# **Expected genetic gain and genotype by environment interaction in almond** (*Dipteryx panamensis* (**Pittier**) **Rec. and Mell**) **in Costa Rica**

León N<sup>1</sup>, Murillo O, Badilla Y<sup>1</sup>, Ávila C<sup>2</sup>, Murillo, R<sup>2</sup>

<sup>1</sup> Instituto Tecnológico de Costa Rica. School of Forest Engineering. Cartago, Costa Rica <sup>2</sup> Universidad Nacional, Research and Forest Services Institute. Heredia, Costa Rica

Corresponding author: O. Murillo, e-mail: olmuga@yahoo.es

# Abstract

Almendro (Dipteryx panamensis (Pittier) Record & Mell.) is one of the native tree species with higher wood density planted in Latin American tropical region, which has given it a reputation and international acceptance in the wood-floors markets. The tree has also a remarkable ecological value, as an important food source and habitat for several endangered fauna. This research studied the potential for genetic improvement and stability of a collection of 29 open-pollinated families of Dipteryx panamensis at 48 months of age. The materials came from three different local native populations northern from Costa Rica. Individual heritabilities for commercial volume registered moderate values, between 9 % and 22 %, while mean family heritability, ranged between 47 % and 74 % for the traits evaluated. No significant genetic differences among provenances for none of the traits evaluated were observed. The genetic correlation for commercial volume between both sites showed a value of r = 0.96, which suggests low G x E interaction. If selected the top 20 individuals from the genetic ranking in commercial volume, an expected genetic gain above 50 % in year 4 can be obtained. The use of improved almendro seeds could have great impact and motivate plantations with this tree species in tropical regions. It is concluded that a unique breeding program for the whole country is feasible.

**Keywords:** almond, Dipteryx panamensis, breeding, genotype x environment interaction

## Introduction

Reforestation with native tree species is becoming more importantly in recent years in Costa Rica (Herrera, 2000; Murillo and Guevara 2013). Some experiences demonstrate a good potential in degraded areas (Guariguata, Rheingans and Montagnini, 1995; Butterfield and Espinoza 1995; Badilla, Murillo and Obando, 2002; Redondo, 2007). Almond (Dipteryx panamensis) is a native tree species of the Fabaceae family, subfamily Papilonaceae, naturally distributed from the lowland forests of Nicaragua to Columbia (Romo, 2005). In Costa Rica, almond is abundant from the lowlands of northern Region, to all Caribbean region (Castañeda et al. 2000). The most recent investigations on growth and potential for plantation of this species in Costa Rica were reported by Delgado et al. (2003); Redondo (2007) and Schmidt (2009). Additionally, after evaluating seven native species from humid regions in Costa Rica, Redondo (2007) reported D. panamensis as the species with highest carbon sequestration.

The first genetic studies on this species in the country reported at 3 years age, an individual heritability ( $(h^{2}i)$  of 0.46 for DBH and 0.29 for total height, with a potential expected genetic gain in DBH of 15 % (Martínez *et al.* 2016).

Tree breeding strategies can be designed as a complement of native trees incorporation efforts into forest production (Resende 1999). Reforestation programs must be based on reliable seed sources, with the highest possible genetic quality (Murillo, Espitia, Castillo, 2012). A network of trials with a broad genetic base should be established as a priority, in order to support subsequent selection and elimination processes in the breeding population. These trials should evaluate all possible potential environments, in order to verify the stability of the materials (Redondo 2007; Pires *et al.* 2011). One of the fundamental issues is whether we should use the same genetic materials across the whole country. This decision is particularly important where there are remarkable edaphic and climatic diversity of conditions throughout the potential reforestation regions. Costa Rica is a land with an extraordinary diversity of ecosystems, which may promote important genotype by environment interactions (GxE), that may reduce genetic gains in breeding programs at national scale (Vencovsky and Barriga 1992; Resende 1999). The existence of this interaction is of major importance and need to be determined (Murillo 2001). However maximization of genetic gain must be balanced with an analysis of relatedness increase within breeding populations. Several approaches have been proposed to monitor loss of genetic diversity (Vencovsky and Barriga 1992; Lindgren, Gea, and Jefferson 1996; Olsson, Lindgren, Li 2001), which demonstrate the importance of considering this subject in order to optimize deployment of genetic materials.

The aim of the present investigation was to optimize the selection of a collection of *Dipteryx panamensis* families, evaluated in provenance/progeny tests at 48 and 46 months of age, established in two contrasting regions of Costa Rica.

## **Material and Methods**

In June 2010, a provenance-progeny trial with 29 open-pollinated families of *Dipteryx panamensis* was established at the Instituto Tecnológico de Costa Rica (ITCR) in Santa Clara, San Carlos, northern Region of Costa Rica (10°22'15″,-84°30'59″, 100 m.elevation) Materials were collected from three natural populations in the northern Region of Costa Rica: Puerto Viejo (Sarapiquí), Crucitas (Cutris, San Carlos) and Coopesanjuan (San Marcos, San Carlos). The ecological Life Zone was classified as Premontane Wet Forest transition to Basal (Holdridge 1967). The temperature fluctuates between 18-24°C. The annual average rainfall ranges between 3500 and 4000 mm, with February and March registering a slight hydric deficit which is noticeable in dry years. The soils are Ultisols, deep, well drained, red, with high organic matter content, very low fertility and a pH around 5,5 (Ortiz and Cordero, 2008).

In South pacific region the trial was established in October 2010 in Puerto Escondido, La Palma of Puerto Jiménez (8°38'N and 83°27'W) in the Osa Peninsula, classified as Premontane Wet Forest transition to Basal, according to Holdridge's life zones classification system (Bolaños, Watson and Tosi, 2005). The annual average rainfall is 4500 mm in average, with total distribution throughout the year and warm temperatures ranging between 26 and 32°C. The soils exhibit high basis saturation and flat topography with 0 to 2 % slopes at 20 m elevation. Soils are classified as Inceptisols (Soil Survey Staff, 2007). No site preparation was applied to any of the two trials

Measurements and observations of seedling adaptive traits

### Genetic tests design

Mother trees were selected in natural populations, spatially separated by no less than 500 m. Seeds from field collections were germinated in greenhouse facilities, separated by family. The genetic test consisted in a provenance/progeny trial established as a randomized complete block design, with six blocks and 6 half-sibs plants per parcel (family within block), following the genetic designed utilized by the Tree Improvement Cooperative GENFORES (Badilla y Murillo, 1998; Murillo, Obando, Badilla and Araya, 2001). Each family was represented by a total of 36 individuals (half sibs) per trial. Trees were initially planted at 3 m x 4 m spacing and, both tests were bordered by two rows of trees.

All single trees in trials were evaluated by measuring its diameter at 1.3 m (DBH in *cm*), commercial height (measured in meters by a telescopic pole). Commercial volume was determined as usual: VolCom = (DBH/100)<sup>2\*</sup> $\pi$ /4\*h<sub>com</sub>\*0.65.

Database was organized in an EXCEL worksheet and analysed trait by trait, through the usage of SELEGEN software (EMBRAPA, Brazil), version 2008 (Resende, 2007). SELEGEN program utilized Restricted Maximum Likelihood (REML) and Best Linear Unbiased Prediction (BLUP) procedures for data analysis and genetic parameter estimations. The following models were utilized:

Model for progenies:	y = Xr + Za + Wp + Ti + e	(1)
Model for provenances:	y = Xr + Za + Wp + Qs + Ti + e	(2)

Where "y" is the data vector; "r" is the vector of the effects of repetition added to the general mean; "a" is the vector of individual additive genetic effects; "p" is the vector of the plot effects (all individuals of a family in one block); "s" is the vector of provenance effects; "i" is the vector of GxE interaction effects; "e" is the vector of residual errors. All effects were considered as random, with the exception of repetition effect as fixed. The capital letters represent the matrices of incidence for the referred effects (Resende 2007).

Effective population size was determined by the function suggested by Vencovsky and Barriga (in Resende, 2002, page 70):

$$Ne = \frac{4*N_f * k_f}{K_f + 3}$$
(3)

Where Nf is the number of families selected, while kf is the average number of "k" individuals selected within the "f" families. Parameters of the function increased with each new individual included in the selection, in particular if the tree belongs to a new family.

The genetic gain was estimated by means of the usual function (Zobel and Talbert, 1984):

$$GG = S^*h^2Fam \tag{4}$$

Where S = differential selection

Fam h<sup>2</sup> = family heritability 
$$\frac{0.25Va}{0.25Va + \frac{Vparc}{6} + \frac{0.75Va + Ve}{6*6}}$$

Where Va is the additive variance component, V Block x Fam is the variation among families within blocks.

Genetic gain estimation considering within families selection was determined as follows:

GG = S \* Famh2 + Sw \* h2w (within family) 
$$\frac{0.75Va}{0.75Va+Ve}$$
 (5)

Where  $S_{w}$  = selection differential within families

Selection scenarios were performed in order to balance genetic gain with loss of diversity, measured by remaining effective population size (Ne).

## Results

#### Provenance analysis

In Table 1 and 2 is presented the analysis by provenances level with respect to commercial volume, without distinction of family structure. As may be observed, the three provenances Coopesanjuan, Sarapiquí and Crucitas, exhibited quite similar genetic values among them. Similarly, in table 2 can be observed, that provenance level registered low values in almost all parameters in comparison with family level. In Table 2 is shown genetic parameters of the 3 provenances at both sites. It can be noted that the provenance mean heritability (Prov h<sup>2</sup>) for commercial volume is 65 %. In terms of estimates accuracy, commercial volume was the variable with highest values.

#### Table 1

Genetic ranking of native provenances of *Dipteryx panamensis* based on individual commercial volume at 48 months of age, northern Region and South Pacific of Costa Rica.

		Genetic Value (u + g)
Order	Provenance	(VolCom/tree in m <sup>3</sup> )
1	Coopesanjuan (San Marcos, San Carlos)	0.0260
2	Puerto Viejo, Sarapiquí	0.0245
3	Crucitas (Cutris, San Carlos)	0.0243

Due to its negligible genetic variation among studied provenances, it was decided to pool all single families as one unique metapopulation for practical decisions.

In Table 2 is shown estimated values for all genetic parameters for traits evaluated in the joint analysis of the two locations. Evaluated traits registered values of individual heritability (h<sup>2</sup>i) ranging between 9 % and 22 %. Meanwhile, family heritability (Fam h<sup>2</sup>) registered higher values, ranging between 47 % and 74 %. Within family heritabilities (h<sup>2</sup>ad) showed very similar values to individual heritability (h<sup>2</sup>i).

Accuracy determines the quality or confidence of genetic parameters estimation. As can be seen, accuracy was very high for commercial volume and diameter, both of which showed values greater than 86 %, whereas for commercial height it was only 68 %.

The "r" value stands for the genetic correlation between sites, which showed values of 96 % for commercial volume and  $r_g = 0.97$  for diameter, but relatively low value for commercial height.

#### Table 2

Genetic parameters of twenty-nine open-pollinated families
of Dipteryx panamensis at 48 months of age evaluated in nort-
hern region and South Pacific of Costa Rica.

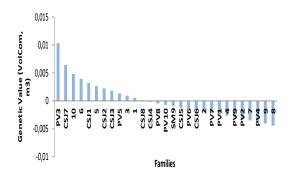
Parameter		Families		Provenances
	DBH (cm)	hcom (m)	VolCom (m3)	VolCom (m3)
Va	0.84	0.34	0.0001	0.0001
V blockxFam	0.22	0.28	0.0001	0.0001
V <sub>GxE</sub>	0.01	0.04	0.0001	0.0001
Ve	3.32	3.07	0.0001	0.0002
VP	4.38	3.73	0.0002	0.0002
Individual h <sup>2</sup> i	0.19 + -0.08	0.09 + -0.05	0.22 + -0.08	
$h^2g$				0.0066 +-0.007
Fam h <sup>2</sup>	0.73	0.47	0.74	
Prov h <sup>2</sup>				0.65
Accuracy	0.86	0.68	0.86	0.81
Within Fam h <sup>2</sup> wf	0.16	0.08	0.19	
r (correlation)	0.97	0.65	0.96	0.90
SEP	0.24	0.21	0.002	0.001
Mean	8.99	5.40	0.02	0.02

DBH= diameter at 1.3 m; hcom= tree commercial height; VolCom=commercial volume; Va= additive genetic variance; V blockxFam= environmental variance among families within blocks); V<sub>GKE</sub> = GxE interaction variance component; Ve= residual variance not explained by the model; VP= Va (Vg) + Vparc + Ve: total phenotypic variance. h<sup>2</sup>i= individual heritability in the strict sense, i.e., of the additive effects; Fam h<sup>2</sup>= mean family heritability, assuming full survival; Within Fam h<sup>2</sup>wf = 0,75Va; Accuracy:  $\sqrt{h^2 family}$ 

 $\frac{0,75Va + Ve}{0,75Va + Ve}$ , Accurac

accuracy of families selection assuming full survival; r = genetic (g) correlation between both locations; SEP= standard deviation of the predicted genotypic value of progeny, assuming full survival.

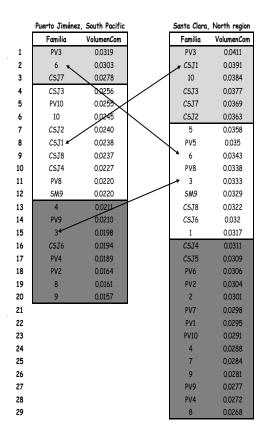
Figure 1 shows the genetic ranking for commercial volume of the 29 open-pollinated families evaluated at both sites. Families PV3, CSJ17, 10, 6 and CSJ1 are located in the top five positions; 12 families position above the mean value of the population and, 17 families remain below the mean value. Families 8, 9, PV4, 7 and PV2 are located in the worst positions at both sites.





Genetic ranking distribution in commercial volume of 29 open-pollinated families of *Dipteryx panamensis* at 48 months of age, evaluated in two environments (northern Region and South Pacific of Costa Rica).

Figure 2 shows the behaviour of both genetic ranking of commercial volume in the two locations. It can be observed, that some of the 29 families registered slight position movements



#### Figure 2

Commercial volume genetic ranking of 29 open-pollinated families of *Dipteryx panamensis* at 48 months age, evaluated in the northern region and south Pacific of Costa Rica.

in both environments, suggesting a low GxE interaction of simple behavior.

## Discussion

#### Analysis of provenances

Provenance heritability values (h<sup>2</sup>g) registered no significant differences among the three populations investigated (Table 1), which suggests that the greatest genetic richness is at the family level within provenances. Hence, breeding efforts should focus on increasing the number of families into the program.

Geographical distances between the three provenances is about 50-75 km, and all of them are located today into small natural forest fragments. Less than 100 years ago this region was covered by a continuous natural forest, suggesting the existence of an almond metapopulation in this region of the country (Gamboa 2008). This condition may explain the small genetic differences among provenances. Future collections need to sample more distant populations and expand the number of families within populations.

The genetic correlation between the two location tests was high for commercial volume (r = 90 %), which supports the hypothesis of maintaining a single national breeding program.

#### Analysis of families

Genetic parameters of the 29 open-pollinated families showed commercial volume with highest value (h<sup>2</sup>i of 0.22) at this age among the traits evaluated, which is very promising for breeding purposes. Therefore, it is a suitable trait for genetic selections.

Given that the accuracy values were very high, genetic selection of best families could be performed in a safe and objective way. However, within families heritability (h<sup>2</sup>ad) reported relatively low values in all traits (< 0.2). Which suggests a reduced intra family variation.

Such a high correlation between both locations (r > 0.96) supports the decision of developing a single national breeding program with almond in the country. Figure 1 shows the genetic ranking of single families in relation to commercial volume, where family PV3, provenance Sarapiqui, was the one to occupy the first position, followed by families CSJ7, 10, 6 and CSJ1, conforming the group of the top 5 families. At the other end of the ranking, the five worst positions are occupied by family 8, native of Crucitas (San Carlos), followed by families 9, PV4, 7 and PV2.

Figure 2 complements the information showing very low GxE interactions between the two tests. Some few families change their position, but they express it in a very slightly manner, that can be described as a GxE interaction of simple type. Almost none of the families in the upper third of the ranking, change its position. Similarly occurs in the lower third of the genetic ranking, where the same families occupy those positions in both locations.

### Expected genetic gain

Given that genetic correlation between both sites was extremely high (r = 0.9 for provenances and r = 0.96 for families), for practical reasons can be decided, whether to base on the results from the most complete and reliable of the two genetic tests. South Pacific location contains only 20 of the 29 families. Additionally, inadequate weed control and strong wind events in this region, could have influenced the results of this test. Accuracy at commercial volume was of 0.89 in San Carlos (northern Region), 0.71 in Osa Peninsula (South Pacific), and 0.86 both sites combined. Therefore, combining both sites will reduce the quality of the estimates and lower heritability values.

Based on the values of the trial in San Carlos a possible selection decision would be collecting seeds from the 20 best individuals in the single trees genetic ranking for commercial volume, regardless of the family or provenance they belong. In this case, a genetic gain of 0.0137 m<sup>3</sup>/tree would be obtained, which represents 51.1 %, which is extremely high. However, the genetic diversity of the resulting commercial subpopulation would be N<sub>e</sub> = 13.12, due to these 20 top individuals belong to only 9 single families. If it is decided to use the family structure and selecting the top 4 individuals from the top 5 families of the ranking, the following will be the results:

## GG TOP 5 families = S\*famh<sup>2</sup>

GG TOP 5 families = (0.0344-0.0268)\*0.799 = 0.00607 m<sup>3</sup>/tree = 22.66 %

 $\begin{array}{l} GG_{Within\,Fam}=S_{within\,fam}*h^2_{within\,fam}\\ GG_{Within\,Fam}=\!0.0231*\!0.30=0.007\ m^3/tree=26\ \%\\ GG=22.66\ (family\ selection)+26.0\ (selection\ within\ top\ families)=48.66\ \%. \end{array}$ 

This value of genetic gain is practically the same as the earlier selection decision. However, the expected genetic diversity need to be estimated and considered

$$N_e = \frac{4N_f \bar{k}_f}{\bar{k}_f + 3 + (\sigma_{k_f}^2 / \bar{k}_f)} \quad N_e = \frac{4*5*5}{5+3+0} = 12.5$$

Can be now observed that genetic diversity (Ne = 12,5 < 13,12) registers a slightly lower value. Therefore, can be decided to simply select the top 20 individuals of the ranking, regardless its family or provenance origin.

## CONCLUSIONS

Future breeding efforts should continue to be performed based on the population established in San Carlos (northern Region), due to its higher reliability and genetic richness.

No differences were observed among the three native provenances investigated. Therefore, all single families can be pooled together in order to conform one single breeding population.

Based on the combined site analysis, the top 5 families relative to commercial volume were PV3, CSJ7, 10, 6 and CSJ1. The strong genetic correlation between both locations (r > 0.96) supports the decision of one single breeding program for almond in the country.

The results confirm that the greatest genetic variability is located between individuals and between families within each provenance, while very little genetic variability exists between the three native populations investigated.

The best deployment decision for reforestation will be collecting seeds from the top 20 single trees from the ranking of commercial volume in Santa Clara, San Carlos, regardless of the family o provenance they belong. It would promote a genetic gain of 0.0137 m3/tree (>50%) at 4 years old.

## Acknowledgement

This research was part of a Research Grant from the Research and Extension Council at the Instituto Tecnológico de Costa Rica. Support was also provided by GENFORES (Tree Improvement and Gene Conservation Cooperative).

## References

- Badilla Y, Murillo O (1998) Propuesta de un diseño de parcela para la investigación con especies nativas. Kurú, 25, 4.
- Badilla Y, Murillo O, Obando G (2002) Reforestación con especies nativas en la zona norte del país. In: Seminario Nacional sobre Especies Nativas. 3-5 de abril, 2002. Universidad Nacional, INISEFOR. Heredia, Costa Rica.
- Bolaños A, Watson V, Tosi J (2005) Mapa ecológico de Costa Rica (Zonas de Vida). San José, Costa Rica, Centro Científico Tropical 206. Esc. 1:200.000. Color.

- Butterfield R, Espinoza M (1995) Screening trial of 14 tropical hardwoods with an emphasis on species native to Costa Rica: fourth year results. New Forests, 9: 135-145. <u>https://doi.org/10.1007/bf00028686</u>
- Castañeda H, Moreira I, Arnáez E, Sánchez E (2000) Estudio morfológico de diferentes estadios ontogénicos de flor, fruto y semilla de Dipterix panamensis (Pittier) Record & Mell (Fabaceae) (Almendro). Tecnológia en marcha, 14(1), 124-132.
- Delgado A, Montero M, Murillo O, Castillo M (2003) Crecimiento de especies forestales nativas en la zona norte de Costa Rica. Agronomía Costarricense, 27(1), 63-78.
- Gamboa N (2008) Regeneración de Dipteryx panamensis en bosques bajo manejo forestal en el paisaje fragmentado del Noreste de Costa Rica. (Tesis M.Sc.) Universidad Estatal a Distancia (UNED). San José, Costa Rica
- Guariguata M, Rheingans R, Montagnini F (1995) Early woody invasion under tree plantations in Costa Rica: implications for forest restoration. Restoration Ecology, 3(4), 252-260.

#### https://doi.org/10.1111/j.1526-100x.1995.tb00092.x

- Herrera M (2000) Evaluación del modelo de desarrollo de plantaciones forestales en la región Huetar Norte. (Tesis de grado). Universidad Nacional. Heredia, Costa Rica.
- Holdridge LR (1967) Life Zone Ecology. San José, Costa Rica, Tropical Science Center. p. 40-43.
- Lindgren D, Gea L, Jefferson P (1996) Loss of Genetic Diversity Monitored by Status Number. Silvae Genetica 45 (1): 52-59.
- Martínez V, Fallas L, Murillo O, Badilla Y (2016) Potencial de mejoramiento genético en Dipteryx panamensis a los 33 meses de edad en San Carlos, Costa Rica. Rev. For. Mesoamericana Kurú 13(30): 3-12. <u>https://doi.org/10.18845/rfmk.v13i30.2455</u>
- Murillo O, Obando G, Badilla Y, Araya E (2001) Estrategia de mejoramiento genético para el Programa de Conservación y Mejoramiento Genético de Especies Forestales del ITCR/FUNDECOR, Costa Rica. Revista Forestal Latinoamericana, 16(30), 273-285.
- Murillo O (2001) Genotype by environment interaction and genetic gain analysis on unbalanced data of Pinus oocarpa provenances. Agronomía Costarricense. 25(1): 21-32.
- Murillo O, Espitia M, Castillo C (2012) Fuentes Semilleras para la Producción Forestal. (1ª ed.). Bogotá, Colombia: Domar S.A.S.
- Murillo O, Guevara V (2013) Estado de los recursos genéticos forestales de Costa Rica. MINAET/FAO/CONAGEBIO, San José, Costa Rica. 159 pp.
- Olsson T, Lindgren D, Li B (2001) Balancing Genetic Gain and Relatedness in Seed Orchards. Silvae Genetica 50 (5-6): 222-227.
- Ortiz E, Cordero S (2008) Atlas Digital de Costa Rica. [CD-ROM]. Instituto Tecnológico de Costa Rica, Cartago, Costa Rica.
- Pires IE, Resende MDV, Da Silva R., Resende Jr MF (2011) Genética Florestal. Vicosa. Ed. Arka. Minas Gerais, Brazil.
- Redondo A (2007) Growth, carbon sequestration, and management of native tree plantations in humid regions of Costa Rica. New Forests, 34: 253–268. https://doi.org/10.1007/s11056-007-9052-9
- Resende MDV (1999) Melhoramiento de essencias florestais. In: BORÉM A (Ed). Melhoramiento de Especies Cultivadas. Editora UFV. Vicosa.
- Resende MDV (2002) Genética Biométrica e Estatística no Melhoramento de Plantas Perenes. EMBRAPA. Brasilia, Brasil. 975p.
- Resende MDV (2007) SELEGEN-REML/BLUP Sistema Estadistico e Seleçao Genetica Computadorizada (Software). Brasilia, Brasil: EMBRAPA
- Romo M (2005) Efecto de la luz en el crecimiento de las plántulas de Dipteryx micrantha Harms «Shihuahuaco» transplantadas a Sotobosque, Claros y Plantaciones. Revista Ecologia Aplicada, 4(1-2), 1-8. https://doi.org/10.21704/rea.v4i1-2.291
- Schmidt F (2009) The effect of the site selection on the growth of Dipteryx panamensis in timber plantations in Costa Rica and Panama. (Tesis M.Sc.) Institute of Forest Growth and Forest Computer Sciences. Faculty of Forest, Geo and Hydro Sciences. University of Technology, Dresden, Germany.
- Soil Survey Staff (2007) Keys to soil taxonomy. Washington, DC, US, USDA-NRCS. p. 435-512.
- Vencovsky R, Barriga P (1992) Genetica biometrica no fitomelhoramento. Ribeirao Preto: Sociedad Brasileira de Genetica.
- Zobel B, Talbert J (1984) Applied Forest Tree Improvement. New York: Jhon Wiley e Sons. 505p.