Research Article

Microsatellite-assisted backcross selection in maize

Luciana Lasry Benchimol[1], Cláudio Lopes de Souza Jr.[2] and Anete Pereira de Souza[3,4]

[1]*Instituto Agronômico de Campinas, Centro de Pesquisa e Desenvolvimento de Recursos Genéticos Vegetais, Campinas, SP, Brazil.*
[2]*Escola Superior de Agricultura "Luiz de Queiroz" Departamento de Genética, Piracicaba, SP, Brazil.*
[3]*Universidade Estadual de Campinas, Centro de Biologia Molecular e Engenharia Genética, Campinas, SP, Brazil.*
[4]*Universidade Estadual de Campinas, Instituto de Biologia, Departamento de Genética e Evolução, Campinas, SP, Brazil.*

Abstract

A microsatellite marker (SSR) was chosen to simulate a target allele and three criteria (02, 04 and 06 markers per chromosome) were tested to evaluate the most efficient parameters for performing marker-assisted backcross (MAB) selection. We used 53 polymorphic SSRs to genotype 186 BC_1 maize (*Zea mays* L.) plants produced by crossing the inbred maize lines L-08-05 (donor parent) and L-14-4B (recurrent parent). The second backcross (BC_2) generation was produced with 180 plants and screened with markers which were not recovered from the first backcross (BC_1) generation. A total of 480 plants were evaluated in the third backcross (BC_3) generation from which 48 plants were selected for parental genotype recovery. Recurrent genotype recovery averages in three backcross generations were compatible with those expected in BC_4 or BC_5, indicating genetic gain due to the marker-assisted backcrossing. The target marker (polymorphic microsatellite PHI037) was efficiently transferred. Six markers per chromosome showed a high level of precision for parental estimates at different levels of maize genome saturation and donor alleles were not present in the selected recovered pure lines. Phenotypically, the plants chosen based on this criterion (06 markers per chromosome) were closer to the recurrent parent than any other selected by other criteria (02 or 04 markers per chromosome). This approach allowed the understanding that six microsatellites per chromosome is a more efficient parameter than 02 and 04 markers per chromosome for deriving a marker-assisted backcross (MAB) experiment in three backcross generations.

Key words: maize, marker-assisted backcrossing (MAB), marker-assisted selection, microsatellites.

Introduction

Backcross breeding in maize (*Zea mays* L.) has been extensively used to transfer favorable alleles for monogenic traits from donor genotypes to elite inbred lines (Openshaw *et al*., 1994) but high heritability polygenic traits have also been transferred through this method (Rinke and Sentz, 1961; Shaver, 1976). Two problems are inherent in backcross breeding, one being the large number of generations required to recover the genome of the inbred parent and the other being linkage drag, *i.e.* the presence of the inbred portions of the donor's genome (which can be linked to non-favorable alleles) surrounding the introgressed allele.

As previously described by Tanksley *et al*. (1989), the main advantage of using DNA markers as opposed to conventional selection is to accelerate the fixation of recipient alleles in non target regions and to identify the genotypes containing crossovers close to target genes (Ribaut and Hoisington, 1998). According to Frisch *et al*. (1999a), molecular markers are used in backcross breeding for two purposes: (i) as a diagnostic tool to trace the presence of a target allele when direct selection is difficult or impossible, such as the case of recessive alleles expressed late in plant development or quantitative trait *loci*. The use of markers as a diagnostic tool was first proposed by Tanksley (1983) and reviewed by Melchinger *et al*. (1990) and the term 'foreground selection' was suggested by Hospital and Charcosset (1997); and/or (ii) to identify individuals with a low proportion of undesirable genome from the donor parent, this approach being called 'background selection' and was first proposed by Tanskley *et al*. (1989) and then by Hillel *et al*. (1990) and was further investigated by Hospital *et al*. (1992) and later reviewed by Viescher *et al*. (1996).

Send correspondence to Luciana Lasry Benchimol. Instituto Agronômico de Campinas, Centro de Pesquisa e Desenvolvimento de Recursos Genéticos Vegetais,Caixa Postal 28, 13001-970 Campinas, SP, Brazil. E-mail: llasry@iac.sp.gov.br.

Several software programs, such as the PLABSIM program (Frisch *et al.,* 2000), are now available to make selection predictions using simulations. Hospital *et al*. (1992) investigated the use of markers for the recovery of the recipient genome during an introgression breeding program and showed that if marker assisted introgression is used it should be performed on three generations. These authors also recommended the use of markers with known map position and a density of two or three markers per 100 cM, because increasing this density results in only small benefits. Jarboe *et al*. (1994) have used the maize genome as a model for simulation and reported that three backcross generations and 80 markers were needed to recover 99% of the recurrent parent genotype.

The efficiency of marker-assisted backcrossing (MAB) selection as a breeding tool has also been evaluated considering the heritability of the target trait (Hospital *et al*., 1997; Knapp, 1998) and by monitoring the target genomic regions simultaneously as opposed to one by one (Hospital *et al*., 1997). Frisch *et al*. (1999a) also developed an approach to increase the efficiency of MAB in which success would depend on the carrier chromosome, the chromosomal position of the target locus, its distance to the flanking marker loci and the number of individuals evaluated. Performing simulations with the published maize map of 80 markers and phenotypic selection, Frisch *et al.* (1999b) also found that increasing the population size from the first backcross (BC_1) generation to the third backcross (BC_3) generation reduced the number of marker data points by as much as 50% without affecting the recurrent parent genotype proportion. Ribaut *et al*. (2002) presented simulations using different strategies using the maize genome as a model to compare allelic introgression with DNA markers through backcrossings.

Ragot *et al.* (1994) demonstrated that MAB could be efficiently used for introgressing a transgene construct containing the *Bt*-gene (the *Bacillus thuringiensis* toxin gene) of a transformed parent in an elite maize inbred, these workers reaching the same level of parent genotype recovery in BC_3 as that expected for the sixth backcross (BC_6) generation. Stuber (1994) using previously mapped favorable quantitative trait *loci* (QTLs) from two inbred lines and successfully transferred them to other inbred lines lacking these QTLs. Single genes with large effects conferring resistance to bacterial blight in rice have also been transferred using marker-assisted selection (Huang *et al*., 1997; Sanchez *et al.*, 2000).

When evaluating the length of the donor segment with the *Vrn-B1 locus* of first and third generation sister wheat lines Salina *et al.* (2003) concluded that if the selection is not molecular marker assisted plants with a long donor segment linked to the target gene may be randomly selected. The objective of our research was to evaluate a maize marker-assisted backcrossing scheme with emphasis on ascertaining how many markers per chromosome should be used and the number of generations in which the genotype of the parental line could be recovered. This research was carried out to verify the differential efficiencies of the number of markers per chromosome in background assisted backcross selection in maize by introgressing a target microsatellite marker.

Material and Methods

Plant material and backcrossing generations

The donor parent was the early-maturing orange flint kerneled inbred maize line L-08-05 (S_7; parent 1, coded as P_1) derived from line IG-1 and the recurrent parent was the early-maturing yellow dent kerneled inbred maize line L-14-4B (S_5; parent 2, coded as P_2) derived from line BR-106. These populations were allocated into different heterotic groups (Naspolini Filho *et al.*, 1981; Souza Jr *et al.*, 1993).

Only one F_1 ear was chosen from the P_1 x P_2 cross. The BC_1 generation resulted from the backcrossing of the F_1 generation with the L-14-4B (P_2) recurrent parent. We genotyped 186 BC_1 plants and selected 12 progenies, four plants for each selection criterion (SC) which were based on the number of markers per chromosomes ($SC_1 = 2$, $SC_2 = 4$ and $SC_3 = 6$). The second backcross (BC_2) generation was derived from the cross of selected BC_1 plants to P_1 plants and 180 BC_2 plants genotyped, of which we selected 15 plants per progeny/selection criterion to give a total of 180 plants. In the BC_3 generation we genotyped 480 BC_3 plants and selected ten plants per progeny/selection criterion, reaching a total of 160 plants per selection criterion. Of the 480 plants genotyped, four plants were selected per progeny/selection criterion, making a total of 16 plants per selection criterion and a total of 48 selected lines.

The differences in the number of plants evaluated in each selection cycle (BC_1 to BC_3) were due to the need to adapt the MAB scheme to practical laboratory condition. Since each PCR-plate supported 96 DNA samples and each microsatellite *locus* needed P_1, P_2 and F_1 control profiles there were 93 empty wells for each plate and two PCR plates were needed for 186 samples, representing a total of 106 amplifications in the first selection cycle (BC_1). The second cycle (BC_2) was made up of the fifteen samples selected from each criterion (SC_1 to SC_3), which represented a total of 60 plants per selection criterion and 180 plants genotyped for the BC_2 generation. In the third selection cycle (BC_3) the number of plants selected per criterion was reduced from fifteen to ten due to the amount of DNA samples to be manipulated. In all, 480 DNA samples were amplified with the non-recovered microsatellites.

Molecular selection

As outlined above, our molecular approach for selecting plants in each backcross generation was based on three selection criteria which differed in respect to the number of

markers per chromosome ($SC_1 = 2$, $SC_2 = 4$ and $SC_3 = 4$). More than 250 microsatellites (synthesized by GIBCO BRL, São Paulo, SP) were screened and used to genotype the BC_1 generation but only 53 were shown to be polymorphic and thus used in the final analysis. Besides their use for polymorphic profiling, markers were also chosen for their map position and distribution along the chromosomes. Maize genome has been dissected into 100 evenly spaced "bins" of approximately 20 cM each (Gardiner *et al.* 1993).

For SC_1 one marker per chromosome arm was chosen (maintaining a standard distance from the centromere to increase recombination and decrease linkage drag) while for SC_2 we chose two markers per chromosome arm and for SC_3, approximately three markers per chromosome arm were selected (Table 1). Molecular criteria were used to compare the recovery of the recurrent parent genotype and the effectiveness of the marker number increase per chromosome used during the selection process.

Table 1 - Microsatellites (SSRs) used in different selection criteria and their respective bins. Saturated (*Sat*) refers to the extra primer-set used to amplify the 48 selected third backcross (BC_3) plants and saturate maize genome. The selection criteria (SC) were the number of markers per chromosomes ($SC_1 = 2$, $SC_2 = 4$ and $SC_3 = 6$).

SSRs	SC	Bin	SSRs	SC	Bin
BNLG 1627	$SC_1/SC_2/SC_3$	1.02	PHI 126	$SC_1/SC_2/SC_3$	6.00
BNLG 2238	SC_2/SC_3	1.04	PHI 077	SC_2/SC_3	6.01
BNLG 1598	SC_3	1.06	BNLG 1371	SC_2/SC_3	6.02
BNLG 615	SC_3	1.07	UMC 1614	*Sat*	6.04
*PHI 037**	All	1.08	NC013	$SC_1/SC_2/SC_3$	6.05
PHI 011	$SC_1/SC_2/SC_3$	1.09	UMC1520	*Sat*	6.06
PHI 120	SC_2/SC_3	1.11	BNLG 1759	SC_3	6.07
BNLG 125	SC_2/SC_3	2.02	UMC 1426	*Sat*	7.00
BNLG 2248	SC_3	2.03	UMC 1632	*Sat*	7.01
BNLG 166	$SC_1/SC_2/SC_3$	2.04	PHI 114	SC_2/SC_3	7.03
DUPSSR 21	SC_3	2.05	BNLG 434	$SC_1/SC_2/SC_3$	7.03
BNLG 1396	SC_3	2.06	DUPSSR 13	SC_2/SC_3	7.04
UMC 1560	*Sat*	2.07	UMC 1407	*Sat*	7.05
BNLG1520	$SC_1/SC_2/SC_3$	2.09	PHI 116	SC_2/SC_3	7.06
UMC1394	*Sat*	3.01	UMC 1139	*Sat*	8.01
BNLG 1144	$SC_1/SC_2/SC_3$	3.03	PHI 119	$SC_1/SC_2/SC_3$	8.02
BNLG 602	SC_3	3.04	PHI 115	SC_3	8.03
MMC 0022	SC_2/SC_3	3.05	BNLG 1863	SC_2/SC_3	8.04
BNLG 420	SC_2/SC_3	3.05	BNLG 1176	SC_3	8.05
UMC 1659	*Sat*	3.07	BNLG 1823	$SC_1/SC_2/SC_3$	8.07
BNLG 1536	*Sat*	3.09	DUPSSR14	SC_2/SC_3	8.09
BNLG 1098	$SC_1/SC_2/SC_3$	3.10	DUPSSR 06	$SC_1/SC_2/SC_3$	9.02
UMC 1276	*Sat*	4.01	BNLG 430	SC_2/SC_3	9.03
UMC 1652	*Sat*	4.04	UMC 1107	*Sat*	9.04
PHI 026	$SC_1/SC_2/SC_3$	4.05	BNLG 1012	SC_2/SC_3	9.05
BNLG 252	SC_3	4.06	BNLG 292	SC_3	9.06
DUPSSR 34	SC_2/SC_3	4.07	BNLG 128	$SC_1/SC_2/SC_3$	9.07
BNLG 2244	*Sat*	4.08	BNLG 619	*Sat*	9.07-9.08
BNLG 1917	$SC_1/SC_2/SC_3$	4.10	PHI 117	*Sat*	10.00
BNLG 589	SC_2/SC_3	4.11	UMC 1318	*Sat*	10.01
BNLG 1006	$SC_1/SC_2/SC_3$	5.00	BNLG1451	$SC_1/SC_2/SC_3$	10.02
PHI 113	SC_2/SC_3	5.03-5.04	BNLG 640	SC_2/SC_3	10.03
DUPSSR 10	SC_2/SC_3	5.04	BNLG 1526	SC_2/SC_3	10.04
BNLG 278	SC_3	5.05-5.06	UMC 1930	*Sat*	10.05
UMC 1019	$SC_1/SC_2/SC_3$	5.06	BNLG 1839	$SC_1/SC_2/SC_3$	10.07
PHI 128	SC_3	5.07			
UMC 1153	*Sat*	5.09			

*PHI037 was transfered as a microsatellite allele for all selection criteria.

A polymorphic microsatellite, PHI037 (bin 1.08), was elected and used as a target locus to be introgressed. Apart from the reliable pattern, PHI037 also presented polymorphic flanking microsatellites (PHI011, at bin 1.09 for SC_1/SC_2; BNLG615, at bin 1.07 for SC_3) that could contribute to diminishing the linkage drag around the locus. Every plant selected satisfied the following conditions: (1) the introgressed marker was heterozygous; (2) the plant showed the highest recovery percentage of the recurrent parent genotype, (3) no donor allele was present in the genotyping of the individual plant.

DNA extraction followed the procedures described in Hoisington *et al.* (1994). Each amplification was carried out using DNA from the donor (P_1) and recurrent (P_2) parent and the F_1 generation. Only the non-recovered microsatellites at BC_1 were used to screen the BC_2 and BC_3 generations. The 48 BC_3 selected plants (BC_{3sel}) were analyzed with microsatellites not used in previous generations, making total of 72 microsatellites spaced at an average of 25 cM (Table 1). The SC_1 and SC_2BC_{3sel} plants were also screened with microsatellites used in genotyping for the other criterion. Reactions were carried out using the Touchdown PCR (Don *et al.,* 1991). Products were separated by 4% (w/v) 1:1 metaphor:agarose gel electrophoresis.

Molecular data analysis

Results were scored as A for the P_1 allele (coded as Genome donor, GD), B for the P_2 allele (coded as Genome recurrent, GR) and H for the F_1 pattern of P_1 and P_2 alleles. The percentage of the P_2 recurrent parent genome present in each genotyped plant was estimated using the expression GR% = [B+ (0.5H)/(B+H+A)] x 100, corresponding to the number of B alleles present in the genotype of each genotyped plant. The introgressed locus was taken from the computation of the total recoveries of recurrent parent genotypes.

The BC_1 and BC_2 percentage recovery means were compared to the expected averages for each backcross generation. The 48 selected BC_3 plants were analyzed considering saturating (BC_{3sel}^{sat}) and non-saturating ($BC_{3sel}^{non-sat}$) conditions. To compare the recovery percentages between the selected plants the data were transformed using the arcsine function. For each selection criterion *t*-tests were performed to compare the percentage recoveries of the P_2 parent BC_3 lines under saturated and non-saturated conditions for an expected average of 93.75% and a saturated average of 100%.

Experimental evaluation and statistical analysis

The 48 BC_3 selected lines were evaluated during the 2002/2003 planting season near the city of Piracicaba in the Brazilian state of São Paulo at three experimental stations (Department of Genetics, Areão, and Caterpillar) belonging to São Paulo State University (ESALq/USP). We evaluated 49 entries (48 selected BC_3 lines plus the P_2 parental line) using a 7 x 7 lattice design with three replicates per lo-

cation. Plots were 4 m long with 0.80 m between rows and 0.20 m between plants within rows and were over-seeded with 40 seeds per plot, later thinned to 20 plants per plot (equivalent to 62,500 plants ha^{-1}). Data for grain yield (GY, grams per plant), number of kernel-rows per ear (NKE), number of rows per ear (NRE), ear length (EL), ear diameter (ED), plant height (PH), ear height (EH), days to silk extrusion (DSE), days to pollen shed (DPS), number of plants per plot (stand) and grain moisture were recorded in all environments. The NKE, NRE, EL, ED, PH, and EH traits were recorded on five competitive plants per plot and the plot means were used for analyses. The DSE and DPS traits were recorded in each plot as days from planting to the time when 50% of plants per plot showed this trait. Grain yield of each plot was adjusted for average stand by covariance analyses, and for 15% grain moisture before analyses.

Analysis of variance was computed for each location and, subsequently, joint analysis of variance was computed across locations for each trait. The BC_3 Line source of variation was partitioned into the three selection criteria (SC_1, SC_2 and SC_3) and among selection criteria (ASC), and the interaction BC_3 lines x Location (BC_3 Lines x L) was partitioned accordingly (SC_1 x L, SC_2 x L, SC_3 x L and ASC x L). The Dunnett test was used to test whether the traits analyzed from each BC_3 selected line differed significantly or not from the P_2 recurrent parental inbred (L-14-4B). All analyses were performed using the SAS software, proc GLM (SAS Institute, 1989). Unless otherwise indicated all stated significances were at p ≤ 0.01.

Results

The three molecular criteria showed an average of 74.3% recovery of the P_2 genotype for the first backcross generation. This average did not differ from the expected value (75%); little variation was detected among the three criteria with the maximum value being 87.5%. The average value of the four selected plants in the BC_1 generation was 83.8% for SC_1, 83.1% for SC_2 and 82.7% for SC_3 (Figures 1a, 1b and 1c respectively). In the BC_2 generation, the three criteria showed an average of 91.2% and maximum value of 98.1%. The mean expected value for this generation, without selection, was 87.5%. The average value of the 16 selected plants in the BC_2 generation was 94.4% for SC_1, 92.5% for SC_2 and 92.1% for SC_3 (Figures 1d, 1e and 1f respectively). In the BC_3 generation, the three criteria showed an average of 96.6%, which was significantly higher than the expected mean (93.75%) for this generation. The maximum value was 100% (without considering the PHI037 allele). The average value of the 16 selected plants in the BC_3 generation was 97.3% for SC_1, 96.2% for SC_2 and 96.4% for SC_3 (Figures 1g, 1h and 1i respectively).

The presence of alleles derived from the donor parent was not detected in the non-saturated condition. In the selected saturated BC_3 lines, some donor alleles were de-

tected in bins that had not been assayed previously for selection (*sat* primers, Table 1). The average number of donor alleles were reduced with the increment of markers per chromosome used for selection, *i.e.* 0.75 ± 0.25 for SC_1, 0.44 ± 0.16 for SC_2, and 0.13 ± 0.09 for SC_3. These results expressed the average frequency of donor alleles per plant and their associated mean errors.

Although the BC_{3sel}^{sat} and $BC_{3sel}^{non-sat}$ conditions considered plants with the same genetic background, the genome saturation estimates of the genetic content for P_2 parent recovery were different. The 16 selected BC_3 lines in three selection criteria were saturated with more markers, and the mean values for P_2 parent recovery showed that the averages for the selected saturated BC_3 lines were significantly lower for SC_1 and SC_2 but were not significant for SC_3. There was a large and significant decrease in P_2 genotype recovery for SC_1 (from 99.8% to 92.4%) and for SC_2 (99.1% to 95.5%) with genome saturation. The increase in the number of markers was also significantly reduced the difference between the averages and standard deviations for $BC_{3sel}^{non-sat}$ and BC_{3sel}^{sat}. All selection criteria, except $BC_{3sel}^{non-sat}$ and BC_{3sel}^{sat} for SC_1, were significantly superior to the expected mean value (93.75%) and significantly inferior to the maximum value (100%). The value for $SC_1/BC_{3sel}^{non-sat}$ did not differ significantly from the maximum value (100%) whereas SC_1/BC_{3sel}^{sat} was significantly
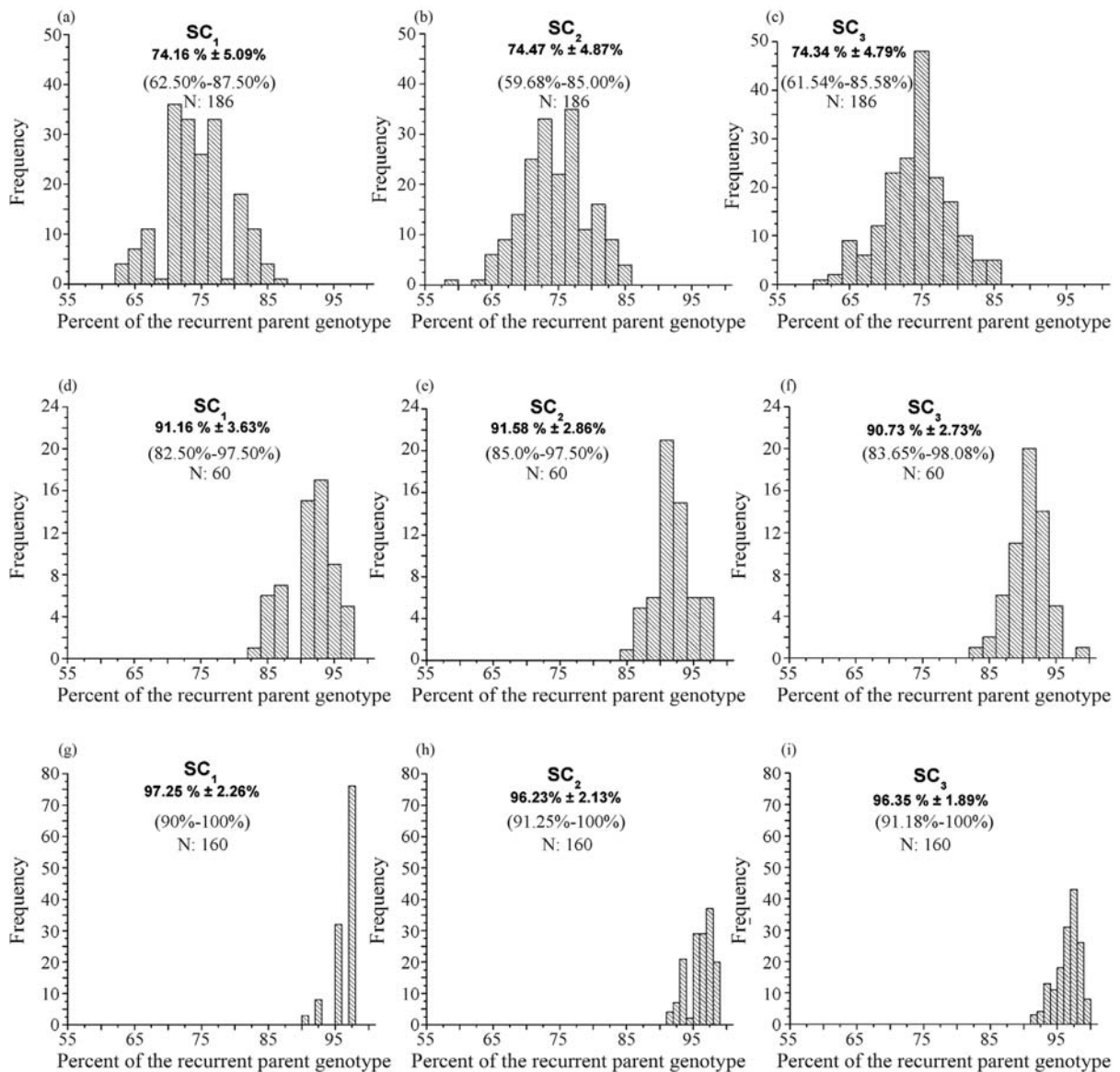

Figure 1 - Distribution of P_2 recurrent parent recovery for the three selection criteria (SC_1, SC_2, SC_3) for the three backcross generations (BC_1, BC_2 and BC_3).

inferior to the expected mean value (93.75%) for this generation. It should be emphasized that the SC_3 criterion differed significantly from SC_2 and SC_1 for BC_{3sel}^{sat} (Table 2).

No significance was detected between different selection criteria for nine phenotypic traits in the variance analyses and the among selection criteria (ASC) level was insignificant for all nine traits. For the BC_3 lines, significance was detected for most of the traits (Table 3). Therefore, the analysis of the phenotypic data did not detect differences among the three selection criteria, but there were significant differences between the BC_3 lines within each selection.

Table 2 - Comparisons of third backcross (BC_3) lines as a percentage of recurrent parent genotype for the expected mean (93.75%) and 100% of the BC_3 generation under non-saturated (*non-sat*) and saturated (*sat*) conditions.

					Selection criteria						
$SC_1^{\dagger}$				$SC_2^{\dagger}$				$SC_3^{\dagger}$			
Mean	Sd	N	R	Mean	Sd	N	R	Mean	Sd	N	R
BC_3^{all}				BC_3^{all}				BC_3^{all}			
97.25**	2.26	160	10	96.23**	2.13	160	8.75	96.35**	1.89	160	8.82
$BC_{3sel}^{non-sat}$				$BC_{3sel}^{non-sat}$				$BC_{3sel}^{non-sat}$			
99.84‡	0.63	16	2.50	99.06**	0.72	16	2.50	97.47**	0.46	16	1.05
BC_{3sel}^{sat}				BC_{3sel}^{sat}				BC_{3sel}^{sat}			
92.40§	1.90	16	5.90	95.50**	1.50	16	4.30	96.90**	1.20	16	3.50

†Sd: standard deviation; N: number of plants analysed; R: range between maximum and minimum values.
*, **: Significant at the 0.05 and 0.01 probability levels, respectively.
§: Significant at the 0.05 probability level for 93.75%, and significant at 0.01 probability level for 100%.

Table 3 - Analyses of variance mean square and significances values for nine traits: grain yield (GY), number of kernel rows per ear (NKE), number of rows per ear (NRE), ear length (EL), ear diameter (ED), plant height (PH), ear height (EH), days to silk extrusion (DSE), and days to pollen shed (DPS). These results are for the third backcross (BC_3) and the selection criteria (SC) were the number of markers per chromosomes ($SC_1 = 2$, $SC_2 = 4$ and $SC_3 = 6$).

S.V.	D.F.	GY ($g\ p^{-1}$)	EL ($cm\ ear^{-1}$)	$ED^{\#}$ ($cm\ ear^{-1}$)	NKE (number)	NRE (number)	PH ($cm\ p^{-1}$)	EH ($cm\ p^{-1}$)	DSE (days)	DPS (days)
						Mean squares				
Locations(L)	2	2,465.05**	4.12*	32.40*	186.15**	2.85*	1,044.21**	634.26**	5.65	3.53
Reps/L	6	401.97**	2.94*	19.87*	45.03**	4.32**	420.31*	252.28*	4.60*	3.00**
Entries(E)	48	683.20**	4.93*	56.56**	49.86**	3.59**	887.20**	481.04**	8.29**	6.33**
BC_3 Lines x Check	1	769.44	6.12	48.92	47.46	4.37	976.42	593.34	7.47	6.39
BC3 Lines	47	681.36**	4.90**	56.72**	49.91**	3.57**	885.30**	478.65**	8.31**	6.33**
SC_1	15	795.39**	5.62**	56.27**	54.58**	3.53**	1,126.36**	584.67**	10.42**	7.80**
SC_2	15	757.68**	5.11**	61.76**	45.57**	3.69**	998.15**	582.81**	10.01**	7.17**
SC_3	15	549.47**	3.91**	57.12**	43.73**	3.36**	588.20	301.07	5.07	4.41
ASC	2	240.65	5.43	19.31	93.68	4.65	459.20	234.23	3.93	3.32
E x L	96	247.54**	1.34	10.26	14.88	1.04	357.42**	187.58**	3.49**	2.54**
(BC_3 Lines x Check) x L	2	364.50*	1.38	12.56	12.25	1.10	705.00*	326.55*	4.67	3.66
BC_3 Lines x L	94	245.05**	1.34	10.21	14.94	1.04	350.02**	184.62**	3.46**	2.52**
SC_1 x L	30	234.50**	1.45	10.27	13.68	1.10	372.16**	194.75**	3.36*	2.60*
SC2 x L	30	268.16**	1.27	9.40	15.96	0.95	358.50**	192.61**	3.62*	2.40*
SC3 x L	30	248.24**	1.31	10.89	14.80	1.06	313.78*	162.17*	3.28*	2.47*
ASC x L	4	127.02	1.17	10.85	17.98	1.18	392.18	217.03	4.46	3.18
Error	234	113.55	1.16	8.78	12.72	0.95	194.59	98.00	1.98	1.51
General mean		85.11	13.89	3.65	34.46	10.40	190.57	110.73	60.64	62.00
SC1 mean		86.12	13.75	3.58	34.20	10.36	192.05	111.01	60.22	62.30
SC2 mean		85.02	14.02	3.76	34.72	10.85	189.62	109.94	60.82	62.00
SC3 mean		84.18	13.68	3.66	34.90	10.10	190.24	110.75	60.74	61.86
Check mean		85.10	14.10	3.61	34.02	10.30	190.36	111.22	60.76	61.87
CV%		12.52	7.75	8.12	10.35	8.70	7.32	8.94	2.32	1.98

The 48 selected BC_3 lines were compared with the P_2 parent line using Dunnett's test (Table 4) without Bonferroni's correction for the probability level (Zar, 1999). For each selection criterion, 120 tests were performed and 88 tests (73.3%) were significant for SC_1, 52 (43.3%) were significant for SC_2 and 8 (6.7%) were significant for SC_3. In relation to the SC_1 criterion, all BC_3 lines differed significantly from the P_2 parent for at least one trait, and six lines differed significantly for nine traits simultaneously while the other ten lines diverged significantly for at least three traits simultaneously. For the SC_2 criterion all BC_3 lines differed significantly from the P_2 parent in at least one trait, with four lines differing simultaneously for all characters and the other 12 lines differing significantly for at least two traits simultaneously. However, for the SC_3 criterion, ten of the sixteen lines did not differ significantly from the P_2 parent for any of the traits, although six lines did show significant differences for two characters simultaneously but none of these lines differed significantly for the PH, EH, DSE or DPS traits, the same behavior being observed in the analysis of variance (Table 3) for these same four traits as regards the SC_3 criterion.

Discussion

The plant breeding community is enthusiastic about marker-assisted selection, but a link between theory and practice is still missing. Parameters are defined and simulations represent suitable approaches providing a strong contribution to the goal of using marker-assisted backcrossing (MAB) in crop improvement (Hospital *et al.,* 1992; Tanksley and Nelson, 1996; Visscher *et al*., 1996; Hospital

Table 4 - Number and percentage of significant Dunnett tests for nine traits in third backcross (BC_3) lines compared to the P_2 recurrent parent inbred line L-14-4B.

Trait[#]	Number of significant tests*
	SC_1	(%)	SC_2	(%)	SC_3	(%)	Total	(%)
GY	14	(87.50)	6	(37.50)	2	(12.50)	22	(45.83)
EL	10	(62.50)	6	(37.50)	1	(6.25)	17	(35.42)
ED	6	(37.50)	4	(25.00)	1	(6.25)	11	(22.92)
NKE	11	(68.75)	7	(43.75)	2	(12.50)	20	(41.67)
NRE	7	(43.75)	5	(31.25)	2	(12.50)	14	(29.17)
PH	10	(62.50)	7	(43.75)	0	(0.00)	17	(35.42)
EH	11	(68.75)	6	(37.50)	0	(0.00)	17	(35.42)
DSE	10	(62.50)	6	(37.50)	0	(0.00)	16	(33.33)
DPS	9	(56.25)	5	(31.25)	0	(0.00)	14	(29.17)
Total	88	(73.33)	52	(43.33)	8	(6.67)	148	(41.11)
Total of tests	120		120		120		360	

[#]EW: ear weight, NKE: number of kernel rows per ear, NRE: number of rows per ear, EL: ear length, ED: ear diameter, PH: plant height, EH: ear height, DSE: days to silk extrusion and DPS: days to pollen shed. *SC1, SC2, SC3: Selection Criteria.

and Charcosset, 1997; Knapp, 1998; Frisch *et al*., 1999a; Frisch *et al*., 1999b; Frisch *et al*., 2000; Reyes-Valdés, 2000). However, most of the theoretical papers related to marker-assisted selection present complex mathematical models, making it difficult to directly derive a practical experiment. Differences in laboratory strategies such as different DNA markers are rarely taken into account in theoretical papers (Ribault *et al*., 2002) and the practical implications of the use of different marker technologies should be considered to achieve selection. Newer and simpler DNA marker systems have been developed and even the technological constraints of early restriction fragment length polymorphism (RFLP) methods seemed to have been overcome (Ribaut *et al*., 1997), with high-density DNA marker maps having been constructed for almost every important crop species (OBrien, 1993). According to Young (1999), more than 400 articles containing the key words 'marker-assisted breeding' or 'marker-assisted selection' can be found in the 1995 to 1999 Current Contents, yet few if any, actually describe the use of marker-assisted technology leading to the release of varieties and few give experimental results validating the efficiency of marker-assisted methodologies in different crop species (Ragot *et al.,* 1994; Stuber, 1994; Salina *et al*., 2003; Lecomte *et al*., 2004).

Sometimes the most efficient strategy may lose its clarity as different theoretical approaches can serve the same purpose. Population increase, when advancing the backcross generations, reduces the number of markers needed in contrast to a constant population size across all generations (Frisch *et al*., 1999b), while another way of reducing the number of markers would be to increase the number of markers in each new backcross generation (Hospital *et al*., 1992). In the study described in this paper the strategy adopted was to maintain the population size in the first and second backcross generations but increase the number of plants in the third backcross generation. The number of markers in the first scheme was reduced to the non-target locus at each generation as only the non-recovered plants were screened in the following generation. Other markers were used to saturate the genome of the selected genotypes obtained from the BC_3 selection.

In order to obtain an economical and viable strategy only a few markers were used to assist selection in the BC_1 and BC_2 generations but from the BC_3 generation other markers were added for the evaluation of the recovered lines and in the end the spacing between markers reached 25 cM across the ten maize chromosomes. Frisch *et al*. (1999a) and Jarboe *et al*. (1994) both used the published maize map of 80 markers for their simulation and found promising results. As has been previously demonstrated (Hospital *et al*., 2002), response to selection would be reduced if a marker from an unknown chromosomal location had been used. Our approach demonstrated that microsatellites could be used efficiently for introgressing a target

allele, simulating a monogenic trait, without any intermediate field selection. The idea of using a marker was to control all the steps for the introgression of the target locus from the donor to the recipient line, this approach being neutral and simple to monitor, although, as in a simulation, no target trait could be followed. All the conditions were adjusted so that they would be practical and fit into the daily laboratory routine. The P_2 recurrent genotype recovery averages in three backcross generations were compatible to the those expected for the fourth and fifth backcrosses (BC_4 and BC_5) which shows that marker-assisted backcrossing produced a genetic gain as regards P_2 recurrent parent recovery, this result agreeing with previously published work (Hospital *et al.*, 1992; Frisch *et al.*, 1999a; Ragot *et al.*, 1994; Jarboe *et al.*, 1994).

Hospital *et al.* (1992) have shown that increasing the number of markers to more than three per non-carrier chromosome was not efficient in early generations, but Ribault *et al.*, (2002) have pointed out that because there is an increased probability of crossover in later generations an increase in the number of markers should be considered as a way of optimizing selection.

The approach outlined in this paper was able to transfer the target allele, the efficiency of transfer at both the molecular and phenotypic level being higher when six markers per chromosome were used. The use of six markers also resulted in a greater level of precision for P_2 recurrent genotype recovery estimates under different conditions of maize genome saturation. Alleles derived from the P_1 donor parent were not detected under non-saturated conditions with six markers per chromosome while under saturated conditions; donor alleles were still detected in unassayed bins for those plants that had been selected using a low number of markers per chromosome.

When the number of markers per chromosome arm was increased there was a reduction in the number of significant tests, indicating that increasing the number of markers during marker-assisted backcrossing resulted in plants that were closer to the P_2 recurrent parent. The SC_3 criterion (six markers per chromosome) was shown to be superior to the SC_1 and SC_2 criteria in that it presented the lowest number of significant tests and the fewest plants with significant tests for more than one trait simultaneously. At the phenotypic level a high number (*i.e.* 6) of markers per chromosome increased the efficiency of MAB in such a way that the final SC_3/BC_3 selected plants were closer to the P_2 parental type than to any of the plants selected using the SC_1 and SC_2 criteria.

This practical approach could be fully optimized by associating field selection for enhanced parental genotype recovery evaluation with a specific transferred phenotypic trait. We agree with the observation by Ribault *et al.* (2002) that the application of a backcross marker-assisted strategy for practical experiments should be on a case-to-case basis

and that it is important to consider the nature of the germplasm involved.

Acknowledgments

References

Don RH, Cox PT, Wainwright BJ, Baker K and Mattick JS (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucl Acids Res 19:4008.

Frisch M, Bohn M and Melchinger AE (1999a) Minimum sample size and optimum positioning of flanking markers in marker-assisted backcrossing for transfer of a target gene. Crop Sci 39:967-975.

Frisch M, Bohn M and Melchinger AE (1999b) Comparison of selection strategies for marker-assisted backcrossing of a gene. Crop Sci 39:1295-1301.

Frisch M and Melchinger AE (2000) The length of the intact donor chromosome segment around a target gene in marker-assisted backcrossing. Genetics 157:1343-1356.

Gaitán-Solís E, Duque MC, Edwards KJ and Tohme J (2002) Microsatellite repeats in bean (*Phaseolus vulgaris*): Isolation, characterization and cross-species amplification in *Phaseolus ssp.* Crop Sci 42:2128-2136.

Gardiner J, Coe EH, Melia-Hancock S, Hoisington DA and Chao S (1993) Development of a core RFLP map in maize using an Immortalized-F_2 population. Genetics 134:917-930.

Hillel J, Schaap T, Haberfeld A, Jeffreys J, Plotzky Y, Cahaner A and Lavi U (1990) DNA fingerprints applied to gene introgression in breeding programs. Genetics 124:783-789.

Hoisington D, Khairallah M and Gonzalez-de-Lion D (1994) Laboratory Protocols: CIMMYT Applied Molecular Genetics Laboratory. 2nd edition. CIMMYT, Mexico DF.

Hospital F, Chevalet C and Mulsant P (1992) Using markers in gene introgression breeding programs. Genetics 132:1199-1210.

Hospital F and Charcosset A (1997) Marker-assisted introgression of quantitative trait *loci*. Genetics 147:1469-1485.

Hospital F, Moreau L, Lacoudre F, Charcosset A and Gallais A (1997) More on the efficiency of marker-assisted selection. Theor Appl Genet 95:1181-1889.

Huang N, Angeles ER, Domingo J, Magpantay G, Singh S, Zhang G, Kumaravadivel N, Bennett J and Khush GS (1997) Marker-assisted selection using RFLP and PCR. Theor Appl Genet 95:313-320.

Jorboe SG, Beavis WD and Openshaw (1994) Prediction of responses to selection in marker-assisted backcross programs by computer simulation. Abstracts of the Second International Conference on Plant Genome, Scherago International Inc., pp 38.

Knapp SJ (1998) Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. Crop Sci 38:1164-1174.

Lecomte L, Duffé P, Baruet M, Servin B, Hospital F and Causse M (2004) Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. Theor Appl Genet, Springer-Verlag 2004, eletronic published a-head of print as original paper: 27 April 2004. Available from: http://link.springer-ny.com/search.htm.

Melchinger AE, Lee M, Lamkey KR, Hallauer AR and Woodman WL (1990) Genetic diversity for restriction fragment length polymorphisms and heterosis for two diallel sets of maize inbreds. Theor Appl Genet 80:488-496.

Naspolini Filho V, Gama EEG, Vianna RT and Môro JR (1981) General and specific combining ability for yield in a diallel cross among 18 maize populations. Rev Bras Genet 4:571-577.

Openshaw SJ, Jarboe SG and Beavis WD (1994) Marker-assisted selection in backcross breeding. Proceedings of the Symposium "Analysis of Molecular Marker Data", Joint Plant Breeding Symposia Series, American Society for Horticultural Science/Crop Science of America, Corvallis, Oregon, pp 41-43.

O'Brien SJ (1993) Genetics Maps: *Locus* Maps of Complex Genomes. Cold Spring Harbor Laboratory, Cold Spring Harbor, pp 41-43.

Ragot M, Biosiolli M, Delbut MF, Dell'Orco A, Malgarini L, Thevenin P, Vernoy J, Vivant J, Zimmermann R and Gray G (1994) Marker-assisted backcrossing: A practical example. In: Berville A and Tersac M (eds) Techniques et Utilizations des Marqueurs Moleculaires. INRA Editions, Versailles, pp 29-31.

Reyes-Valdés MH (2000) A model for marker-based selection in gene introgression breeding programs. Crop Sci 40:91-98.

Ribaut JM, Jiang C, González-de-Léon D, Edmeades GO and Hoisington DA (1997) Identification of quantitative trait *loci* under drought conditions in tropical maize. 2. Yield components and marker-assisted selection strategies. Theor Appl Genet 94:887-896.

Ribaut JM and Hoisington D (2002) Simulation experiments on efficiencies of gene introgression by backcrossing. Crop Sci 42:557-565.

Rinke EH and Sentz JC (1961) Moving corn-belt germplasm northward. Ann Hybrid Corn Industry Conf 16:53-56.

Salina E, Dobrovolskaya O, Efremova T, Leonova I and Röder MS (2003) Microsatellite monitoring of recombination around the *Vrn-B1 locus* of wheat during early backcross breeding. Plant Breed 122:116-119.

Sanchez AC, Brar DS, Huang N, Li Z and Khush GS (2000) Sequence tagged site marker-assisted selection for three bacterial blight resistance genes in rice. Crop Sci 40:792-797.

SAS Intitute (1989) SAS/SAT Use'rs Guide, version 6, v 2, SAS Institute, Cary, 846 pp.

Shaver DL (1976) Conversions for earliness in maize inbreds. Maize Genet Coop Newslttr 50:20-23.

Souza Jr CL, Santos MX, Mangavaca R and Gama EEG (1993) Estimativas de parâmetros genéticos na interpopulação de milho BR-105 x BR-106 e suas implicações no melhoramento. Pesq Agropec Bras 28:473-479.

Stuber CW (1994) Enhancement of grain yield in maize using marker-facilitated introgression of QTLs. Proceedings of the Symposium "Analysis of Molecular Marker Data", Joint Plant Breeding Symposia Series, American Society for Horticultural Science/Crop Science of America, Corvallis, Oregon, pp 44-46.

Tanksley SD (1983) Molecular markers in plant breeding. Plant Mol Biol Rep 1:3-8.

Tanksley SD and Nelson JC (1996) Advanced backcross QTL analysis: A method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet 92:191-203.

Tanksley SD, Young ND, Paterson AH and Bonierbale MW (1989) RFLP mapping in plant breeding: New tools for an old science. Bio/Tech 7:257-264.

Visscher PMS, Haley CS and Thompsom R (1996) Marker assisted introgression in backcross breeding programs. Genetics 144:1923-1932.

Young ND (1999) A cautiously optimistic vision for marker-assisted breeding. Mol Breed 5:505-510.

Young ND and Tanksley SD (1989) Restriction Fragment length polymorphism maps and the concept of graphical genotypes. Theor Appl Genet 77:95-101.

Yu K, Park SJ, Poysa V and Gepts P (2000) Integration of simple sequence repeat (SSR) markers into a molecular linkage map of common bean (*Phaseolus vulgaris* L.). J Her 91:429-434.

Zar JH (1999) Biostatistical Analysis. 4th edition. Prentice Hall Inc., Upper Saddle River, New Jersey, 663 pp.