancestry in plant species using a pedigree reconstruction algorithm and application to an oil palm breeding population. Theor Appl Genet. doi:10.1007/s00122-014-2273-3

Degen B, Streiff R, Ziegenhagen B (1999) Comparative study of genetic variation and differentiation of two pedunculate oak (*Quercus robur*) stands using microsatellite and allozyme loci. Heredity 83:597–603. doi:10.1038/sj.hdy.6886220

Delvalle P, Antanasio M, Ayala M, Svirz I, Petkoff J (2003) Ensayos de orígenes de *Prosopis alba* grises (algarrobo blanco). Revista Instituto Nacional de Tecnología Agropecuaria. Estación Experimental Colonia Benítez "Augusto G. Schulz" : 1–8

Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. J Stat Softw 22(4):1–20

Dray S, Dufour AB, Chessel D (2007) The ade4 package-II: two-table and K-table methods. R News 7(2):47–52

El-Kassaby YA, Lstiburek M (2009) Breeding without breeding. Genet Res 91:111–120

El-Kassaby YA, Klapste J, Guy RD (2012) Breeding without breeding: selection using the genomic best linear unbiased predictor method (G-BLUP). New For 43:631–637

Endelman JB (2011) Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genomes 4:250–255. doi:10.3835/plantgenome2011.08.0024

Endelman JB, Jannink JL (2012) Shrinkage estimation of the realized relationship matrix. G3: genes. Genomes, Genet 2:1405–1413. doi:10.1534/g3.112.004259

Ewens M, Gezan S, Felker R (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

FAO (2007) Situación de los Bosques del Mundo, Roma, Italia. http://www.fao.org/docrep/009/a0773s/a0773s00.htm

Felker P, Guevara JC (2003) Potential of commercial hardwood forestry plantations in arid lands—an economic analyses of *Prosopis* lumber production in Argentina and the United States. For Ecol Manag 186:271–286

Felker P, Smith D, Wiesman C, Bingham RL (1989) Biomass production of *Prosopis alba* clones at two non-irrigated field sites in semiarid south Texas. For Ecol Manag 29:135–150

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. Agrofor Syst 53:65–76

Felker P, Ewens M, Velarde M, Medina D (2008) Initial evaluation of *Prosopis alba* griseb clones selected for growth at seawater salinities. Arid Land Res Manag 22:334–345

Fernandez J, Toro MA (2006) A new method to estimate relatedness from molecular markers. Mol Ecol 15:1657–1667

Ferreyra LI, Bessega C, Vilardi JC, Saidman BO (2007) Consistency of population genetics parameters estimated from isozyme and RAPDs dataset in species of genus Prosopis (Leguminosae, Mimosoideae). Genetica 131:217–230

Fienieg ES, Gakusera P (2013) The use and integration of molecular DNA information in conservation breeding programmes: a review. J Zoo Aquar Res 1(2):44–51

Frentiu FD, Clegg SM, Chittock J, Burke T, Blows MW, Owens IPF (2008) Pedigree-free animal models: the relatedness matrix reloaded. Proc R Soc Lond B Biol Sci 275:639–647
Gill JL (1965) Effects of finite size on selection advance in simulated populations. Aust J Biol Sci 18:599–617
Hadfield JD (2010) MCMC methods for multi-response generalised linear mixed models: the MCMCglmm R Package. J Stat Softw 33(2):1–22
Hadfield D, Richardson DS, Burke T (2006) Towards unbiased parentage combining genetic, behavioral and spatial data in a Bayesian framework. Mol Ecol 15:3715–3730
Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes 2:618–620
Hazel LN (1943) The genetic basis for constructing selection indexes. Genetics 28:476–490
Johnson R (2004) Marker-asisted selection. Plant Breed Rev 24(1):293–309
Kruuk LEB (2004) Estimating genetic parameters in natural populations using the "animal model". Philos Trans R Soc Lond B 359:873–890
Lande R (1979) Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution 33:402–416
Lande R (1980) Genetic variation and phenotypic evolution during allopatric speciation. Am Nat 116:463–479
Lindgren D, Mullin TJ (1997) Balancing gain and relatedness in selection. Silvae genet 46:124–129
Lopez C, Maldonado A, Salim V (2001) Variación genética de Progenies de *Prosopis alba*. Investigación Agraria. Sistemas y Recursos Forestales 10(1):59–68
Lstiburek M, Klapste J, Kobliha J, El-Kassaby YA (2012) Breeding without breeding. effect of gene flow on fingerprinting effort. Tree Genet Genomes 8:873–877
Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. Genetics 152:1753–1766
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. Mol Ecol Notes 5:487–489
Namkoong G, Kang HC, Brouars JS (1988) Tree breeding: principles and strategies. Monographs on theoretical and applied genetics. Spinger, New York
Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590
Nielsen R, Tarpy DR, Kern Reeve H (2003) Estimating effective paternity number in social insects and the effective number of alleles in a population. Mol Ecol 12:3157–3164
Olsson T, Lindgren D, Li B (2001) Balancing genetic gain and relatedness in seed orchads. Silvae Genet 50:222–227
Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–290
Pometti CL, Pizzo B, Brunetti M, Macchioni N, Ewens M, Saidman BO (2009) Argentinean native wood species: physical and mechanical characterization of some *Prosopis* species and Acacia aroma (Leguminosae; Mimosoideae). Bioresour Technol 100(6):1999–2004
R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0. http://www.r-project.org/
Ritland K (1996a) Estimators of pairwise relatedness and individual inbreeding coefficients. Genet Res 67:175–185
Ritland K (1996b) A marker-based method for inferences about quantitative inheritance in natural populations. Evolution 50(3):1062–1073
Roig FA (1993) Aportes a la Etnobotánica del Género *Prosopis*. In: Contribuciones Mendocinas a la quinta Reunión Regional para América Latina y el Caribe de la Red de Forestación del CIID, Unidades de Botánica y Fisiología vegetal, Iadiza (eds.), p 99–121
Rojas S, Ch YKC, Nagao EO (2011) Diversidade genética em acessos do banco de germoplasma de camucamu (Myrciaria dubia [H.B.K.] McVaugh) do INPA usando marcadores microssatélites (EST-SSR). Revista Corpoica Ciencia y Tecnología Agropecuaria 12(1):51–64
Salto CS (2011) *Variación genética en progenies de polinización abierta de Prosopis alba Griseb. de la Región Chaqueña.* Maestría en genética vegetal Thesis. Área de mejoramiento genético UNR – Concordia
Sanchez L, Yanchuk AA, King JN (2008) Gametic models for multitrait selection schemes to study variance of response and drift under adverse genetic correlations. Tree Genet Genomes 4:201–212
Sillampää MJ (2011) On statistical methods for estimating heritability in wild populations. Mol Ecol 20:1324–1332

Stoehr M, Yanchuk AA, Xie Ch, Sanchez L (2008) Gain and diversity in advances generation coastal Douglas fir selections for seed production populations. Tree Genet Genomes 4:193–200. doi:10.1007/s11295-007-0100-6

Verga A (2005) Recursos genéticos, mejoramiento y conservación de especies del género *Prosopis*. Mejores árboles para más forestadores: El Programa de Producción de Material de Propagación Mejorado y el Mejoramiento Genético en el Proyecto Forestal de Desarrollo Edición: Carlos A. Norberto. SAGPyA-BIRF

Verga A (2009) Domesticación de especies forestales nativas del Parque Chaqueño (Algarrobo) http://inta.gob.ar/proyectos/pnfor-044341

Wang J (2011) COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. Mol Ecol Resour 11(1):141–145

Wang J, Brekke P, Huchard E, Knapp LA, Cowlishaw G (2010) Estimation of parameters of inbreeding and genetic drift in populations with overlapping generations. Evol Int J Organic Evol 64:1704–1718

Weir RP, Lindgren D (1996) Effecive family number following selection with restrictions. Biometrics 52:525–535

White TL, Neale DB (2007) Data analysis-mixed models, variance components and breeding values. In: White TL, Adams WT, Neale DB (eds) Forest genetics, p 395–437. ISBN 978-1-84593-285-5. doi: 10.1079/9781845932855.0395

Wilson AJ, Reale D, Clements MN, Morrissey MM, Postma E, Walling CA, Kruuk LEB, Nussey DH (2010) An ecologist's guide to the animal model. J Anim Ecol 79:13–26

Wong CK, Bernardo R (2008) Genome wide selection in oil palm: increasing selection gain per unit time and cost with small populations. Theor Appl Genet 116:815–824

Yanchuk AD, Sanchez L (2011) Multivariate selection under adverse genetic correlations: impacts of population sizes and selection strategies on gains and coancestry in forest tree breeding. Tree Genet Genomes 7:1169–1183