

**EFICIÊNCIA DOS DIALELOS CIRCULANTES
NA ESCOLHA DE GENITORES,
AVALIADA COM SIMULAÇÃO DE DADOS**

RUBEN DELLY VEIGA

1998

DATA DE DEVOLUÇÃO
no registro» CIRCULAÇÃO E EMPRESTII

Ruben Delly Veiga

**EFICIÊNCIA DOS DIALELOS CIRCULANTES
NA ESCOLHA DE GENITORES,
AVALIADA COM SIMULAÇÃO DE DADOS**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Curso de Doutorado em Agronomia, área de concentração em Genética e Melhoramento de Plantas, para obtenção do título de "Doutor".

Orientador

Daniel Furtado Ferreira

Prof. Dr/ UFLA

LAVRAS
MINAS GERAIS – BRASIL
1998

**Ficha Catalográfica preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Veiga, Ruben Delly

**Eficiência dos dialelos circulantes na escolha de genitores, avaliada com
simulação de dados / Ruben Delly Veiga. – Lavras : UFLA, 1998**

95 p. : il.

Orientador : Daniel Furtado Ferreira

Tese (Doutorado) – UFLA.

Bibliografia.

**1. Dialelo parciais. 2. Simulação. 3. Eficiência. 4. Melhoramento Genético. I.
Universidade Federal de Lavras. II. Título.**

CDD-631.523

RUBEN DELLY VEIGA

**EFICIÊNCIA DOS DIALELOS CIRCULANTES
NA ESCOLHA DE GENITORES,
AVALIADA COM SIMULAÇÃO DE DADOS**

Tese apresentada à Universidade Federal de Lavras,
como parte das exigências do Curso de Doutorado em
Agronomia, área de concentração em Genética e
Melhoramento de Plantas, para obtenção do título de
"Doutor".

APROVADA em 27 de novembro de 1998.

Eduardo Bearzoti

Prof. Dr. / UFLA

Elaine Aparecida de Souza

Prof. Dr. / UFLA

Isaias Olívio Geraldi

Prof. Dr. / ESALQ-USP

Magno Antônio Patto Ramalho

Prof. Dr. / UFLA



Daniel Furtado Ferreira

Prof. Dr. / UFLA

(Orientador)

LAVRAS
MINAS GERAIS - BRASIL

*“A razão sozinha não tem
nenhuma capacidade de revisão: ela
o consegue somente em combinação
com a observação”*

Sir Francis Bacon, 1620

DEDICO

SUMÁRIO

RESUMO	i
ABSTRACT.....	iii
1 INTRODUÇÃO.....	01
2 REFERENCIAL TEÓRICO.....	03
2.1 Escolha de populações segregantes.....	03
2.2 Cruzamentos dialélicos.....	04
2.3 Análise de cruzamentos dialélicos.....	08
2.4 O Emprego de dialelos dirculantes.....	11
2.5 Uso da simulação de dados.....	15
3 METODOLOGIA.....	21
3.1 Dialelos simulados.....	21
3.2 Simulação dos dialelos circulares.....	25
3.3 Eficiência dos dialelos circulares.....	26
3.3.1 Correlação classificatória da capacidade geral de combinação (CGC) dos pais.....	26
3.3.2 Correlação classificatória da capacidade específica de combinação (CGC) dos pais e híbridos.....	26
3.3.3 Proporção de estimativas de CGC dentro do intervalo de confiança...	26
3.3.4 Proporção de estimativas de CEC dentro do intervalo de confiança...	27
3.4 Dados de dialelos completos obtidos na literatura.....	28
4 RESULTADOS E DISCUSSÃO.....	29
4.1 Dialelos simulados.....	29
4.2 Dialelos de dados experimentais obtidos na literatura.....	49
5 CONCLUSÕES.....	59
6 REFERÊNCIAS BIBLIOGRÁFICAS.....	61
7 ANEXOS.....	67

LISTA DE TABELAS

Tabela	Página
1 Número de cruzamentos dialélicos em função do número de pais , na presença e na ausência dos recíprocos (Haluaer e Miranda Filho, 1981)	06
2 Combinações híbridas para dialelo circulante com $s = 3$, sendo $p = 6$ genitores.....	07
3 Esquema do dialelo completo envolvendo p pais e os híbridos (método II) do modelo de Griffing (1956)	09
4 Restrições nas soluções para o dialelo circulante, método 2 do modelo de Griffing (1956).....	12
5 Resumo das análises de variância e intervalos de confiança (LI e LS a 95%) para a herdabilidade, considerando as simulações do dialelo com e sem dominância, para diferentes herdabilidades e considerando duas repetições na avaliação das combinações híbridas.....	32
6 Resumo das análises de variância e intervalos de confiança (LI e LS a 95%) para a herdabilidade, considerando as simulações do dialelo com e sem dominância, para diferentes herdabilidades e considerando dez repetições na avaliação das combinações híbridas.....	33
7 Resumos das análises de variância dos dialelos e intervalos de confiança para a herdabilidade (LI e LI) com 95% de confiança, considerando os experimentos com dados reais.....	51

LISTA DE FIGURAS

Figura	Página
1 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman (r_c) entre as estimativas de capacidade geral de combinação (CGC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades, número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo com dominância.....	36
2 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman (r_c) entre as estimativas de capacidade geral de combinação (CGC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades, número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo com dominância.....	37
3 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman (r_c) entre as estimativas de capacidade geral de combinação (CGC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades, número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo sem dominância.....	38
4 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman (r_c) entre as estimativas de capacidade geral de combinação (CGC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades, número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de	

	cada caso, no modelo sem dominância.....	39
5	Proporção das estimativas da capacidade geral de combinação (CGC) obtidas no dialelo circulante, que coincidem com o intervalo de confiança para CGC do dialelo completo, considerando-se cinquenta simulações para diferentes valores de herdabilidade do carácter e número de cruzamentos de cada genitor (s), para o modelo com dominância.....	41
6	Proporção das estimativas da capacidade geral de combinação (CGC) obtidas no dialelo circulante, que coincidem com o intervalo de confiança para CGC do dialelo completo, considerando-se cinquenta simulações para diferentes valores de herdabilidade do carácter e número de cruzamentos de cada genitor (s), para o modelo com dominância.....	42
7	Proporção das estimativas da capacidade geral de combinação (CGC) obtidas no dialelo circulante, que coincidem com o intervalo de confiança para CGC do dialelo completo, considerando-se cinquenta simulações para diferentes valores de herdabilidade do carácter e número de cruzamentos de cada genitor (s), para o modelo sem dominância.....	43
8	Proporção das estimativas da capacidade geral de combinação (CGC) obtidas no dialelo circulante, que coincidem com o intervalo de confiança para CGC do dialelo completo, considerando-se cinquenta simulações para diferentes valores de herdabilidade do carácter e número de cruzamentos de cada genitor (s), para o modelo sem dominância.....	44
9	Estimativas das médias, mínimas e máximas das correlações	

	classificatórias de Spearman entre as estimativas de capacidade específica de combinação (CEC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades, número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo com dominância.....	47
10	Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman entre as estimativas de capacidade específica de combinação (CEC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades, número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo com dominância.....	48
11	Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman entre as estimativas de capacidade geral de combinação (CGC) no dialelo completo e no circulante, considerando-se experimentos: A - de Ferreira (1993) e B – Napolini Filho et al. (1981)	52
12	Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman entre as estimativas de capacidade específica de combinação (CEC) no dialelo completo e no circulante, considerando-se experimentos: A – de Ferreira (1993) e B – Napolini Filho et al. (1981)	54
13	Proporção das estimativas da capacidade geral de combinação (CGC) obtidas no dialelo circulante, que coincidem com o intervalo de confiança para CGC do dialelo completo, considerando cinquenta simulações para o número de cruzamentos de cada genitor (s), considerando-se experimentos: A - de Ferreira (1993) e B – Napolini Filho et al. (1981)	55

RESUMO

VEIGA, Ruben Delly. **Eficiência dos dialelos circulantes na escolha de genitores, avaliada com simulação de dados.** Lavras: UFLA, 1998. 100p. (Tese - Doutorado em Genética e Melhoramento de Plantas).

Os cruzamentos dialélicos consistem em um processo de escolha de populações segregantes, em programas de melhoramento genético, que pode ser inviabilizado se o número de genitores for elevado. Neste caso, os dialelos circulantes são uma alternativa, que, embora apresentem uma série de vantagens, é pouco utilizada, em razão da inexistência de informações sobre a sua eficiência em relação ao dialelo completo. O objetivo deste trabalho foi avaliar, por meio de simulação de dados, a eficiência de dialelos circulantes, na obtenção de estimativas de capacidade geral de combinação (CGC) e capacidade específica de combinação (CEC). Foram simuladas linhagens de 100 genes considerados independentes, com efeitos iguais, sem epistasia, que foram utilizadas para gerar dialelos completos, pelo método II do modelo de Griffing (1956). Considerou-se situações de herdabilidade de 10%, 20%, 50% e 75%, em modelos com efeitos aditivos e com dominância completa. A partir daí, foram simulados todos os dialelos circulantes possíveis em cada caso, pelo algoritmo proposto por Kempthorne e Curnow (1961). Foram utilizados também dados de experimentos disponíveis na literatura (Ferreira, 1993; Napolini Filho et al., 1981) para simulações de dialelos circulantes. Utilizou-se a linguagem do sistema SAS - *Statistical Analysis System* - (SAS, 1995) para: correlacionar, por meio da correlação de Spearman, as estimativas de CGC e CEC obtidas nos dialelos circulantes, cada um deles aleatorizado 50 vezes, com aquelas obtidas no dialelo completo; determinar as proporções de estimativas nos circulantes que ocorreram nos intervalos de confiança dos completos. Os resultados permitiram concluir que: o uso da simulação de dados por computador demonstrou eficiência e coerência com os resultados experimentais obtidos; os dialelos circulantes demonstraram ser tão eficientes quanto o completo, tanto na classificação dos pais quanto à CGC e à CEC, como na magnitude das estimativas desses parâmetros; no dialelo circulante, o número de cruzamentos (s) de cada pai afeta as estimativas de CGC e de CEC, embora, mesmo com um valor mínimo de s é possível obter uma boa concordância com as estimativas obtidas nos dialelos completos; especialmente em condições de baixa herdabilidade, pode ser vantajoso aumentar o número de cruzamentos de cada pai, embora a melhoria na eficiência ocorra com aumentos de s até no máximo a

* Comitê Orientador: Daniel Furtado Ferreira – UFLA, Magno Antônio Patto Ramalho – UFLA, César Augusto Brasil Pereira Pinto – UFLA e João Bosco dos Santos - UFLA

metade do número de pais envolvidos, com o que, já se consegue uma alta eficiência dos dialelos circulantes em todos os casos.

ABSTRACT

VEIGA, Ruben Delly. Efficiency of circulant diallels on parental choice, evaluated by data simulation. Lavras: UFLA, 1998. 95p. (Doctorate thesis in Genetics and Plant Breeding)*.

Diallel crosses consist in a process of choosing segregating populations in genetic improvement programs, which may become unfeasible if the number of parents is high. In this case, circulant diallels are an alternative that, although presenting a number of advantages, is little utilized due to the lack of information on their efficiency in relation to the complete diallel. The objective of this work was to evaluate by means of data simulation, the efficiency of circulant diallels, in obtaining estimates of general combining ability (GCA) and specific combining ability (SCA). Lines of 100 genes considered independent were simulated, with equal effects, without epistasis, which were used to generate complete diallels, by method II of Griffing's model (1956). Situations of heritability of 10%, 20%, 50% and 75% in models with additive effects and complete dominance were considered. From there on, all the possible circulant diallels were simulated in each case by the algorithm proposed by Kempthorne and Curnow (1961). Data of available experiments in the literature were also utilized (Ferreira, 1993; Napolini Filho et al., 1981) for simulations of circulant diallels. The language of the SAS system - Statistical Analysis System - (SAS, 1995) was employed to: correlate, by means of Spearman's correlation, the estimates of GCA and SCA obtained in circulant diallels each of randomized 50 times, with those obtained in the complete diallel; to determine the estimative ratios of estimates in the circulant ones that happened in the confidence interval of trust of the complete ones. The results allowed to end that: the use of the data simulation by computer showed both efficiency and consistency with the obtained experimental results; the circulant diallel demonstrated to be as efficient as the complete one on both in the parents' classification and concerning GCA and SCA and in the magnitude of the estimates of these parameters; in the circulant diallel, the number of each parents' crossings (s), it affects the estimates of GCA and of SCA although, even with a least value of s it is possible to obtain a good agreement with the estimates obtained in the complete diallel; mainly under conditions of low heritability it may be advantageous to raise the number of crosses of each parent, although the improvement in the efficiency takes place with increased s up to, in the maximum, the half of the

* Guidance Committee: Daniel Furtado Ferreira - UFLA (Major Professor), Magno Antônio Patto Ramalho - UFLA, César Augusto Brasil Pereira Pinto - UFLA and João Bosco dos Santos.

number of parents concerned, this way, a high efficiency of the circulant diallels in every case already may be obtained.

1 INTRODUÇÃO

Nas espécie vegetais existe, em geral, uma quantidade inumerável de linhagens e cultivares que podem participar dos programas de hibridação para a obtenção das populações segregantes, fazendo-se necessária a utilização de critérios de escolha. A identificação de metodologias eficientes para essa escolha é a área de melhoramento que provavelmente tem recebido a maior atenção dos pesquisadores (Baenziger e Peterson, 1991; Bernardo, 1996; Charcosset, A. et al., 1998; Souza, 1997; Abreu, 1997).

Entre as metodologias disponíveis, os cruzamentos dialélicos são os mais amplamente utilizados em quase todas as espécies cultivadas. Uma das vantagens dessa metodologia é que ela não só contribui para o processo decisório de escolha de pais e populações segregantes, como também pode fornecer informações sobre o controle genético dos caracteres.

A principal restrição aos cruzamentos dialélicos é que, com o aumento do número de pais, o número de combinações híbridas a serem obtidas e avaliadas torna-se proibitivo. Uma opção é o emprego dos cruzamentos dialélicos parciais que separam os pais em dois grupos, obtendo-se cruzamentos apenas entre os grupos, reduzindo assim o número de híbridos a serem obtidos (Hallauer e Miranda Filho, 1981).

Uma outra alternativa que pode dar maior flexibilidade aos cruzamentos dialélicos, com relação ao número de pais envolvidos, são os dialelos circulantes. Este tipo de dialelo, embora tenha sido proposto há algumas décadas (Kempthorne e Curnow, 1961) e possibilite o envolvimento de um número

muito superior de pais, tem sido pouco utilizado. Isso ocorre porque os melhoristas não têm informações sobre a sua eficiência em relação aos dialelos completos.

Desta forma, foi realizado o presente trabalho, com o objetivo de avaliar a eficiência de dialelos circulantes em relação ao dialelo completo, quanto às estimativas da capacidade geral de combinação (CGC) e da capacidade específica de combinação (CEC), a partir de simulação de dados em computador, tomando todos os circulantes possíveis de um mesmo dialelo completo, considerando diferentes níveis de herdabilidade, na presença e ausência de dominância. O estudo de avaliação da eficiência dos dialelos circulantes também foi complementado a partir de dados experimentais disponíveis na literatura, com simulação de circulantes possíveis em cada caso.

2 REFERENCIAL TEÓRICO

2.1 Escolha de populações segregantes

Tanto em programas de melhoramento de plantas autóгамas como de alógamas, a escolha da população segregante na qual proceder a seleção é a etapa crucial. Isto porque uma escolha mal feita compromete todas as etapas posteriores que demandam grande quantidade de recursos e de tempo (Hallauer, 1990). Essa escolha não é uma atividade simples pois, dado um número de cultivares ou populações disponíveis, as possibilidades de combinações entre elas, duas a duas são infinitas.

Assim, há a necessidade de metodologias eficientes de escolha. Em função desse fato, essa é provavelmente a área, nos programas de melhoramento, que tem recebido maior atenção. Segundo Baenzinger e Peterson (1991), os procedimentos de escolha das populações podem ser agrupados em duas categorias: a primeira abrange aqueles que utilizam o comportamento dos genitores *per se* e, a segunda, o comportamento dos genitores pela sua descendência.

Na primeira categoria estão alguns procedimentos, sendo que o mais comum deles é a própria média dos possíveis genitores. Há, contudo, inúmeras outras opções que utilizam, por exemplo, técnicas multivariadas (Ferreira, 1993; Morrison, 1967), o BLUP (Bernardo, 1996) e marcadores moleculares (Duarte, 1998).

Na segunda categoria estão os procedimentos que baseiam a decisão no desempenho dos genitores em cruzamentos. Neste grupo, os cruzamentos dialélicos são mais comuns.

2.2 Cruzamentos dialélicos

Cruzamentos dialélicos, ou simplesmente dialelos, podem ser conceituados como sendo as combinações de p genitores, dois a dois. Para se emitir qualquer comentário sobre dialelos é importante também, inicialmente, conceituar capacidade geral (CGC) e específica de combinação (CEC), termos que foram utilizados inicialmente por Sprague e Tatum (1942). A capacidade geral expressa o desempenho médio do genitor i em combinações com os p genitores envolvidos e a capacidade específica, refere-se ao desvio ocorrido em cada combinação, em relação ao esperado, considerando as CGC de cada genitor.

Há várias opções de cruzamentos dialélicos. A primeira consiste nos cruzamentos dialélicos completos ou balanceados, em que todos os cruzamentos dos p genitores, dois a dois, estão presentes, sendo obtidas $p(p-1)/2$ combinações híbridas, com ou sem os recíprocos. Desse modo, as avaliações implicam em p^2 combinações, se forem envolvidos as combinações híbridas, seus recíprocos e os genitores. É evidente que há outras possibilidades, que serão realçadas posteriormente (Cruz e Regazzi, 1994).

A segunda refere-se aos dialelos parciais completos. Nesse caso, os genitores envolvidos estão separados em dois grupos, contendo cada qual p_1 e p_2 , sendo que $p_1 + p_2 = p$. Os pares de combinações híbridas são obtidos apenas entre genitores de grupos diferentes, perfazendo $p_1.p_2$ combinações a serem avaliadas.

Os dialelos completos ou parciais são denominados incompletos quando há falta de uma ou mais combinações híbridas. A única diferença em relação aos

outros anteriormente comentados é que não há possibilidade de utilização das expressões ordinárias desenvolvidas para análise ou os *softwares* específicos para tal. Porém, eles podem ser analisados empregando-se o método dos quadrados mínimos (Cruz e Regazzi, 1994; Veiga e Cavalcanti, 1997).

A terceira opção são os cruzamentos dialélicos circulantes, propostos por Kempthorne e Curnow (1961). Nesse caso, os p pais estão presentes em um número s de combinações híbridas, que deve estar entre 2 e $p-1$ genitores, sendo os pais também avaliados. Se por ventura não houver possibilidade de incluir os pais, o número s de combinações híbridas deve estar entre 3 e $p-1$.

Alguns autores fizeram comparações entre os modelos disponíveis de dialelos (Cruz e Vencovsky, 1989), ou propuseram adaptações de modelos existentes para descrever situações particulares (Geraldí e Miranda Filho, 1988).

A principal diferença entre essas opções de cruzamentos dialélicos, como já mencionado, é o número de combinações híbridas obtidas e avaliadas. Hallauer e Miranda Filho (1981) traçaram comentários a esse respeito, comparando os dialelos completos e parciais. A Tabela 1 mostra o número de combinações considerando a presença ou não dos recíprocos e os genitores, de acordo com estes autores. Fica claro que, com o aumento de genitores, o número de combinações híbridas torna-se impraticável. Uma primeira opção para contornar esta dificuldade é a de usar os dialelos parciais.

Se forem considerados, por exemplo, um número de 100 genitores envolvidos, no dialelo completo deveriam estar presentes, sem os recíprocos, 4950 híbridos. Este grupo de 100 genitores poderia ser subdividido em dois outros, de acordo com algum critério genético ou agrônômico. Dependendo do número de genitores em cada grupo, o número de combinações híbridas seria diferente. O maior número ocorre quando a divisão é feita pela metade; para 50 genitores, em cada grupo do exemplo citado, seriam necessárias 2500 combinações híbridas, ou seja, praticamente a metade do dialelo completo. Se a

separação não for proporcional, (por exemplo, 30 genitores no grupo 1 e 70 no grupo 2), o número de combinações passa a ser de 2100. É fácil visualizar que, quanto mais desproporcional a divisão, menor será o número de combinações híbridas a serem obtidas e testadas. É por esta razão que esse tipo de dialelo tem sido utilizado em várias espécies cultivadas (Takeda, 1990).

TABELA 1 Número de cruzamentos dialélicos em função do número de pais , na presença e na ausência dos recíprocos (Hallauer e Miranda Filho, 1981)

Pais	Sem os recíprocos		Com os recíprocos	
	Híbridos	Total	Híbridos	Total
5	10	15	20	25
6	15	21	30	36
7	21	28	42	49
8	28	37	56	64
9	36	45	72	81
10	45	55	90	100
15	105	120	210	225
20	190	210	380	400
50	1225	1275	2450	2500
100	4950	5050	9900	10000
...
p	$p(p-1)/2$	$p(p+1)/2$	$p(p-1)$	p^2

Neste aspecto, o dialelo circulante é ainda mais vantajoso que o dialelo parcial, razão pela qual foi sugerido nos casos em que o melhorista tem um grande número de genitores a serem avaliados (Gonçalves, 1987 e Cruz e Regazzi, 1994).

O algoritmo para determinar o número e identificar as combinações híbridas foi proposto por Kempthorne e Curnow (1961) e está descrito em Cruz

e Regazzi (1994). Segundo estes autores, com p genitores e s cruzamentos por genitor ($2 \leq s \leq p - 1$), o número total de combinações é $p.s$, porém como os recíprocos não necessitam ser incluídos, o número de combinações híbridas total passa a ser $ps/2$.

Para se identificar quais cruzamentos devem ser realizados (y_{ij}), é necessário salientar que $i = 1, 2, \dots, p$, $j = k+i, k+i+1, \dots, k+i-1+s$, sendo k obtido pela expressão $k = (p+1-s)/2$, com a restrição evidente de que k deve ser um número inteiro. Depreende-se, então, que s não pode assumir qualquer valor, para determinado valor de p . Ou seja, se p for um número par, o s deve ser ímpar. O contrário também é verdadeiro. Além disso, se os pais estiverem incluídos no dialelo circulante, o número s pode ser igual a 2, caso contrário, s deve iniciar em 3. Valores de j acima de p devem ser reduzidos por múltiplos de p , para que j esteja entre i e p . Outra observação refere-se ao limite máximo do valor de s , que, se for igual a $p-1$, o dialelo circulante corresponderá ao completo.

Seja um exemplo em que $p = 6$, isto é, 6 genitores envolvidos. Neste caso s pode ser 3 ou 5, pois p é par. Considerando $s = 3$, tem-se que $k = (p+1-s)/2 = (6+1-3)/2 = 2$. Assim, as combinações híbridas a serem obtidas serão em número de $p.s/2 = 6.3/2 = 9$ (Tabela 2).

TABELA 2 Combinações híbridas para um dialelo circulante com $s = 3$, $p = 6$ genitores.

Pai 1	Pai 2					
	1	2	3	4	5	6
1	$y_{1,1}$		$y_{1,3}$	$y_{1,4}$	$y_{1,5}$	
2		$y_{2,2}$		$y_{2,4}$	$y_{2,5}$	$y_{2,6}$
3			$y_{3,3}$		$y_{3,5}$	$y_{3,6}$
4				$y_{4,4}$		$y_{4,6}$
5					$y_{5,5}$	
6						$y_{6,6}$

Voltando ao exemplo com 100 genitores, s pode assumir qualquer valor ímpar entre 3 e 99. Novamente considerando-se o menor valor, o número de combinações híbridas totais será de 150, número bem inferior aos 4950 necessários nos cruzamentos dialélicos completos. Infelizmente, não são disponíveis informações precisas a respeito do valor mínimo ou ideal de s para que os cruzamentos sejam realizados sem prejuízo nas inferências genéticas a serem obtidas a partir do dialelo circulante.

2.3 Análise dos cruzamentos dialélicos

Segundo Vencovsky (1970), evolução ocorrida nos objetivos do melhoramento foi responsável por modificações e aprimoramento das bases teóricas e das técnicas para pesquisas com caracteres quantitativos. No início, as pesquisas foram orientadas para estimação de componentes de variância genética, geralmente funções lineares de quadrados médios. Segundo o autor, surgiram alguns inconvenientes, levando à necessidade de experimentos de larga escala para contorná-los. Em consequência, foram desenvolvidas as bases teóricas que permitiram estudar propriedades genéticas de populações a partir de médias. Pesquisas dentro desta linha têm mostrado utilidade no estudo de populações, cruzamentos e populações afins, quanto a aspectos ligados a efeitos epistáticos na expressão das médias, bem como da variabilidade genética.

Quando os cruzamentos dialélicos são tais que os pais representam a população, caracteriza-se o modelo fixo. Neste caso, o objetivo não é a estimação de componentes de variância, e, sim, a estimação de efeitos de cruzamentos específicos e do efeito de todos os cruzamentos que envolvem um pai comum.

Dentre os modelos propostos para análise de tabelas dialélicas, os mais utilizados correspondem aos métodos de Jinks e Hayman (1953), Griffing (1956) e Gardner e Eberhart (1966). O modelo de Jinks e Hayman se destaca

pela quantidade de informações genéticas em espécies estudadas (Vencovsky, 1970), porém com a imposição de uma série de restrições para sua validade. O método de Gardner e Eberhart (1966) tem como principal vantagem um estudo detalhado da heterose, com sua decomposição em vários componentes.

Griffing (1956) propôs quatro métodos de análise dialélica, sendo que o denominado método II é um dos mais utilizados, e caracteriza-se por incluir os híbridos e os genitores, sem os recíprocos. A Tabela 3 apresenta um esquema de cruzamentos dialélicos para o tal método. Maiores detalhes para os demais métodos do modelo de Griffing e demais modelos citados podem ser encontrados em Cruz e Regazzi (1994).

TABELA 3 Esquema do dialelo completo envolvendo p pais e os híbridos (método II) do modelo de Griffing (1956).

Pai1	Pai2						p
	1	2	3	4	5	...	
1	$y_{1,1}$	$y_{1,2}$	$y_{1,3}$	$y_{1,4}$	$y_{1,5}$...	$y_{1,p}$
2		$y_{2,2}$	$y_{2,3}$	$y_{2,4}$	$y_{2,5}$...	$y_{2,p}$
3			$y_{3,3}$	$y_{3,4}$	$y_{3,5}$...	$y_{3,p}$
4				$y_{4,4}$	$y_{4,5}$...	$y_{4,p}$
5					$y_{5,5}$...	$y_{5,p}$
...						...	
p							$y_{p,p}$

Neste caso, o modelo estatístico é expresso por:

$$y_{ij} = m + g_i + g_j + s_{ij} + \bar{e}_{ij} ; \quad (i) \quad (1)$$

em que:

y_{ij} : valor médio do híbrido ($i \neq j$), ou do genitor ($i=j$);

$i, j = 1, 2, \dots, p$ (número de pais) ;

m : média geral ;

g_i, g_j : efeitos da capacidade geral de combinação do i -ésimo e j -ésimo

- país ;
- s_{ij} : efeitos da capacidade específica de combinação no cruzamento entre o i -ésimo e j - ésimo país ;
- $\bar{\epsilon}_{ij}$: erro experimental médio.

O modelo estatístico pode ser também representado, em notação matricial (Searle, 1971 e Cruz e Regazzi, 1994), da seguinte maneira:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{e} \quad ; \quad (\text{ii})$$

em que :

- \mathbf{Y} : vetor composto pelas médias obtidas pelos híbridos e pelos genitores, com dimensões $p(p+1)/2 \times 1$;
- \mathbf{X} : matriz de delineamento do modelo, com $p(p-1)/2$ linhas e $1+p+p(p+1)/2$ colunas;
- \mathbf{B} : vetor de parâmetros, com dimensões $1+p+p(p+1)/2 \times 1$;
- \mathbf{e} : vetor composto pelas médias dos erros experimentais dos cruzamentos, com dimensões $1+p+p(p+1)/2 \times 1$.

A análise de variância decorrente do modelo descrito desdobra o efeito de tratamentos em efeitos de capacidade geral e específica de combinação. As estimativas desses parâmetros, representadas pelo vetor \mathbf{b} são obtidas pela solução das equações normais descritas em (iii), com restrições impostas às soluções (Cruz e Regazzi, 1994):

$$\mathbf{X}'\mathbf{X} \mathbf{b} = \mathbf{X}'\mathbf{Y} \quad ; \quad (\text{iii})$$

2.4 O Emprego de dialelos circulantes

O estudo da capacidade combinatória de um grupo maior de pais, conforme já salientado neste capítulo, requer a obtenção e a avaliação de várias combinações híbridas, o que pode inviabilizar os estudos (Cruz e Regazzi, 1994). Este problema pode ser contornado com o uso de dialelos circulantes, apesar de que, com o número de cruzamentos reduzido, os estimadores das funções estimáveis ficam sujeitos a uma elevação de suas variâncias como consequência da elevação da variância residual. Estes delineamentos consistem em uma amostragem dos cruzamentos possíveis entre os pais em estudo, tratando-se de metodologia altamente eficiente sob o ponto de vista prático, permitindo estimar a capacidade de combinação de linhagens em cruzamento, ao mesmo tempo podendo-se detectar combinações híbridas promissoras (Gonçalves, 1987). Este autor propôs um esquema de dialelo circulante para avaliar 25 linhagens endógamas, cruzadas aos pares, de modo que cada uma pudesse aparecer em três cruzamentos de forma circulante. As conclusões obtidas puderam mostrar que o modelo reduzido foi eficiente na avaliação do potencial genético de linhagens endógamas. A primeira contribuição neste tipo de delineamento, apresentando método de análise, refere-se a Gilbert (1958) o qual sugeriu que, na análise de um dialelo parcial de 15 a 20 genitores, cada um deveria estar presente em 8 híbridos.

O modelo estatístico para o dialelo circulante, considerando o modelo de Griffing (1956) descrito por Cruz e Regazzi (1994) é definido como:

$$y_{ij} = m + g_i + g_j + s_{ij} + \bar{e}_{ij} , \quad (\text{iv})$$

em que :

y_{ij} : valor médio da combinação híbrida ($i \neq j$), ou do genitor ($i=j$);

$i = 1, 2, \dots, p$;

$j = i+k, i+k+1, \dots, i+k+s-1$; $k = (p+1-s)/2$;

- m** : média geral ;
- g_i, g_j** : efeitos da capacidade geral de combinação do i-ésimo e j-ésimo pais ;
- s_{ij}** : efeitos da capacidade específica de combinação entre o i-ésimo e j-ésimo pais ;
- \bar{e}_{ij}** : erro experimental médio, admitido homocedástico e com distribuição normal $N(0, \sigma^2/r)$.

O modelo pode também ser representado seguindo a notação matricial descrita em (ii), observadas as dimensões das matrizes envolvidas. As estimativas dos parâmetros do modelo são obtidas pela solução das equações normais dadas por (iii). Como acontece com o dialelo completo, a matriz de delineamento não possui inversa única, sendo necessário a imposição de restrições nas soluções, para que possa ser obtida solução única para os parâmetros genéticos, capacidade geral de genitores e capacidade específica entre pares de genitores, de interesse dos melhoristas. No caso dos circulantes, no método II do modelo de Griffing, as restrições impostas são descritas conforme a Tabela 4.

TABELA 4 Restrições nas soluções para o dialelo circulante, método II do modelo de Griffing (1956)

Estimadores	Número de Estimativas	Número de restrições	Tipos de restrições	Grans de liberdade
\hat{m}	1	-	-	1
\hat{g}_i	p	1	$\sum_i g_i = 0$	p-1
\hat{s}_i	ps/2 + p	p	$2s_{ii} + \sum_{j(i)} s_{ij} = 0$	ps/2
Total	ps/2 + 2p + 1	(p+1)	-	ps/2 + 2

Deve ser observado que, em função da variação dos índices que controlam os pais no modelo (v) ($i = 1, 2, \dots, p$; $j = i+k, i+k+1, \dots, i+k+s-1$; $k = (p+1-s)/2$), tem-se em relação aos parâmetros do modelo: $\sum_i g_i = 0$ e $\sum_{j(i)} g_j \neq 0$.

Além da obtenção direta das estimativas para os parâmetros do modelo, utilizando a notação matricial, Cruz e Regazzi (1994) utilizam uma descrição proposta por Kempthorne e Curnow (1961) que simplifica as análises em geral, principalmente no caso de desenvolvimento de programas computacionais, como foi o caso deste trabalho. Os autores se baseiam no fato de que, no método 2 do modelo de Griffing (1956), as equações normais podem ser expressas por:

$$Y_{..} = \left[\frac{ps}{2} + p \right] \hat{m} + (s+2) \sum_i \hat{g}_i + \frac{1}{2} \sum_i (2\hat{s}_{ii} + \sum_{j(i)} \hat{s}_{ij}) \quad (v)$$

$$2Y_{ii} + \sum_{j(i) \neq i} Y_{ij} = (s+2)\hat{m} + (s+4)\hat{g}_i + \sum_{j(i)} \hat{g}_j + 2\hat{s}_{ii} + \sum_{j(i)} \hat{s}_{ij} \quad (vi)$$

Tem-se ainda que:

$$Y_{ij} = \hat{m} + \hat{g}_i + \hat{g}_j + \hat{s}_{ij} \quad (vii)$$

Considerando-se as restrições, obtém-se os estimadores:

$$\hat{m} = \frac{Y_{..}}{p(s+2)} \quad e \quad \hat{s}_{ij} = y_{ij} - \hat{m} - \hat{g}_i - \hat{g}_j$$

As estimativas de \hat{g}_i são obtidas pela solução do sistema:

$$A\hat{G} = Q ;$$

em que:

\hat{G} : vetor de dimensões $p \times 1$ dos estimadores da capacidade geral de estimação;

Q : vetor de dimensões $p \times 1$ cujos elementos são dados por:

$$2y_{ii} + \sum_{j(i) \neq i} y_{ij} - (s+2)\hat{m} ;$$

A : matriz de dimensões $p \times p$ cujos elementos são:

$$a_{ii} = s+4;$$

$$a_{ij} = 1, \text{ se } j = k+1, k+i+1, \dots, k+i+s-1 \text{ (} 1 \leq j \leq p \text{)};$$

$$a_{ij} = 0, \text{ nos outros casos.}$$

Algumas considerações têm sido feitas em relação à eficiência dos dialelos circulantes. Suas desvantagens se devem à perda de informações para os híbridos ausentes nos cruzamentos e ainda pelas estimativas que são questionáveis para a avaliação da capacidade combinatória quando o número de cruzamentos é pequeno. Além disto, os poucos graus de liberdade associados aos efeitos da capacidade específica de combinação podem trazer dificuldades de detectar diferenças significativas nos ensaios realizados (Cruz e Regazzi, 1994 e Hallauer e Miranda Filho, 1981).

Vários trabalhos comparam delineamentos diversos com os parciais, embora nenhum deles aponte uma solução quanto ao número de cruzamentos para cada pai que seja eficiente na determinação de estimativas de parâmetros genéticos. Murty, Arunachalam e Anand (1967) compararam estimativas de capacidade de combinação e de variâncias obtidas em dialelo completo ($p = 10$) com as obtidas em dialelos circulantes ($s = 3, 5, 7$), seguindo o modelo de Kempthorne e Curnow (1961). Os autores, concluem que as estimativas tendem a ser subestimadas para valores de s inferiores a $p/2$. Murty e Anan (1966), citados por Gonçalves (1987), também compararam dialelos parciais com o

completo e mostraram que o parcial foi melhor do que o completo quando o número de pais no parcial é igual ou superior à metade do total.

Hallauer e Miranda Filho (1981), comentam que Kempthorne e Curnow (1961) discutem a eficiência dos dialelos circulantes na estimativa de componentes genéticos, quanto ao número de cruzamentos de cada genitor, e concluem que as estimativas de componentes de variância e de herdabilidade devem ser similares às obtidas nos dialelos completos.

De maneira geral, no caso de dialelos parciais, procura-se reduzir o número de cruzamentos (s), procurando fazer com que as estimativas dos componentes de variância estejam racionalmente próximas das análises do dialelo completo (Gonçalves, 1987). Além destas observações, do ponto de vista genérico, não se tem conhecimento, em literatura, de estudos de comparação entre dialelos completos e os possíveis dialelos circulantes que poderiam ser realizados para um mesmo conjunto de pais.

2.5 Uso da simulação de dados

Até o século XVII, o conhecimento era fortemente influenciado pelos postulados de filósofos como Platão, Aristóteles ou Euclides, com supostos poderes de previsão, algumas vezes amparados por métodos dedutivos. Com o surgimento da filosofia científica, ou simplesmente o método científico, as atenções se voltaram para os fatos: observação do sistema físico, formulação de uma hipótese, soluções para o modelo matemático e realização de experiência para testar as hipóteses. Desta forma, a autoria da metodologia científica é atribuída, dentre outros, a Sir Francis Bacon, que, no ano de 1620, escreveu “A razão sozinha não tem nenhuma capacidade de previsão; ela o consegue somente em combinação com a observação” (Naylor, 1971).

Os primeiros indícios de simulação de dados surgiram com a utilização do método de Monte Carlo, através de Von Neuman, em 1940, com blindagem

de reatores nucleares (Morgan, 1995), e os primeiros trabalhos utilizando simulação de dados, na área de Genética são de autoria de Fraser (1957a) e Fraser (1957b).

Vários autores definiram simulação de dados. Entre eles destacam-se algumas definições: “imitar, através de programa de computador, comportamento de um sistema real, para estudar seu funcionamento em condições alternativas” (Dachs, 1988); “variedade de técnicas úteis, conectadas com as regras de um modelo de alguma espécie” (Morgan, 1995); “técnica numérica para realizar experiências em um computador digital, as quais envolvem certos tipos de modelos lógicos que descrevem o comportamento de um sistema” (Naylor, 1971). Esta última definição torna-se mais interessante no que diz respeito à simulação de dados aplicada a modelos biológicos ou a pesquisadores de modo geral, uma vez que enfatiza a necessidade do envolvimento de modelos nos estudos de simulação.

O uso da simulação de dados tem uma grande diversidade de áreas de aplicação, basicamente sob duas linhas de atuação: problemas matemáticos completamente determinísticos, cuja solução é difícil, ou em problemas que envolvem o processo estocástico (Monte Carlo), cuja técnica de simulação tem base probabilística ou estocástica (Hammersley, 1964). Exemplo de aplicação de processo puramente determinístico pode ser visto no estudo de Gosset, em 1908, a respeito da distribuição amostral da variável definida como t de Student para tamanhos diferentes de amostras.

A simulação de dados com base em processos probabilísticos tem enorme aplicação na definição de modelos matemáticos, certamente com base em processos de geração de números aleatórios, em que a geração de distribuições de probabilidade são garantidas pelo teorema de probabilidade integral: “Se U é uma variável aleatória com distribuição no intervalo $(0,1)$ e se

F é uma função de distribuição qualquer, então a função F inversa de U tem distribuição desejada” (Morgan, 1995).

Esta teoria respalda e implementa as funções geradoras de números aleatórios para todas as distribuições conhecidas, disponíveis em programas computacionais, a exemplo do sistema SAS (SAS, 1995).

Em Genética e Melhoramento de Plantas, muitos estudos por simulação de dados têm sido utilizados. Brown e Rothery (1993) utilizaram simulação por computador e exemplos reais em biologia para demonstrar propriedades e métodos de ajuste de modelos. Sua aplicação se estende a formas determinísticas e estocásticas para modelos representativos de populações e processos, com variações amostrais, ambientais e genéticas. Boote, Jones e Pckering (1996) descreveram a potencialidade e as limitações do uso de modelos em produção, bem como das entradas necessárias à construção de modelos. Os autores apontam as possibilidades de incluir nos modelos aspectos com a interação de genética, fisiologia e ambiente. As limitações apontadas estiveram ligadas à escolha adequada dos modelos, quanto ao nível de complexidade a ser adotado em cada situação.

Vários trabalhos utilizam simulação de dados, enfocando aspectos de seleção. Partner et al. (1993) trabalharam com seleção em espécies produtivas de forma hipotética, para avaliar efeito da seleção em caracteres genéticos quantitativos. Conforme o esquema utilizado, linhagens são cruzadas obtendo-se geração F_1 e, posteriormente aplica-se seleção truncada em gerações subsequentes até F_9 . A simulação apresentada permitiu o uso de diferentes métodos de melhoramento, em diferentes tipos de interação alélica. Villanueva e Kennedy (1993) comparam a técnica de índice de seleção com seleção em tandem, obtendo expressões algébricas para respostas à seleção por índices. Os autores concluíram que, embora os índices de seleção sejam mais eficientes quanto aos valores agregados pelo melhoramento, sua eficiência relativa diminui

após repetidos ciclos de seleção, e a redução na eficiência aumenta com maior intensidade de seleção e herdabilidade.

Outros autores têm introduzido, também, questões ligadas à seleção recorrente nos trabalhos de simulação. Kervella, Robert e Fouilloux (1993) estudaram a influência da recombinação na variabilidade genética. Através da simulação de programas de seleção recorrente com frequências de recombinação alta, média e baixa. Concluíram que, em taxas médias e altas de recombinação, as mudanças na média e variância genética são quase idênticas, enquanto que em frequências baixas de recombinação a variância genética tende a ficar mais alta após o segundo ciclo de seleção. Mather e Kannenberg (1992) utilizaram simulação para investigar o efeito da correlação genética causado por ligação ou pleiotropia, em resposta a repetidos ciclos de seleção em milho. A correlação genética entre produção de grãos e maturidade influenciou a resposta indireta da maturidade quando a seleção foi feita apenas para produção. Esses autores destacaram ainda que o conhecimento da associação entre correlação genética entre características e o desequilíbrio de ligação é útil na previsão de respostas à seleção.

Seleção assistida por marcadores tem sido objeto de estudos por simulação de dados em computador. Edwards e Page (1994) avaliaram as respostas à seleção assistida por marcadores e compararam-nas com as de seleção recorrente fenotípica em espécies alógamas anuais. Os resultados evidenciaram que a seleção assistida demonstrou eficiência entre 44,7 e 99,5% em relação à seleção recorrente fenotípica, dependendo do modelo genético considerado. Moreno (1992), utilizou simulação de Monte Carlo e regressão múltipla para verificar se o uso de QTL's no melhoramento de plantas deve estar associado à magnitude de efeitos aditivos e dominantes no modelo. Ferreira (1995) utilizou os recursos de programação em simulação de dados para avaliar a eficiência de métodos de mapeamento de locos quantitativos (QTLs) e da seleção assistida

por marcadores moleculares. Outros autores trabalharam na linha de simulação de seleção assistida por marcadores, incluindo outros fatores além dos citados, a exemplo de Bearzoti (1997), Zhang e Smith (1993) e Euclides (1996). Este último desenvolveu e utilizou o sistema computacional GENESYS para avaliação de métodos de seleção clássicos e associados a marcadores moleculares.

Certamente, o uso de simulação de dados tem sido uma ferramenta bastante útil, com grande amplitude de aplicação. Em todos os casos, a simulação propiciou obter informações e estudar fenômenos de forma adequada que, algebricamente, são extremamente complexos.

Vários autores têm recorrido à simulação de dados para estudos relacionados com cruzamentos dialélicos, seja para comparações de métodos e análises dialélicas ou para obtenção de parâmetros genéticos. Assim, Cruz e Vencovsky (1989) usaram a simulação para comparar alguns métodos da metodologia proposta por Griffing (1956) e da metodologia proposta por Gardner e Eberhart (1966). Os autores puderam descrever os parâmetros genéticos do modelo em termos de efeitos e de frequências alélicas, e apontaram a interpretação genética mais adequada para os contrastes envolvendo os parâmetros do modelo.

Oliveira (1995) apresentou alternativas na escolha de pais em um programa de melhoramento do feijoeiro, utilizando simulação de dados com a finalidade de avaliar a estimação de parâmetros genéticos e o desempenho de populações segregantes na geração F_{∞} . Nesta linha, Zhu e Weir (1996) utilizaram simulação de Monte Carlo para comparar métodos de estimação de parâmetros genéticos, incluindo efeito materno e efeito ligado ao sexo.

Na área de desenvolvimento de software, são muitas as alternativas de sistemas computacionais, que, em geral possuem alternativas de simulação de cruzamentos dialélicos e de análises estatísticas relacionadas. Dentre eles, citam-

se: DIALLEL (Burow e Coors, 1994); PAGIS (Oliveira et al., 1994); MENDEL (Schwarzbach et al., 1991); GREGOR (Tinker e Mather, 1993) e GENES (Cruz, 1997).

3 METODOLOGIA

Este trabalho foi realizado em duas etapas. Na primeira foram simulados dialelos com 50 pais e na segunda foram utilizados dados disponíveis na literatura.

3.1 Dialelos simulados

Inicialmente considerou-se um dialelo completo com 50 pais. Para isto foram simuladas 50 linhagens contendo 100 genes, de efeitos iguais e distribuição independente. Na simulação consideraram-se dois tipos de interação alélica: aditivo e de dominância completa, desconsiderando-se a ocorrência de interações epistáticas. Sendo assim, o valor genotípico atribuído a cada linhagem foi obtido pelo somatório dos efeitos genéticos de cada loco. Os valores genotípicos das combinações híbridas envolvendo duas linhagens foram obtidos de modo semelhante ao relatado para as linhagens genitoras.

Na constituição genotípica das linhagens, procurou-se aleatorizar a distribuição dos locos, de modo que fossem obtidas linhagens com, no mínimo, dois locos com alelos favoráveis, até o total de 100. Desse modo, obteve-se uma população de linhagens com uma ampla variabilidade genética.

O valores fenotípicos das linhagens e dos híbridos da geração F_1 foram estabelecidos atribuindo-se erros aleatórios aos valores genotípicos. Esses erros

foram atribuídos conforme diferentes herdabilidades, no sentido amplo, considerando a seleção entre pais e híbridos (h^2), isto é, 10%, 20%, 50% e 75%.

A expressão de h^2 nesse caso é fornecida por:

$$h^2 = \frac{\sigma_G^2}{\left(\sigma_G^2 + \frac{\sigma_e^2}{r}\right)} \quad ;$$

em que:

σ_G^2 é a variância genética obtida a partir dos efeitos genéticos das linhagens e híbridos;

r é o número de repetições em que o dialelo seria avaliado experimentalmente; no caso foram consideradas duas situações, isto é, $r = 2$ e $r = 10$;

σ_e^2 é a variância do erro; no presente contexto, como a h^2 foi pré-fixada e a variância genética é conhecida, obteve-se σ_e^2 pela expressão :

$$\sigma_e^2 = \left(\frac{1-h^2}{h^2}\right) \cdot r \cdot \sigma_G^2$$

Definindo-se a variância residual e considerando-se que os erros apresentam distribuição normal, com média zero, os efeitos dos erros foram gerados pela função RANNOR do sistema SAS – *Statistical Analysis System* (SAS, 1995).

Uma vez obtidos os valores fenotípicos das linhagens e dos híbridos nas repetições (duas ou dez), efetuou-se a análise de variância e a partir dos resultados médios procedeu-se a análise dialélica adotando-se o procedimento de Griffing (1956), método II, modelo I, apresentado a seguir:

$$y_{ij} = m + g_i + g_j + s_{ij} + \bar{e}_{ij} \quad ;$$

em que :

- y_{ij} : valor médio da combinação híbrida ($i \neq j$) ou do genitor ($i = j$);
 $i, j = 1, 2, \dots, p$ (número de genitores) ;
- m : média geral ;
- g_i, g_j : efeitos da capacidade geral de combinação do i -ésimo e j -ésimo genitores ;
- s_{ij} : efeitos da capacidade específica de combinação entre o i -ésimo e j -ésimo genitores ;
- \bar{e}_{ij} : erro experimental médio.

O modelo estatístico pôde ser também representado em notação matricial (Searle, 1971 e Cruz e Regazzi, 1994), da seguinte maneira:

$$Y = XB + e \quad ;$$

em que :

- Y : vetor composto pelas médias obtidas pelos híbridos e pelos genitores, com dimensões $p(p+1)/2 \times 1$;
- X : matriz de delineamento do modelo, com $p(p-1)/2$ linhas e $1+p+p(p+1)/2$ colunas;
- B : vetor de parâmetros, com dimensões $1+p+p(p+1)/2$;
- e : vetor composto pelas médias dos erros experimentais dos cruzamentos com dimensões $1+p+p(p+1)/2 \times 1$.

A partir da análise de variância foi estimada a herdabilidade no sentido amplo para a seleção entre tratamentos avaliados, visando compará-la com o valor paramétrico.

O intervalo de confiança (IC) de h^2 foi obtido pela expressão de Knapp et al. (1985), descrito em Souza (1997), conforme expressões abaixo:

$$\hat{h}^2 = \frac{QMTrat - QMErro}{QMTrat}$$

em que:

\hat{h}^2 : estimativa da herdabilidade ao nível de médias de tratamentos;
 QMTrat : quadrado médio de tratamentos da análise dialélica (Cruz, 1997);

QMErro : quadrado médio do resíduo da análise de variância correspondente aos valores de parcela.

$$LI = \left\{ 1 - \left[\left(\frac{QMTrat}{QMErro} \right) \cdot F_{(1-\alpha/2; gl_{Res}, gl_{Trat})} \right]^{-1} \right\} ;$$

$$LS = \left\{ 1 - \left[\left(\frac{QMTrat}{QMErro} \right) \cdot F_{(\alpha/2; gl_{Res}, gl_{Trat})} \right]^{-1} \right\}$$

em que:

LI : limite inferior do IC para herdabilidade ;

LS : limite superior do IC para herdabilidade ;

$F_{(1-\alpha/2; gl_{Res}, gl_{Trat})}$: valor do percentil da distribuição F de Snedecor, com graus de liberdade gl_{Res} e gl_{Trat} (resíduo e tratamento, respectivamente), e probabilidade de F exceder este valor é $1-\alpha/2$;

$F_{(\alpha/2; gl_{Res}, gl_{Trat})}$: valor do percentil da distribuição F de Snedecor, com graus de liberdade gl_{Res} e gl_{Trat} (resíduo e tratamento, respectivamente), e probabilidade de F exceder este valor sendo $\alpha/2$.

Estas análises foram feitas 16 vezes, correspondendo às combinações de duas interações alélicas, quatro herdabilidades e dois valores para repetição no

experimento. Para isso, utilizou-se o método dos quadrados mínimos conforme apresentado por Cruz e Regazzi (1994).

3.2 Simulação dos dialelos circulantes

Para cada um dos dialelos completos referenciados anteriormente, com exceção das situações com 10 repetições, foram simulados todos os dialelos circulantes possíveis, isto é, o número de cruzamentos em que cada pai está envolvido (s). Na presente situação, como o número de pais é par (50), o valor de s foi sempre ímpar, variando de 3 a 49 (Kempthorne e Curnow, 1961).

Para cada um dos 8 dialelos completos já mencionados, foram simulados 50 dialelos circulantes para cada valor de s, sendo que, na obtenção das 50 simulações de cada situação, os pais foram aleatorizados utilizando a função RANUNI (SAS, 1995). Desta forma, foram simulados 9600 cruzamentos dialélicos circulantes. Cada um deles foi gerado utilizando os procedimentos descritos por Kempthorne e Curnow (1961). Utilizando as médias dos pais e dos correspondentes híbridos obtidos, foi realizada a análise dialélica, pelo método dos quadrados mínimos, conforme procedimento apresentado por Cruz e Regazzi (1994).

Os recursos computacionais disponíveis para geração dos resultados utilizaram o ambiente operacional Windows 95, em microcomputadores compatíveis com IBM PC, com capacidade de memória RAM de 64 MB e processador Pentium de 300Mhz de velocidade de processamento. Os programas desenvolvidos para obtenção de estimativas dos parâmetros do dialelo, simulação dos dialelos circulantes, bem como as demais análises, foram elaboradas com os recursos do sistema SAS, nos módulos de matrizes e de programação (SAS, 1995). Alguns dos programas desenvolvidos para simulação encontram-se no Anexo B.

3.3 Eficiência dos dialelos circulantes

A eficiência dos circulantes em relação ao dialelo completo foi medida utilizando-se alguns critérios descritos a seguir.

3.3.1 Correlação classificatória da capacidade geral de combinação (CGC) dos pais

A partir das estimativas da CGC de ambos os dialelos foi estimada a correlação classificatória de Spearman (r_c) (Snedecor e Cochran, 1971). Como foram simulados 50 dialelos circulantes para cada situação, quanto aos valores de s , h^2 e interação alélica, foi possível identificar os valores mínimo, médio e máximo para cada situação.

3.3.2 Correlação classificatória da capacidade específica de combinação (CGC) dos pais e híbridos

De modo análogo ao descrito no item 3.3.1, foram estimados os valores mínimo, médio e máximo da correlação classificatória de Spearman entre estimativas de CEC no dialelo completo e no dialelo circulante. Como no dialelo circulante há apenas uma fração das combinações híbridas do dialelo completo, evidentemente as estimativas das correlações classificatórias foram obtidas a partir das combinações híbridas comuns.

3.3.3 Proporção de estimativas de CGC dentro do intervalo de confiança

A partir da análise do dialelo completo, item 3.1, foram obtidas as estimativas da capacidade geral de combinação para cada pai (\hat{g}_i), sendo que o intervalo de confiança (IC) foi obtido pela expressão:

$$\hat{g}_i \pm t_{\alpha/2} s(\hat{g}_i) ;$$

em que:

\hat{g}_i : estimativa de CGC do pai ;

$t_{\alpha/2}$: é o quantil superior da distribuição de t de Student para graus de liberdade $g|_{Res}$, cuja $Prob(t > t_{\alpha/2}) = \alpha/2$.

$s(\hat{g}_i)$: erro padrão da estimativa dado por:

$$s(\hat{g}_i) = \sqrt{\frac{(p-1) \cdot QMerro}{p(p+2) \cdot r}} ;$$

em que:

QMerro : quadrado médio do resíduo ou erro da análise de variância;

p : número total de pais.

Como para cada uma das situações foram realizadas 50 simulações do dialelo circulante, foi possível estimar a proporção de casos em que a estimativa de g_i pertencia ao intervalo de confiança do respectivo pai no dialelo completo. A partir da informação de cada pai foi obtida a proporção média de coincidência.

3.3.4 Proporção de estimativas de CEC dentro do intervalo de confiança

A partir da análise do dialelo completo, item 3.1, foram obtidas as estimativas da capacidade específica de combinação para cada pai ou híbrido (\hat{S}_{ij}), sendo que o intervalo de confiança (IC) foi obtida pela expressão:

$$\hat{S}_{ij} \pm t_{\alpha/2} s(\hat{S}_{ij}) ;$$

em que:

\hat{S}_{ij} : estimativa de CEC do cruzamento entre os genitores i e j;

$s(\hat{S}_{ij})$: erro padrão da estimativa, dado por:

$$s(\hat{S}_{ij}) = \sqrt{\frac{p(p-1) \cdot QMerro}{(p+1) \cdot (p+2) \cdot r}} , \text{ se } i=j ;$$

$$s(\hat{s}_{ij}) = \sqrt{\frac{(p^2 + p + 2) \cdot Q_{\text{Merro}}}{(p + 1) \cdot (p + 2)r}}, \text{ se } i \neq j.$$

Também no caso da CEC, para cada uma das situações, foram realizadas 50 simulações do dialelo circulante, possibilitando estimar a proporção de casos em que a estimativa de s_{ij} pertencia ao intervalo de confiança do respectivo pai ou híbrido no dialelo completo. A partir desta informação foi obtida a proporção média de coincidência.

3.4 Dados de dialelos completos obtidos na literatura

Foram utilizados os dados médios dos pais e dos respectivos híbridos de dois dialelos com a cultura do milho. O primeiro, cujos resultados de produção de espigas (t/ha) foram publicados por Ferreira (1993), envolveu 28 variedades de milho avaliadas em duas repetições, no município de Sete Lagoas (MG). O segundo, publicado por Napolini Filho et al. (1981) no qual produtividades médias de produção de espigas (t/ha) de 18 populações de polinização aberta foram avaliadas em duas localidades, Sete Lagoas (MG) e Piracicaba (SP), em duas repetições. Os resultados médios encontram-se nos Anexos A e B.

A partir dos dados médios da produção de espigas (t/ha) e o quadrado médio do erro foi efetuada a análise dialélica utilizando o modelo de Griffing (1956), método II, cujos procedimentos foram semelhantes ao descrito no item 3.1.

Simularam-se dialelos circulantes, variando o número de cruzamentos por genitores de 3 a 27 no caso do experimento de Ferreira (1993) e de 3 a 17 no experimento de Napolini Filho et al. (1981). Em todos os casos, novamente 50 simulações foram consideradas.

As comparações entre os dialelos circulantes e o completo foi efetuada de modo análogo aos apresentados nos itens 3.3.1 a 3.3.4.

4 RESULTADOS E DISCUSSÃO

4.1 Dialelos simulados

Nas Tabelas 5 e 6 são apresentados os resumos das análises de variância, considerando-se o dialelo proposto com 50 linhagens, com e sem dominância e diferentes herdabilidades (h^2) ao nível de média das combinações híbridas para caracteres envolvidos. Inicialmente, chamam a atenção, as estimativas do coeficiente de variação experimental (CV). Como era esperado, com o aumento da h^2 houve redução nessa estimativa, ou seja, houve melhoria na precisão experimental, evidenciando que a simulação foi coerente com a proposta.

Destaque também deve ser dado ao fato de que, com a alteração no número de repetições de 2 para 10, no experimento de avaliação das combinações híbridas, as estimativas dos coeficientes de variação aumentaram (Tabelas 5 e 6), resultado que a princípio pode parecer estranho. Entretanto, é importante ressaltar que as estimativas foram obtidas a partir de uma herdabilidade (h^2) pré estabelecida. Assim, como a h^2 é fornecida pela expressão $h^2 = \sigma_G^2 / (\sigma_G^2 + \sigma_e^2 / r)$, infere-se que $\sigma_e^2 = r(\sigma_G^2(1 - h^2)) / h^2$. Desse modo, como a variância genética (σ_G^2) é também fixa, com o aumento de r evidentemente o componente da variância do erro, σ_e^2 , aumenta. Contudo, a amplitude do intervalo de confiança da herdabilidade (h^2) é superior quando se emprega um menor número de repetições, ou seja, com maior número de repetições, a estimativa de h^2 é muito mais fidedigna, embora a estimativa do

CV tenha sido maior. É oportuno também salientar que, independentemente do número de repetições, com o aumento da herdabilidade ocorre redução na contribuição do erro experimental, a amplitude do intervalo de confiança para h^2 paramétrica reduz.

Como era esperado também, uma vez que as linhagens simuladas apresentavam uma ampla variação na constituição genética, detectou-se diferença significativa ($P < 0,01$) para a capacidade geral de combinação em todos os casos considerados, realçando que a simulação foi pertinente com a proposta.

Vale ressaltar o que ocorreu com a significância observada para a fonte de variação capacidade específica de combinação (CEC). Quando foi considerado o modelo sem dominância, a CEC não foi significativa em todos os casos, exceto quando se considerou 10 repetições e h^2 75%. Esse resultado é coerente, pois a CEC só será diferente de zero na presença de dominância. Isso é facilmente observado na expressão apresentada por Vencovsky (1987), que mostra que a CEC do híbrido ij em um loco é dada por $2[(\bar{p}-p_i)(p_j-\bar{p})\delta]$. Nessa expressão, \bar{p} é a frequência média do alelo favorável no loco considerado em todas as n linhagens; p_i e p_j são as frequências do referido alelo nas linhagens i e j, respectivamente, e δ o desvio do heterozigoto em relação à média dos homozigotos, ou seja, o efeito de dominância. Essa expressão é similar à apresentada por Falconer (1981) que demonstra ser a heterose (h) função da divergência genética (Y) ao quadrado e da dominância (δ), ou seja, $h = Y^2\delta$. Assim, desde que $\delta = 0$, a CEC deve ser nula em todos os casos. A exceção ocorrida, com h^2 75%, só pode ser explicada considerando um erro tipo I, isto é, rejeição de uma hipótese verdadeira (Mood, Graybill e Boes, 1974).

Nos modelos com dominância, contudo, esperava-se que a CEC fosse significativa em todas as situações, o que de fato ocorreu quando se utilizou 10 repetições. Entretanto, com 2 repetições, isso só foi verificado com h^2 a partir de

50% (Tabelas 5 e 6). Para justificar este resultado até certo ponto contraditório, com 2 repetições, a explicação mais plausível, está associada ao poder do teste. Com maior número de repetições, mesmo com baixa h^2 , há um aumento no poder do teste F, no caso, em detectar a significância (Mood, Graybill e Boes, 1974).

Este aspecto é muito importante para demonstrar que as inferências genéticas só podem ser corretamente retiradas dos cruzamentos dialélicos se as avaliações dos híbridos forem efetuadas em experimentos com número suficiente de repetições. Isso é particularmente importante para caracteres de baixa herdabilidade como é o caso da produtividade de grãos. Considerando que nos dialelos completos, quando o número de genitores é elevado, o número de combinações híbridas (Tabela 1) a serem avaliadas assume uma magnitude tal que praticamente inviabiliza a condução de experimentos com número de repetições que possibilite precisão compatível a que as inferências genéticas necessitam. Esse é um argumento muito forte para o emprego de dialelos parciais, e principalmente circulantes, ocorrendo redução drástica do número de híbridos a serem avaliados (Hallauer e Miranda Filho, 1981; Cruz e Regazzi, 1994).

TABELA 5 Resumo das análises de variância e intervalos de confiança (LI e LS a 95%) para a herdabilidade, considerando as simulações do dialelo com e sem dominância, para diferentes herdabilidades e considerando duas repetições na avaliação das combinações híbridas.

FV	GL	Herdabilidades (modelos com dominância)											
		10%			20%			50%			75%		
		QM	F	P%	QM	F	P%	QM	F	P%	QM	F	P%
Trat	1274	190276,4	0,99	57	102122,7	1,28	0,0	40909,6	2,03	0,0	26294,7	4,02	0,0
CGC	49	441322,4	2,29	0,0	500010,0	6,28	0,0	443102,9	22,0	0,0	387915,2	59,3	0,0
CEC	1225	180234,5	0,94	86	86207,2	1,08	8,7	24821,9	1,23	1,27	11829,9	1,81	0,0
Erro	1275	192343,6			79606,6			20116,4			6539,4		
Média		1105,1			1093,4			1101,0			1097,0		
CV (%)		39,69			25,80			12,88			7,37		
h ² (LI)		-12,8			13,0			45,1			72,2		
h ² (LS)		9,4			30,2			55,9			77,7		

FV	GL	Herdabilidades (modelos sem dominância)											
		10%			20%			50%			75%		
		QM	F	P%	QM	F	P%	QM	F	P%	QM	F	P%
Trat	1274	141734,2	0,98	64	76121,9	1,27	0,0	30903,7	2,05	0,0	19855,0	4,05	0,0
CGC	49	407932,1	2,83	0,0	469928,2	7,87	0,0	429139,2	28,4	0,0	384584,1	78,4	0,0
CEC	1225	131086,3	0,91	95	60369,2	1,01	43	14974,3	0,99	57	5265,8	1,07	11
Erro	1275	144277,8			59713,3			15089,4			4905,3		
Média		1007,1			997,1			1003,6			1000,2		
CV (%)		37,71			24,51			12,24			7,00		
h ² (LI)		-13,6			12,4			45,5			72,4		
h ² (LS)		8,8			29,7			56,2			77,9		

TABELA 6 Resumo das análises de variância e intervalos de confiança (LI e LS a 95%) para a herdabilidade, considerando as simulações do dialelo com e sem dominância, para diferentes herdabilidades e considerando dez repetições na avaliação das combinações híbridas.

FV	GL	Herdabilidades (modelos com dominância)											
		10%			20%			50%			75%		
		QM	F	P%	QM	F	P%	QM	F	P%	QM	F	P%
Trat	1274	1051196	1,18	0,0	501121	1,27	0,0	196187	1,98	0,0	137772	4,24	0,0
CGC	49	2891548	3,25	0,0	2261282	5,72	0,0	2073311	20,9	0,0	2081349	64,0	0,0
CEC	1225	977582	1,10	1,12	430819	1,09	1,96	121102	1,22	0,0	60028	1,85	0,0
Erro	11475	891036			395546			99036			32501		
Média		1106,5			1095,7			1096,3			1094,5		
CV (%)		85,3			57,4			28,7			16,5		
h ² (LI)		5,39			14,5			45,3			74,4		
h ² (LS)		24,0			27,4			53,5			78,3		

FV	GL	Herdabilidades (modelos sem dominância)											
		10%			20%			50%			75%		
		QM	F	P%	QM	F	P%	QM	F	P%	QM	F	P%
Trat	1274	780717	1,17	0,0	373953	1,26	0,0	146504	1,97	0,0	103927	4,26	0,0
CGC	49	2574856	3,85	0,0	2147088	7,24	0,0	2025199	27,3	0,0	2045952	83,9	0,0
CEC	1225	708952	1,06	8,2	303028	1,02	31,6	71356	0,96	82,7	26246	1,08	3,28
Erro	11475	668371			296701			74287			24379		
Média		1007,1			997,1			1003,6			1000,2		
CV (%)		37,71			24,51			12,24			7,00		
h ² (LI)		-13,6			12,4			45,5			72,4		
h ² (LS)		8,8			29,7			56,2			77,9		

O emprego de dialelos circulantes ainda é muito pequeno, sobretudo pela desconfiança dos melhoristas sobre a qualidade das informações geradas a partir deles. Infelizmente não foram encontradas informações, na literatura, sobre o grau de confiabilidade dos dialelos circulantes em relação aos dialelos completos. Pensando em preencher esta lacuna, optou-se por um modo exequível de se obter tais informações por meio de simulações de dados. Considerando-se as 50 linhagens referidas anteriormente, foi possível simular todas as possibilidades de dialelos circulantes, isto é, variou-se o número de cruzamentos (s) para cada pai, atendendo às restrições impostas pelo algoritmo proposto (Kempthorne e Curnow, 1961).

Quando se realiza um dialelo, uma das preocupações é classificar os pais, com relação à sua capacidade geral de combinação (CGC). Assim, pode-se inferir sobre a eficiência dos dialelos circulantes, em relação aos dialelos completos, a partir da classificação obtida dos pais, com relação à CGC, nas duas situações. Nas Figuras 1,2,3 e 4 são apresentadas as estimativas das correlações classificatórias de Spearman, considerando diferentes valores de h^2 e s, na presença e ausência de dominância, respectivamente. Chama a atenção, inicialmente, o fato de que, com h^2 acima de 50%, as correlações obtidas são superiores a 0,9, mesmo com o menor valor de s utilizado, independente da interação alélica. Infere-se, então, que para caracteres de alta h^2 , os dialelos circulantes fornecem praticamente a mesma classificação dos dialelos completos quanto à capacidade de combinação dos pais envolvidos.

Esse resultado é muito expressivo, pois pode reduzir acentuadamente o trabalho dos melhoristas e, sobretudo, possibilitar que se explore com maior intensidade o grande número de genitores disponíveis. No exemplo considerado, essa vantagem fica bem evidente. Para envolver todas as 50 linhagens no dialelo completo, deveriam ser obtidas 1225 combinações híbridas, além disso elas

deveriam ser testadas em ensaios com número razoável de repetições, como já comentado, Já no dialelo circulante com $s = 3$, seria necessário obter e testar apenas 75 combinações híbridas, o que é muito mais factível. Contudo, a maioria dos caracteres de importância econômica, como produção de grãos, apresenta estimativas de h^2 geralmente inferiores a 50%, tanto para plantas autógamias (Ramalho, Santos e Zimmerman, 1993; Singh, Ceccarelli e Hamblin, 1993) como para alógamas (Lamkey e Hallauer, 1986). Na simulação, constatou-se que quando h^2 é baixa, a classificação dos pais com relação à CGC variou de forma mais acentuada com o valor de s . Porém, mesmo com $s = 3$, $h^2 = 10\%$ e na presença de dominância, o coeficiente de correlação de Spearman médio foi superior a 0,5. Com $s = 11$ esse valor já ultrapassa 0,7, e com $s = 25$, metade do número de pais, já atinge novamente o valor de 0,9. Esses resultados permitem inferir que os dialelos circulantes possibilitam classificar adequadamente os pais com relação à CGC, mesmo quando o carácter considerado tem baixa herdabilidade. Nesse caso, para maior segurança, pode-se realizar o dialelo circulante com maior número de cruzamentos por pai (s maior).

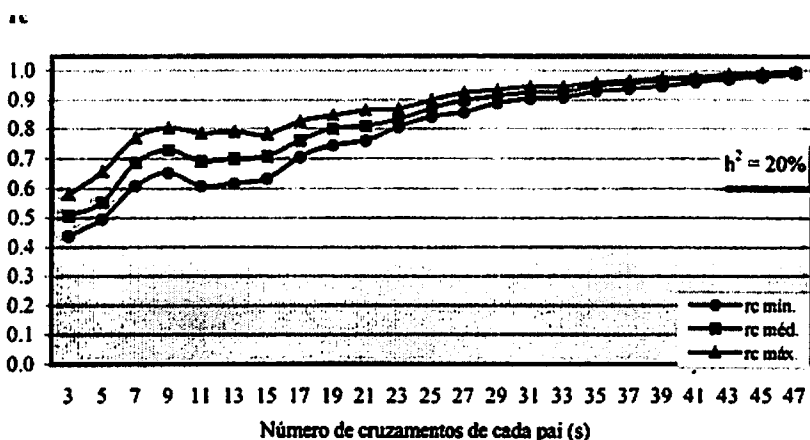
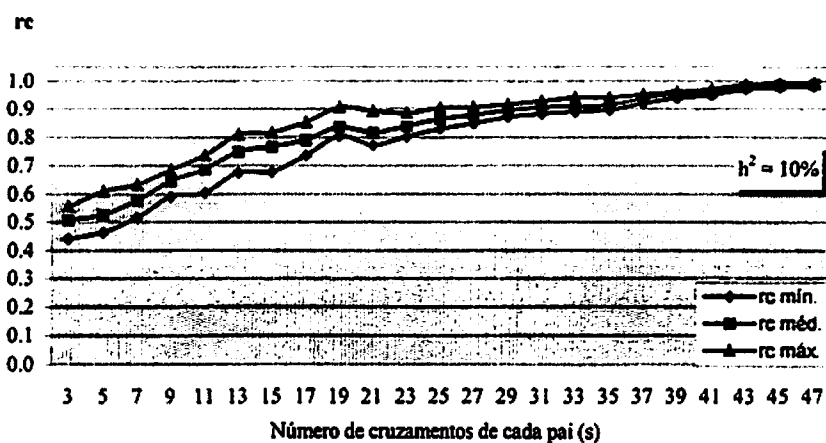


FIGURA 1 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman (rc) entre as estimativas de capacidade geral de combinação (CGC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades ($h^2 = 10\%$ e $h^2 = 20\%$), número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo com dominância.

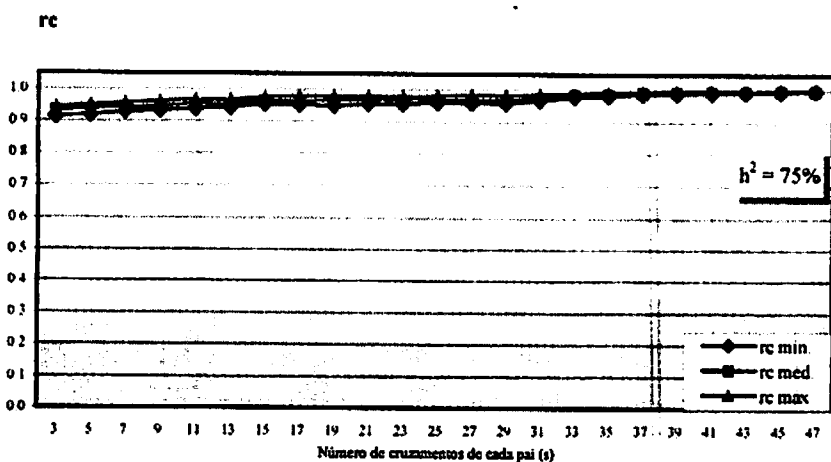
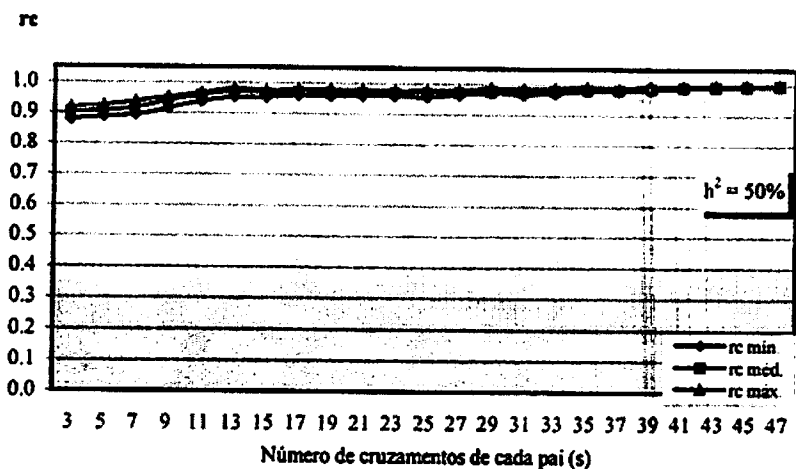


FIGURA 2 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman (r_c) entre as estimativas de capacidade geral de combinação (CGC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades ($h^2 = 50\%$ e $h^2 = 75\%$), número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo com dominância.

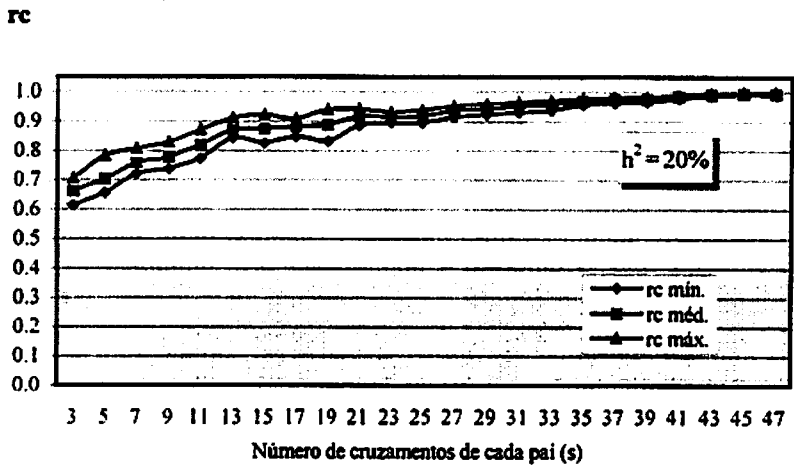
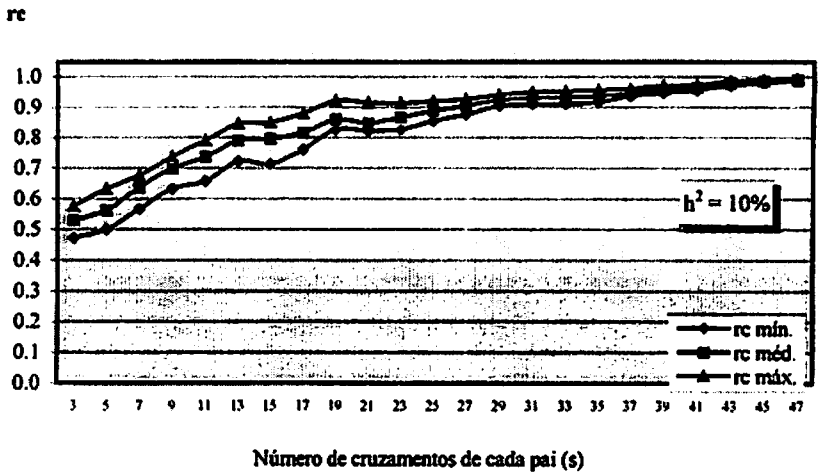


FIGURA 3 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman (rc) entre as estimativas de capacidade geral de combinação (CGC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades ($h^2 = 10\%$ e $h^2 = 20\%$), número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo sem dominância.

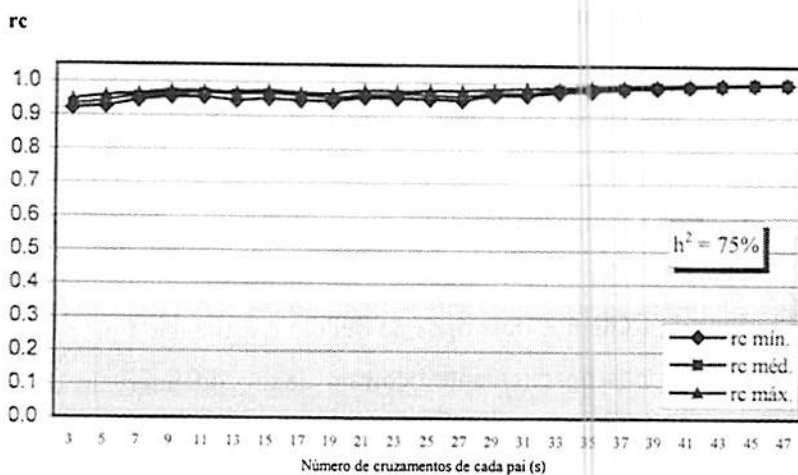
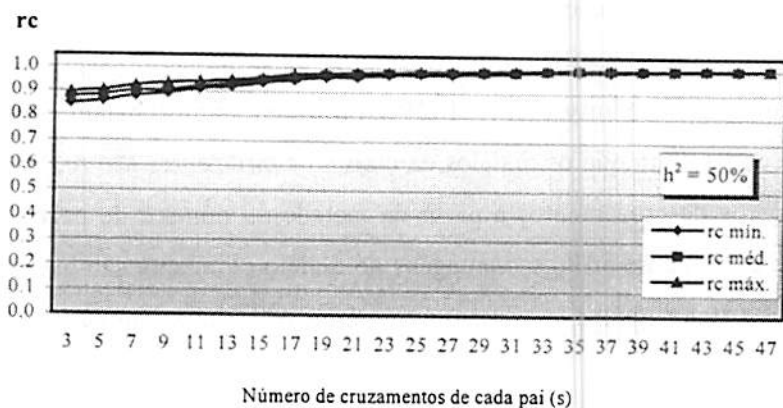


FIGURA 4 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman (rc) entre as estimativas de capacidade geral de combinação (CGC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades ($h^2 = 50\%$ e $h^2 = 75\%$), número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo sem dominância.

Mesmo com esta última recomendação de aumentar o valor de s , buscando uma maior segurança, é necessário enfatizar que esse valor não necessita ultrapassar 50% do número de pais, até certo ponto contrariando as conclusões de Murty e Anand (1966) e Murty, Arunachalam e Anand (1967), que concluíram que os dialelos parciais e os circulantes são recomendados para números de cruzamentos a partir da metade do número de pais. Mesmo nesse caso extremo de 50%, as vantagens do dialelo circulante são marcantes. Com p igual a 50 e s igual a 25, seriam necessários 625 cruzamentos, ao invés dos 1225 já comentados.

Alguns melhoristas, além da classificação dos genitores com relação à CGC, podem estar interessados nas próprias estimativas das capacidades de combinação, sobretudo por permitirem verificar qual o grau de discernimento entre os genitores, inclusive por meio de testes estatísticos apropriados. Esse poderia ser um outro argumento contra a utilização do dialelo circulante, pois sem envolver todas as combinações híbridas, as estimativas de CGC teriam magnitudes não comparáveis às obtidas no dialelo completo.

Para avaliar a efetividade ou não desse argumento, foram obtidas as estimativas de CGC nos dois tipos de dialelo e estimada a probabilidade de que os valores obtidos no circulante ocorram dentro do intervalo de confiança da CGC do dialelo completo (Figuras 5, 6, 7 e 8). A proporção apresentada em cada situação refere-se às estimativas de CGC no dialelo circulante, para cada um dos pais, que ocorreram no intervalo de confiança da CGC obtida no dialelo completo, para o referido pai. Em princípio, os resultados obtidos são semelhantes aos encontrados para as estimativas da correlação de Spearman relatados anteriormente, ou seja, com o aumento no número de pais envolvidos no cruzamento, maior s , a proporção de valores dentro do intervalo de confiança aumenta.

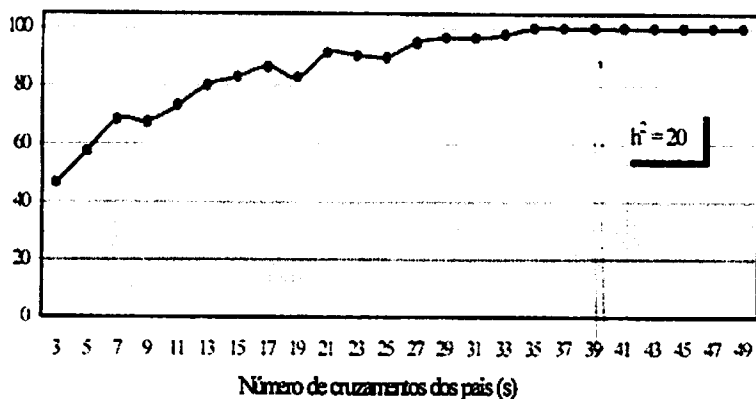
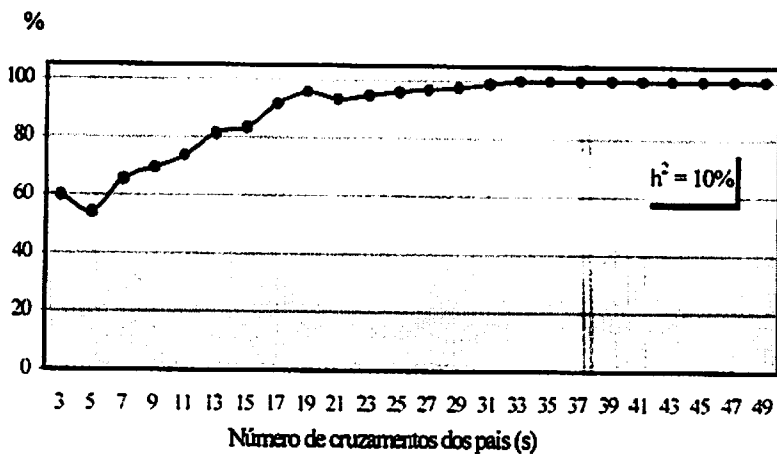


FIGURA 5 Proporção das estimativas da capacidade geral de combinação (CGC) obtidas no diallelo circulante, que coincidem com o intervalo de confiança para CGC do diallelo completo, considerando-se cinquenta simulações para diferentes valores de herdabilidade ($h^2 = 10\%$ e $h^2 = 20\%$), do carácter e número de cruzamentos de cada genitor (s), para o modelo com dominância.

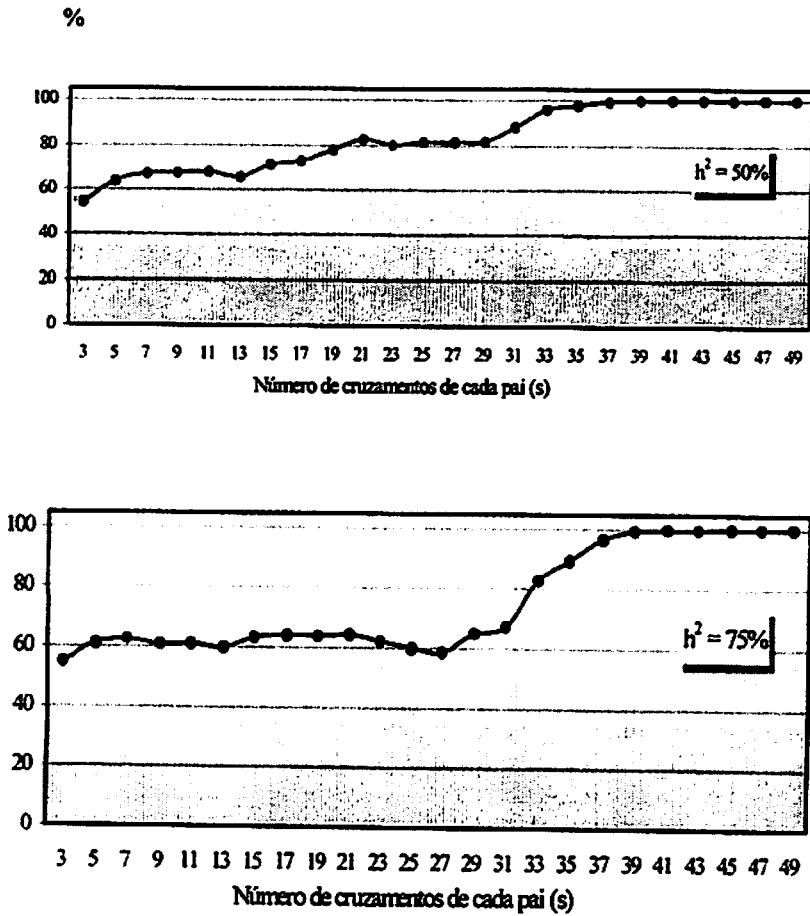


FIGURA 6 Proporção das estimativas da capacidade geral de combinação (CGC) obtidas no dialelo circulante, que coincidem com o intervalo de confiança para CGC do dialelo completo, considerando-se cinquenta simulações para diferentes valores de herdabilidade ($h^2 = 50\%$ e $h^2 = 75\%$), do carácter e número de cruzamentos de cada genitor (s), para o modelo com dominância.

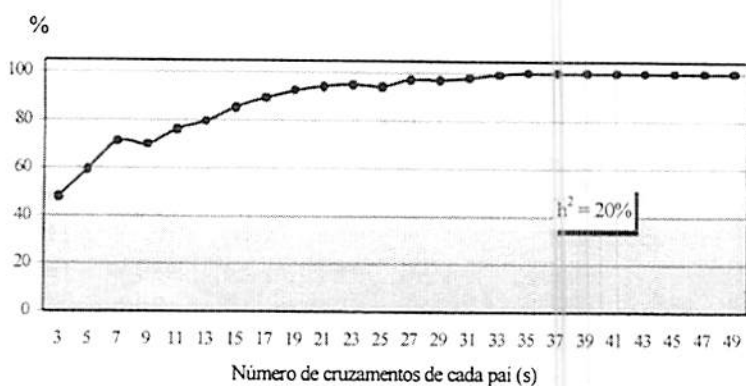
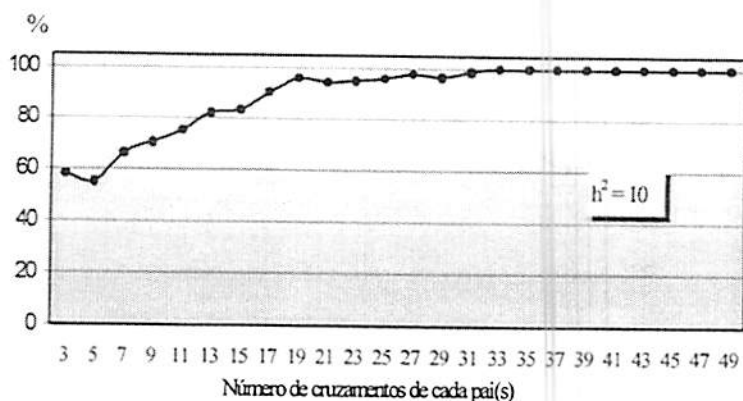


FIGURA 7 Proporção das estimativas da capacidade geral de combinação (CGC) obtidas no dialelo circulante, que coincidem com o intervalo de confiança para CGC do dialelo completo, considerando-se cinquenta simulações para diferentes valores de herdabilidade ($h^2 = 10\%$ e $h^2 = 20\%$), do carácter e número de cruzamentos de cada genitor (s), para o modelo sem dominância.

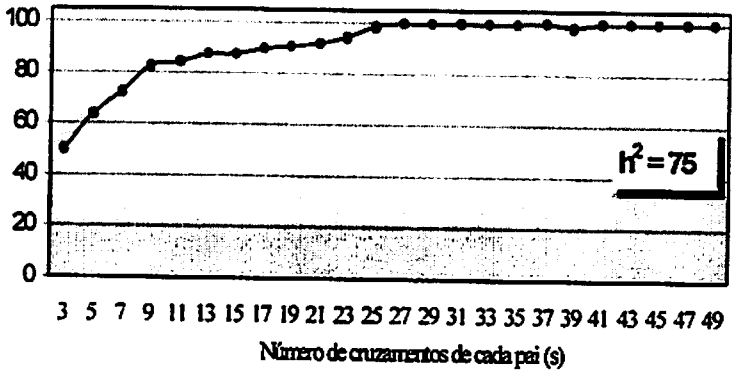
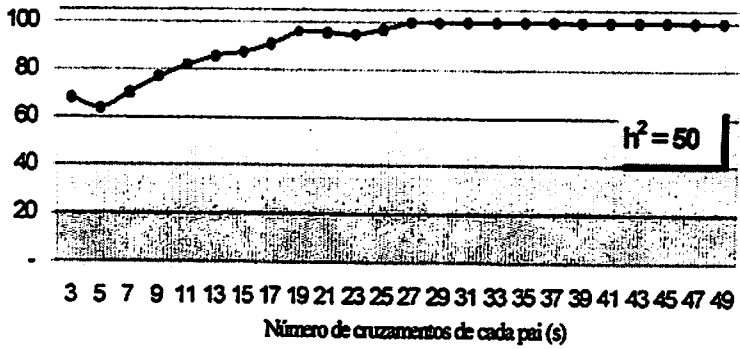
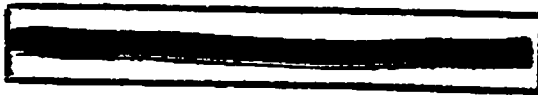


FIGURA 8 Proporção das estimativas da capacidade geral de combinação (CGC) obtidas no dialelo circulante, que coincidem com o intervalo de confiança para CGC do dialelo completo, considerando-se cinquenta simulações para diferentes valores de herdabilidade ($h^2 = 50\%$ e $h^2 = 75\%$), do carácter e número de cruzamentos de cada genitor (s), para o modelo sem dominância.

Há, contudo, uma diferença marcante no que se refere à variação com a herdabilidade do caráter. Com o aumento de h^2 , em geral, a proporção de ocorrência no intervalo reduziu, para um mesmo s . Isso é facilmente explicado considerando que, com o aumento de h^2 , o erro é menor, o intervalo de confiança é mais curto, reduzindo a chance de cair no intervalo considerado. Essa variação só não é mais expressiva porque as estimativas de CGC obtidas no dialelo circulante, com maior h^2 , pela mesma razão já comentada, oscilou menos e aumenta o número de casos que recaem no intervalo considerado.

As proporções obtidas novamente realçam que os dialelos circulantes fornecem praticamente as mesmas estimativas de CGC do dialelo completo, tanto na presença de dominância (Figuras 5 e 6) e especialmente na ausência da mesma (Figuras 7 e 8). Veja que, mesmo com s igual a 3, h^2 de 10%, na presença de dominância, 60% dos valores de CGC do dialelo circulante ocorreram no intervalo de CGC do dialelo completo. Novamente deve ficar evidenciado que, se o melhorista está interessado num ajustamento mais estreito das estimativas, a opção é utilizar um maior número de cruzamentos por genitor. Esse número contudo, seria, no máximo 50% do valor de p .

As estimativas da capacidade geral de combinação são úteis para direcionar os futuros trabalhos de hibridação, isto é, genitores com alta CGC devem dar boas combinações híbridas em cruzamentos com os demais genitores (Vencovsky, 1987). Entretanto, quando é realizado o dialelo, o melhorista está interessado em dar continuidade ao seu trabalho, isto é, identificar o par ou pares de combinações híbridas com potencial de permitir identificar linhagens promissoras (Ramalho, Santos e Zimmerman, 1993). Para isso, é necessário obter informações sobre a capacidade específica de combinação (CEC).

Como já salientado, elas só serão diferentes de zero na presença de dominância (Vencovsky, 1987; Falconer, 1981) e quanto maior a divergência dos dois genitores envolvidos. Assim, uma alta estimativa de CEC, pelo menos

em princípio, indica que aquela combinação híbrida possui um maior número de locos em heterozigose e, portanto, a população segregante dela derivada terá maior amplitude de variação (Abreu, 1997), evidentemente ampliando a chance de selecionar linhagens com desempenho superior aos pais. Do exposto, a estimativa de CEC é de fundamental importância.

No dialelo circulante, analogamente ao dialelo completo, pode-se obter estimativas de CEC dos pares de combinações envolvidas. Novamente foram obtidas por simulação, de maneira semelhante às estimativas de CGC, a correlação classificatória de Spearman e a proporção de casos de estimativas de CEC do circulante que ocorreram no intervalo de confiança da mesma CEC no dialelo completo (Figuras 9 e 10). Observa-se que, no caso das estimativas da correlação de Spearman, os valores foram superiores àqueles obtidos para CGC, para os mesmos valores de s . Com s igual a 3 e mesmo com h^2 de 10%, a correlação média foi próxima a 0,8. Com s igual a 11, os valores da correlação já são superiores a 0,9 e a partir daí, há uma estabilização.

A proporção de ocorrência da estimativa de CEC no dialelo circulante no intervalo de confiança da CEC do dialelo completo também evidencia que o dialelo circulante fornece estimativas de CEC muito similares àquelas do dialelo completo. A proporção de casos no intervalo de confiança, mesmo com s igual a 3, foi superior a 75%, para todas as herdabilidades. Novamente, com s igual a 11, os valores foram superiores a 90%, estabilizando a partir daí (Figura 9 e 10).

O fato mais relevante, contudo, é que no dialelo circulante só é estimada uma pequena proporção das capacidades específicas de combinação possíveis. Novamente considerando o exemplo com p igual a 50, no dialelo completo seriam obtidas 1225 combinações híbridas e, conseqüentemente igual número de estimativas da CEC, enquanto que no dialelo circulante, com s igual a 3, apenas 75. Isto é, ficariam faltando 1150 estimativas de CEC.

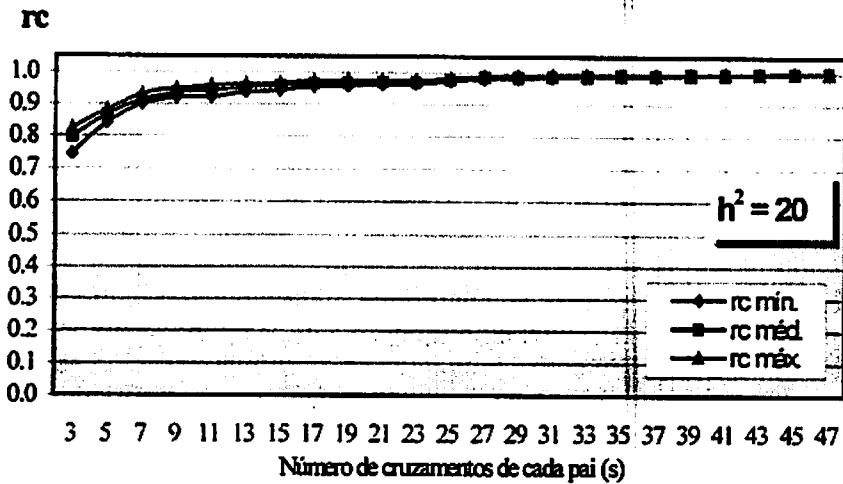
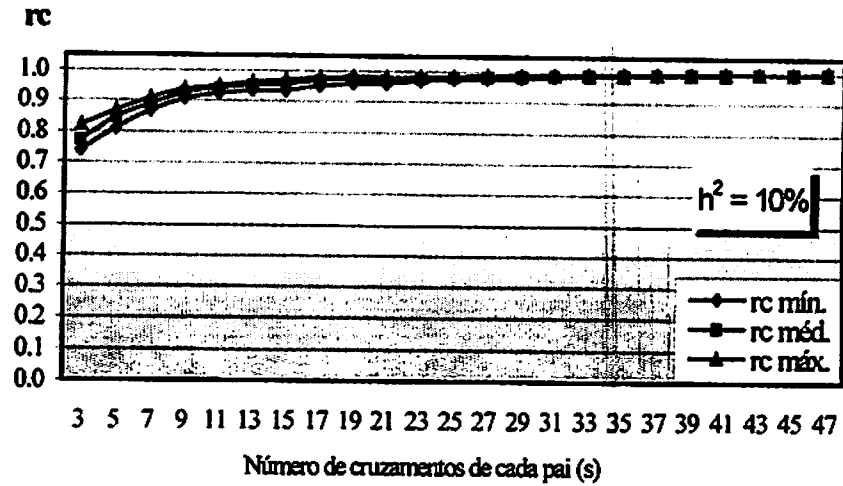


FIGURA 9 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman entre as estimativas de capacidade específica de combinação (CEC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades ($h^2 = 10\%$ e $h^2 = 20\%$), número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo com dominância.

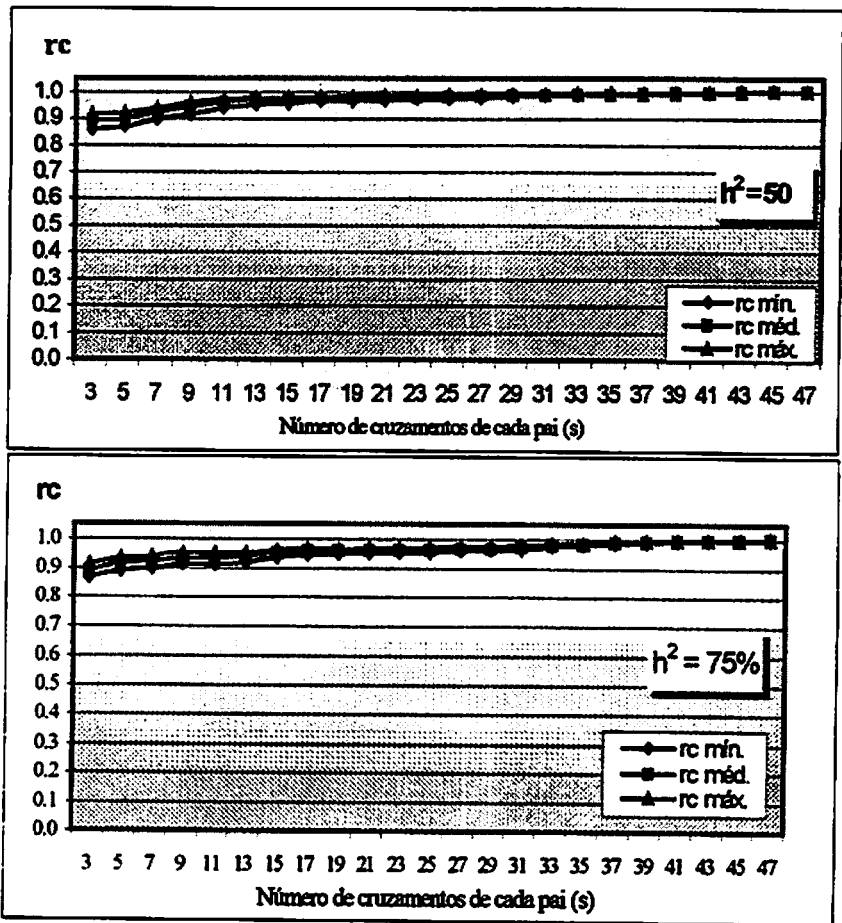


FIGURA 10 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman entre as estimativas de capacidade específica de combinação (CEC) no dialelo completo e no circulante, considerando-se diferentes herdabilidades ($h^2 = 50\%$ e $h^2 = 75\%$), número de cruzamentos de cada pai (s) no circulante, provenientes de cinquenta simulações de cada caso, no modelo com dominância.

Como já salientado, estas estimativas de capacidades específicas ausentes são fundamentais ao trabalho dos melhoristas, o que se tornaria uma séria restrição ao emprego desse tipo de dialelo. Essa restrição poderia ser atenuada com algumas alternativas. A primeira delas é que, como já comentado, a CEC é tanto mais expressiva quanto maior a divergência dos pais, assim, os melhoristas poderiam, no momento de realizar os cruzamentos, orientá-los de modo tal a favorecer as combinações entre genitores mais divergentes. Assim procedendo, só seriam obtidas as combinações que efetivamente são mais promissoras, tornando-se uma vantagem adicional a favor dos dialelos circulantes, pois seria obtido e principalmente avaliado um número restrito de combinações. Para identificar os pais mais divergentes, há algumas estratégias, tais como a origem das linhagens, coeficiente de parentesco, o uso de análise multivariada ou o emprego de marcadores (Baenziger e Peterson, 1991; Pereira Reis, 1998; Abreu, 1997).

A segunda alternativa para atenuar o limitado número de estimativas de CEC envolvidas no dialelo circulante é a possibilidade de se utilizar procedimentos genético-estatísticos que permitam estimar os valores de CEC em falta. Nesse contexto, algumas opções têm sido apresentadas, como o emprego dos modelos mistos e a predição por meio do BLUP (Bernardo, 1995) e predição pelo uso de marcadores, em que os autores comparam modelos de regressão com o uso de marcadores como componentes, com predição por meio de BLUP (Charcosset et al., 1998).

4.2 Dialelos de dados experimentais obtidos na literatura

Na literatura, são escassos os dialelos envolvendo um grande número de pais. Na realidade, há na literatura um enorme número de publicações envolvendo cruzamentos dialélicos, em várias espécies, contudo, o número de pais em estudo é quase sempre inferior a 15 (Gama et al., 1995; Melo, 1996;

Freire Filho, 1988; Otubo, 1994; Santos, 1984; Lopes, 1984). Em apenas dois trabalhos, com a cultura do milho, em que foram utilizados 18 pais (Naspolini Filho et al., 1981) e 28 pais (Ferreira, 1993). Essa informação reforça o argumento de que o potencial dos dialelos tem sido pouco explorado pelos melhoristas, em função de restringir o número de pais envolvidos. Esse fato confirma as observações anteriores de que os cruzamentos dialélicos circulantes deveriam ser mais utilizados.

Os resultados das análises de variância obtidos nos experimentos com a cultura do milho por Naspolini Filho et al. (1981) e Ferreira (1993) estão apresentados na Tabela 7. Constatou-se diferença significativa entre os tratamentos em ambos os casos. O mesmo ocorreu para a fonte de variação capacidade geral de combinação. Contudo, no que se refere à capacidade específica de combinação, observou-se diferença significativa apenas na análise realizada com as médias publicadas do experimento com milho obtido com 18 pais.

Vale ressaltar que ocorreu diferença marcante na estimativa de h^2 para seleção entre os híbridos. Com as médias advindas dos experimentos com 28 pais, obteve-se h^2 igual a 28,2%, enquanto que no caso do experimento com 18 pais, a estimativa de h^2 foi de 63,1%. Depreende-se, então, que no caso de Ferreira (1993) provavelmente ocorreu o fato já comentado com os resultados dos dialelos simulados com duas repetições que, mesmo tendo dominância, a capacidade específica de combinação só foi detectada com h^2 acima de 50%. Este fato foi inclusive salientado por Ferreira (1993), que não detectou efeito significativo para fonte de variação CEC, embora a depressão por endogamia dos pais fosse elevada e significativa, indicando a presença de dominância e também de divergência genética. Ressalta-se que, da mesma forma que os experimentos simulados, o experimento de Ferreira (1993) foi conduzido com duas repetições.

Tabela 7 Resumos das análises de variância dos dialelos e intervalos de confiança para a herdabilidade (LI e LS) com 95% de confiança, considerando os experimentos com dados reais.

Fontes de Variação	Experimentos com dados reais							
	Experimento 1 (Milho - Ferreira, 1993)				Experimento 2 (Milho - Napolini Filho et al., 1981)			
	G. L.	QM	F	P %	G. L.	QM	F	P %
Trat.	405	1,8763	1,39	0,00	170	2,2711	2,71	0,04
CGC	27	7,8212	5,81	0,00	17	10,4695	12,58	0,00
CEC	378	1,4517	1,08	22,96	153	1,3601	1,62	0,00
Resíduo	405	1,3471			1020	0,8378		
Médias (t/ha)		7,25				4,27		
C. V. %		16,01				21,44		
\hat{h}^2		39,28				63,11		
h^2 LI		12,74				54,00		
LS		40,93				70,97		

Na Figura 11, são apresentadas as correlações classificatórias de Spearman envolvendo as estimativas de CGC obtidas no dialelo experimental completo e as diferentes simulações dos dialelos circulantes possíveis. Verifica-se que, em média, já no circulante com 3 cruzamentos por genitor, a correlação foi próxima a 0,6, para o dialelo de menor h^2 (Ferreira, 1993) e próxima a 0,9 para o de maior h^2 (Napolini Filho et al., 1981). Observando-se as Figuras 1 e 2, que apresentam os dados simulados com $h^2 = 20\%$, semelhante à obtida por Ferreira (1993), e a com $h^2 = 75\%$, que é semelhante à obtida por Napolini Filho et al. (1981), verifica-se que os resultados foram muito coerentes. Assim, com 11 cruzamentos por genitor, com os dialelos circulantes na simulação (22% dos 50 pais), e $h^2 = 20\%$, a correlação média foi superior a 0,7.

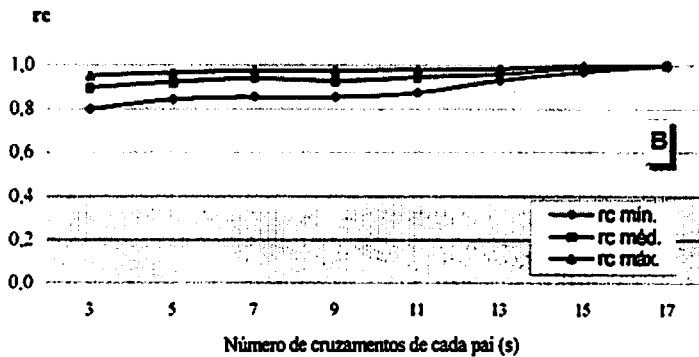
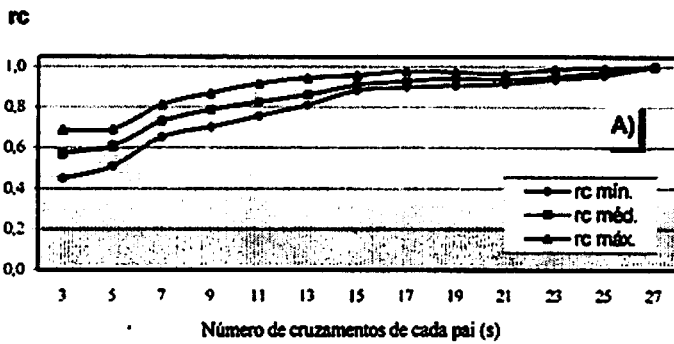


FIGURA 11 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman entre as estimativas de capacidade geral de combinação (CGC) no dialelo completo e no circulante, considerando-se experimentos: A - de Ferreira (1993) e B - Naspolini Filho et al. (1981).

Com os dados de Ferreira (1993), 22% de 28 pais correspondem a 6 e a correlação média foi também próxima de 0,7. A mesma observação é válida quando se compara os resultados de Napolini Filho et al. (1981) com a $h^2 = 75\%$ no dialelo simulado.

Comparando as mesmas figuras, verifica-se que a amplitude de variação nas estimativas mínimas e máximas de correlação de Spearman foram muito semelhantes.

Os mesmos comentários são válidos quando se considera a classificação das estimativas da Capacidade Específica de Combinação (Figura 12).

As proporções de casos de estimativas de CGC, das cinquenta simulações do dialelo circulante que ocorreram no intervalo de confiança para as CGC no dialelo completo, estão mostradas na Figura 13. Verifica-se que mesmo com três pais envolvidos em cada cruzamento ($s = 3$), a proporção está em torno de 50% no experimento com 28 pais e de 70% com 18 pais. A partir de 50% dos pais envolvidos ($s = 14$ com 28 pais ou 9 com 18 pais), essa proporção já ultrapassa 90% de coincidência.

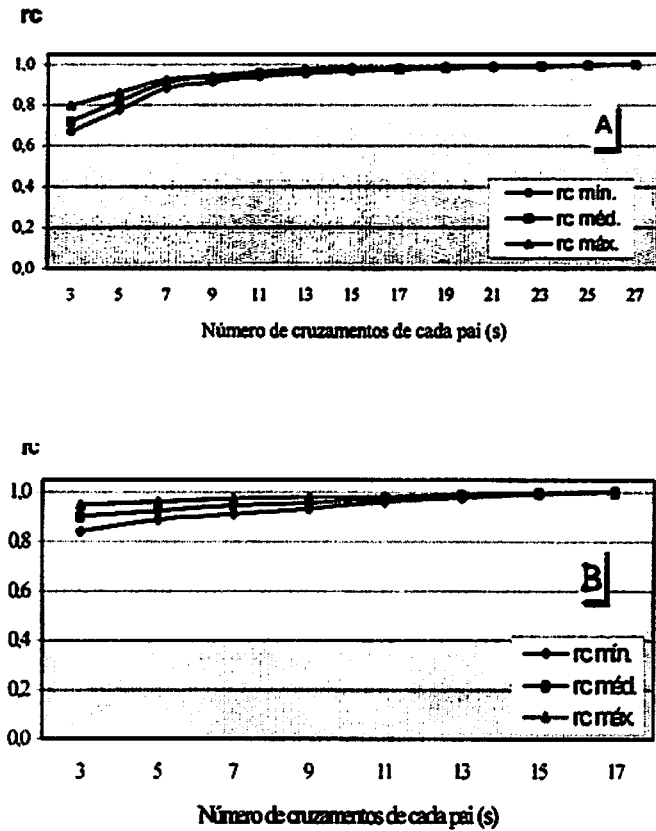


FIGURA 12 Estimativas das médias, mínimas e máximas das correlações classificatórias de Spearman entre as estimativas de capacidade específica de combinação (CEC) no dialelo completo e no circulante, considerando-se experimentos: A - de Ferreira (1993) e B - Naspolini Filho et al. (1981).

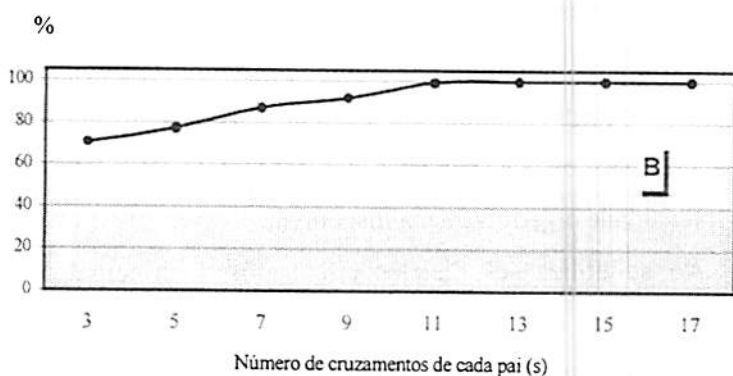
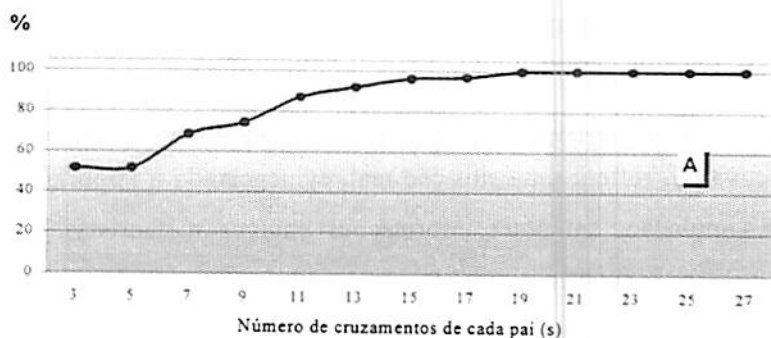


FIGURA 13 Proporção das estimativas da capacidade geral de combinação (CGC) obtidas no dialelo circulante, que coincidem com o intervalo de confiança para CGC do dialelo completo, considerando cinquenta simulações para o número de cruzamentos de cada genitor (s), considerando-se experimentos: A - de Ferreira (1993) e B - Napolini Filho et al. (1981)

Vale a pena comparar novamente esse resultado com aqueles apresentados no dialelo simulado (Figuras 5 e 6). Fica evidente novamente que nas situações comparáveis de h^2 , já salientadas, os resultados foram muito semelhantes.

Todos estes resultados mostram que a simulação do dialelo efetuada foi efetiva em refletir uma situação real, evidenciando a importância dos trabalhos de simulação, em obter informações para os melhoristas, sobretudo porque possibilita envolver situações em que seria praticamente impossíveis serem comparadas ao nível de campo.

Evidenciou-se que os cruzamentos dialélicos circulantes possibilitam aos melhoristas resultados bem semelhantes aos que seriam obtidos com os dialelos completos. Este fato, como já evidenciado, reduz consideravelmente o número de cruzamentos a serem obtidos e avaliados, possibilitando que, na escolha dos pais ou das populações segregantes, seja possível explorar com maior intensidade o germoplasma disponível.

Ao se utilizar os dialelos circulantes, a principal indagação é com relação ao número de cruzamentos de cada genitor, ou seja, o valor de s . Ficou evidenciado que este valor depende da herdabilidade para a seleção de híbridos, podendo ser tanto menor quanto maior for a h^2 . Contudo, mesmo com h^2 baixa, ele não necessita ultrapassar a 50% do número total de genitores. Em princípio, esse resultado é contraditório ao publicado por Murty, Arunachalam e Anand (1967), e por Murty e Anand (1966), que salientam que o dialelo circulante é eficiente com número de cruzamentos por genitor a partir da metade do número total de pais. Ao que parece, os estudos com o objetivo de avaliar a eficiência de circulantes não haviam abrangido até então o número necessário de pais e não haviam atingido a amplitude que os recursos de simulação possibilitaram no presente trabalho.

Um argumento que poderia ser aventado é o de que os melhoristas não possuem condições a "priori" de saber qual h^2 será obtida para a seleção dos híbridos no dialelo, ficando difícil, assim, ter informações sobre a magnitude do valor de s . Um bom critério seria utilizar as estimativas do coeficiente de variação (CV) para a característica de interesse, que é comumente obtido nas condições experimentais em que os híbridos serão avaliados. Se as estimativas do CV são normalmente de magnitude média ou alta, indicando baixa precisão, é um indício de que a h^2 a ser obtida será pequena e, nessa condição, o número de cruzamentos por pai deverá ser maior. Em caso de boa precisão, a h^2 deverá ser alta e o s poderá ser pequeno.

No dialelo circulante, o número de cruzamentos a serem avaliados é muito inferior ao dos dialelos completos, há possibilidade que os experimentos sejam conduzidos com maior número de repetições. Esse procedimento, como já mostrado, aumenta a chance de se detectar significância para a capacidade específica, quando existe dominância, aliada à presença de divergência entre os pais.

5 CONCLUSÕES

- O uso da simulação de dados por computador demonstrou eficiência e coerência com os resultados experimentais obtidos, permitindo destacar a importância da precisão experimental, quanto à questão do poder do teste estatístico, sobretudo em condições de baixa herdabilidade, o que é bastante comum em características de interesse agrônomo.
- Os dialelos circulantes demonstraram eficiência comparável a dos completos, tanto na classificação dos pais quanto à CGC e à CEC, como na magnitude das estimativas desses parâmetros.
- No dialelo circulante, o número de cruzamentos de cada pai (s), afeta as estimativas de CGC e de CEC; entretanto mesmo com um valor mínimo de s , é possível obter uma boa concordância nas estimativas em relação aos dialelos completos.
- Especialmente em condições de baixa herdabilidade, pode ser vantajoso aumentar o número de cruzamentos de cada pai; contudo, a melhoria na eficiência ocorre com o aumento de s até o máximo da metade do número de pais envolvidos, com o que já se consegue uma alta eficiência dos dialelos circulantes em todos os casos.

6 REFERÊNCIAS BIBLIOGRÁFICAS

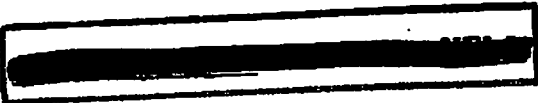
- ABREU, A.F.B. Predição do potencial genético de populações segregantes do feijoeiro utilizando genitores inter-raciais. Lavras: UFLA, 1997. 79p. (Tese-Doutorado em Genética e Melhoramento de Plantas).
- BAENZIGER, P.S.; PETERSON, E.J. Genetic variation: Its origin and use for breeding self-pollinated species. In: STALKER, H.T.; MULTRIPHY, J.P. Plant breeding in the 1990's, North Caroline State University, p.69-100, 1991.
- BEARZOTI, E. Simulação de seleção recorrente assistida por marcadores moleculares em espécies autógamas. Piracicaba: ESALQ, 1997, 230p. (Tese-Doutorado em Genética e Melhoramento de Plantas).
- BERNARDO, R. Genetic Models for predicting maize single-cross performance en unbalanced yield trial data. *Crop Science*, v.35, p.141-147, 1995.
- BERNARDO, R. Best linear unbiased prediction of maize single-cross performance. *Crop Science*, v.36, p. 50-56, 1996.
- BOOTE, K.J.; JONES, J.W.; PICKERING, N.B. Potencial uses and limitations of crop models. *Agronomy Journal*, Gainesville, v.88, n.5, p.704-716, 1996.
- BROWN, D.; ROTHERY, P. Models in biology: mathematics, statistics and computing. Cambridge, UK: Agriculture & Food Research Council, Babraham Institute, 1993, 688p.
- BUROW, M.D. ; COORS, J.G. DIALLEL: a microcomputer program for the simulation and analysis of diallel crosses. *Agronomy-Journal*, v.86, n.1, p.154-158, 1994.
- CHARCOSSET, A.; BONISSEAU, B.; TOUCHEBEUF, O.; BURSTIN, J.; DUBREUIL, P.; BARRIÈRE, Y.; GALLAIS, A.; DENIS, J.B. Prediction of maize hibrid silage performance using marker dada: comparison of several models for specific combining ability. *Crop Science*, v.38, p.38-44, 1998.
- CRUZ, C.D. Programa GENES: aplicativo computacional em genética e estatística. Viçosa: UFV, Imprensa Universitária, 1997, 442p.

- CRUZ, C.D.; VENCOVSKY, R. Comparação de alguns métodos de análise dialélica. *R. Bras. Genet.*, Ribeirão Preto, v.12, p.425-438, 1989.
- CRUZ, C.D.; REGAZZI, A.J. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, Imprensa Universitária, 1994, 390p.
- DACHS, J.N.W. **Estatística computacional**. São Paulo: Livros Técnicos e Científicos Editora S. A. 1988, 236p.
- DUARTE, J.M.; **Estudo de divergência genética em raças de feijão por meio de marcadores RAPD**. Lavras: UFLA, 1998, 78p. (Dissertação-Mestrado em Genética e Melhoramento de Plantas).
- EUCLYDES, R.F. **Uso de sistema para simulação Genesys na avaliação de métodos de seleção clássicos e associados a marcadores moleculares**. Viçosa, 1996, 149p. (Tese-Doutorado em Genética e Melhoramento de Plantas)
- FALCONER, D.S. **Introdução à genética quantitativa**. Trad. SILVA, M.A. de; SILVA, J.C. Viçosa: Imprensa Universitária da UFV, 1981. 279p.
- FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores RAPD, RFLP em análises genética**. Brasília: EMBRAPA-CENARGEN, 1995, 220p.
- FERREIRA, D.F. **Eficiência de Métodos de Mapeamento de Locos Quantitativos (QTLs) e da Seleção Assistida por Marcadores Moleculares Através de Simulação**. Piracicaba: ESALQ, 1995, 210p. (Tese-Doutorado em Genética e Melhoramento de Plantas).
- FERREIRA, D.F. **Métodos de avaliação da divergência em milho e suas relações com os cruzamentos dialélicos**. Lavras, 1993. 72p. (Dissertação - Mestrado em Genética e Melhoramento de Plantas).
- FRASER, A.S. **Simulation of genetic system by automatic digital computers. I. Introduction**. *Aust. J. Biol. Sci.* 10:484-491, 1957a.
- FRASER, A.S. **Simulation of genetic system by automatic digital computers. II. Effects of linkage on rates of advance under selection**. *Aust. J. Biol. Sci.* 10:492-499, 1957b.
- FREIRE Filho, F.R. **Análise genética de um dialélio entre genótipos precoces de soja (*Glycine max* (L.) Merrill)**. Piracicaba: ESALQ, 1988, 245 p. (Dissertação-Mestrado em Genética e Melhoramento de Plantas).
- GAMA, E.E.G. e; HALLAUER, A.R.; LOPES, M.A.; PARENTONI, S.N.; SANTOS, M.X.; GUIMARÃES, P.E.O. **Combining ability among fifteen early cycle maize populations in Brazil**. *Brazilian Journal of Genetic*, v.18, n.4, p.569-577, 1995.

- GARDNER, C.O.; EBERHART, S. A. Analysis and interpretation of the variety cross diallel and related populations. *Biometrics*, Raleigh, v.22, p. 439-452, 1966.
- GERALDI, I. O. ; MIRANDA FILHO, J. B. de. Adapted models for the analysis of combining ability of varieties in partial diallel crosses. *Revista Brasileira de Genética*, Ribeirão Preto, v.11, p.419-430, 1988.
- GILBERT, N.E. Diallel cross in plant breeding. *Heredity*. Edinburgh, v. 12, p.477-92, 1958.
- GRIFFING, B. Concept of general and specific combining ability in relation to diallel crossing systems. *Australian Journal of Biological Science*, Bombay, v.9, p.463-493, 1956.
- GONÇALVES, P.S. Esquema circulante de cruzamentos para avaliação de linhagens de milho (*Zea mays* L.) ao nível interpopulacional. Piracicaba: ESALQ, 1987, 140p. (Tese-Doutorado em Genética e Melhoramento de Plantas).
- HALLAUER, A.R. Methods used in developing maize inbreds. *Maydica*, v.35, p.1-16, 1990.
- HALLAUER, A.R. ; MIRANDA FILHO, J.B. de. *Quantitative genetics in maize breeding*. Ames: Iowa State University Press, 1981. 468p.
- JINKS, J.L.; HAYMAN, B.I. The analysis of diallel crosses. *Maize Genetics Corporations News Letter*, Ithaca, v.27, p.48-54, 1953.
- KEMPTHORNE, O.; CURNOW, R.N. The partial diallel cross. *Biometrics*, v.17, p. 229-50, 1961.
- KERVELLA, J.; ROBERT, N.; FOUILLOUX, G. Influence of recombination on genetic variability. II. Simulation studies. *Agronomie*, v.13, n.5, p.371-379, 1993.
- KNAPP, S.J.; STROUP, W.W.; ROSS, W.M. Exact confidence intervals for heritability on a progeny mean basis. *Crop Science*, Mdison, v.25, n.1, p.192-194, Jan./Fev. 1985.
- LAMKEY, K.R.; HALLAUER, A.R. Herdability estimated from recurrent selection experiments in maize. *Maydica*, Ames, v.32, p. 61-78, 1987.
- LOPES, A. de M. *Análise Genética dos componentes de produção num dialelo entre seis cultivares de arroz (*Oryza sativa* L.) em dois regimes hídricos*. Viçosa: UFV, 1984, 135p. (Tese - Doutorado em Genética e Melhoramento).
- MATHER, D.E.; KANNENBERG, L.W.; Genetic simulation of ratio-based recurrent selection. *Maydica*, v.37, n.2, p.169-172, 1992.

- MIRANDA Filho, J.B.; GERALDI, I.O.; Na adapted model for the analysis of partial diallel crosses. *Ver. Brasil. Genet.* v.7,4, p.677-688, 1984)
- MELO, L.C. Escolha de genitores visando a obtenção de cultivares de feijoeiro tolerantes à baixa temperatura na fase adulta. Lavras: UFLA, 1996, 80p. (Dissertação - Mestrado em Genética e Melhoramento de Plantas).
- MOOD, A.; GRAYBILL, F.A; BOES, D.C. Introduction to the theory os statistics. Tokyo: McGraw-Hill Kogaskusha, 1974, 382p.
- MORRISON, D.F. Multivariate statistical methods. New York: McGraw-Hill, 1967, 415p.
- MORGAN, B. J. T. Elements of simulation. 6ed. London:Chapman & Hall, 1995, 351p.
- MURTY, B.R.; ANAND, I.J. Combining ability and diversity in some varieties of *Linus uritatissemum*. *Indian J. Genet.* v.26, p.21-36, 1966.
- MURTY, B.R.; ARUNACHALAM, V.; ANAND, I.J. Diallel and partial diallel ana'lysis of some yield factors in *Linus uritatissemum*. *Heredity*, v.22, p.35-41, 1967.
- NASPOLINI Filho, V.; GAMA, E.E.G. e; VIANNA, R.T.; MÔRO, J.R. General and specific combining ability for yield in a diallel cross among 18 maize populations (*Zea mays* L.), *Ver. Brasil. Genet.*, v.4, n.4, p.571-577, 1981.
- NAYLOR, T. H.; BALINTFY, J. L.; BURDICK, D. S.; CHU, K. Técnicas de simulação em computadores. São Paulo: Editora Vozes Ltda, trad, 1971, 401p.
- PARTNER, P. R. L; SMITH, M. L.; SPOOR, W.; CLARKSON, M. I. Computer simulation of selection in a hypothetical crop species. *Computer Applications in the Biosciences*, Edinburgh, UK, v.9, n.5, p.597-605, 1993.
- OLIVEIRA, E.B. de; RESENDE, M.D.V. de; HIGA, A.R.; OIADA, G.R.P.V. PAGIS - a software for genetic analysis and selection index for forest experiments. Documentos Centro Nacional de Pesquisa de Florestas, EMBRAPA Centro Nacional de Pesquisa de Florestas, Colombo, PR, Brazil, v.24, p.19, 1994.
- OTUBO, S.T. Controle genético da tolerância do feijoeiro (*Phaseolus vulgaris* L.) à baixas temperaturas na fase de germinação. Lavras: UFLA, 1994. 50p. (Dissertação - Mestrado em Genética e Melhoramento de Plantas).
- RAMALHO, M.A.P.; SANTOS, J.B. dos; ZIMMERMANN, M.J.O. Genética quantitativa em plantas autógamas: aplicações ao feijoeiro. Goiânia: UFG, 1993. 271p.

- REIS, W.P. **Divergência genética entre cultivares brasileiros de trigo recomendados no Brasil.** Piracicaba: ESALQ, 1998. 77p. (Tese-Doutorado em Genética e Melhoramento de Plantas).
- SANTOS, J.B. dos. **Controle genético de caracteres agronômicos e potencialidades de cultivares de feijão (*Phaseolus vulgaris* L.) para o melhoramento genético.** Piracicaba: ESALQ, 1984. 223p. (Tese-Doutorado em Genética e Melhoramento de Plantas).
- SAS Institute. **SAS language and procedures: Usage.** Version 6, 1st ed. Cary NC: SAS Institute Inc., 1995, 373 p.
- SCHWARZBACH, E.; PESEK, J.; HERMAN, M.; HARTMANN, J. **The Mendel-program for the simulation of stochastic genetic processes in teaching and breeding research. Biometrics in plant breeding: proceedings of the eighth meeting of the Eucarpia section Biometrics on plant breeding.** p.439-440, jul. 1991.
- SEARLE, S. R. **Linear Models.** New York: John Wiley & Sons, 1971, 532p.
- SINGH, S.P.; CECCARELLI, S.; HAMBLIN, J. **Estimation of heritability from varietal trials data. Theoretical and Applied Genetics,** Berlin, v.86, p.437-441, 1993.
- SNEDECOR, G. W.; COCHRAN, W. **Statistical methods.** 6. ed. Ames, Iowa. The Iowa State College Press, 1978. 485p.
- SOUZA, E.A. **Alternativas experimentais na avaliação de progênies em programas de melhoramento genético vegetal.** Piracicaba: ESALQ, 1997, 122p. (Tese-Doutorado em Genética e Melhoramento de Plantas).
- SPRAGUE, G. F. ; TATUM, L. A. **General vs. specific combining ability in single crosses of corn. Journal of American Society of Agronomy,** Madison, v.34, p.923-932, 1942.
- TAKEDA, C. **Avaliação de progênies de feijoeiro do cruzamento "ESAL 501" x "A 345" em diferentes densidades de plantio.** Lavras, 1990. 82p. (Dissertação - Mestrado em Genética e Melhoramento de Plantas).
- TINKER, N. A.; MATHER, D. E. GREGOR: **software for genetic simulation. Journal-of-Heredity,** Lakeshore, Canada, v.84, n.3, p.237, 1993.
- VEIGA, R.D.; CAVALCANTI, J.J.V. **Uma aplicação do sistema SAS em modelos incompletos de análise dialélica.** In: 42^o SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, Recife, 1997. **Programas e Resumos ...** Recife: UFRPE, 1997, p34.



VENCOVSKY, R. Alguns aspectos teóricos e aplicados relativos a cruzamentos dialélicos de variedades. Piracicaba: ESALQ, 1970, 59p. (Tese-Livre Docência Escola Superior de Agricultura "Luiz de Queiroz"/ USP).

VENCOVSKY, R. Herança quantitativa. In: PATERNIANI, E. Melhoramento e produção do milho no Brasil. Campinas: Fundação Cargill, 1978. p.122-201.

VILLANUEVA, B.; KENNEDY, B.W. Index versus tandem selection after repeated generations of selection. *Theoretical and applied genetics*, v.85, n.6, p.706-712, 1993.

ZHANG, W.; SMITH, C. Computer simulation of marker-assisted selection utilizing linkage disequilibrium. *Theoretical and applied genetics*, v.83, p.813-820, 1992.

ZHU, J.; WEIR, B. S. Diallel analysis for sex-linked and maternal effects. *Theoretical and Applied Genetics*, v. 92, n.1, p.1-9, 1996.

WEI, R. Predicting genetic diversity and optimizing selection in breeding programmes. *Wei, R. P.*, v.1, p.64, 1995.

ANEXOS

ANEXO A

Página

TABELA 1A Valores médios para produção de espigas (t/ha) de 28 variedades de milho, ensaio conduzido por Ferreira (1993), em Sete Lagoas (MG), em duas repetições..... 69

TABELA 2A Valores médios para produção de espigas (t/ha) de 18 populações de milho, ensaio conduzido por Napolini Filho et al. (1981)..... 70

ANEXO B

Página

TABELA 1B Valores médios para produção de espigas (t/ha) de híbridos F_1 , obtidos de 28 variedades de milho, ensaio conduzido por Ferreira (1993), em Sete Lagoas (MG), em duas repetições..... 71

TABELA 2B Valores médios para produção de espigas (t/ha) de híbridos F_1 , obtidos de 18 populações de milho, ensaio conduzido por Napolini Filho et al. (1981) em Sete Lagoas (MG) e Piracicaba (SP), em duas repetições..... 72

ANEXO C

Página

Programas utilizados para simulação de uma das situações estudadas, linguagem SAS (*Statistical Analysis System*)..... 73

TABELA 1A. Resultados médios para produção de espigas (t/ha) de 28 variedades de milho em cruzamentos dialélicos (Ferreira, 1993), ensaio conduzido em Sete Lagoas (MG), em duas repetições

Variedades	Médias	Variedades	Médias
01. CMS - 01	5,58	16. CMS - 28	5,79
02. CMS - 02	5,56	17. CMS - 29	8,05
03. CMS - 03	6,59	18. CMS - 90	7,19
04. CMS - 04N	7,04	19. BR - 136	6,08
05. CMS - 04C	7,85	20. CMS - 39	8,36
06. BR - 105	8,18	21. CMS - 50	6,96
07. BR - 106	7,44	22. SINT. ELITE	7,44
08. BR - 107	5,70	23. PH - 4	4,95
09. BR - 111	7,18	24. CUNHA	7,61
10. BR - 112	6,11	25. BAIH-TUSON	3,37
11. CMS - 14C	5,27	26. SARACUSA	8,61
12. CMS - 15	7,04	27. NITROFLINT	6,08
13. CMS - 22	6,15	28. NITRODENT	7,55
14. CMS - 23	6,18	-	-
15. BR - 126	5,80	-	-

TABELA 2A. Resultados médios para peso de espigas (t/ha) de 18 populações de milho em cruzamentos dialélicos (Naspolini Filho et al, 1981), ensaio conduzido em Sete Lagoas (MG) e Piracicaba (SP), em duas repetições

Populações	Médias	Populações	Médias
01. CMS - 21	3,64	16. CMS - 06	5,23
02. CMS - 23	3,95	17. CMS - 07	4,06
03. CMS - 22	3,35	18. CMS - 10	3,67
04. CMS - 24	4,24	-	-
05. CMS - 16	2,57	-	-
06. CMS - 17	2,65	-	-
07. CMS - 05	4,13	-	-
08. CMS - 04	4,53	-	-
09. CMS - 03	4,27	-	-
10. CMS - 02	3,45	-	-
11. CMS - 01	3,01	-	-
12. CMS - 14	4,33	-	-
13. CMS - 12	4,07	-	-
14. CMS - 11	4,34	-	-
15. CMS - 15	3,81	-	-

TABELA 1B. Resultados médios para produção de espigas (t/ha) de híbridos obtidos de 28 variedades de milho (Ferreira, 1993), ensaio conduzido em Sete Lagoas (MG), em duas repetições

Pais	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	4,5	6,9	5,6	6,7	7	6	5,6	6,3	6,4	7,7	7,2	6,9	6,5	7,7
2		5,6	7,2	6,4	6,1	6,8	7,4	7,7	6,8	6,1	7,1	5,7	5,4	6,5
3			7,2	7,2	8,4	6,9	6,8	8,8	7,1	7,3	7,4	6,6	7	6,9
4				7,1	7,6	8,7	5,9	7,8	6,7	5,9	8	7,6	7	7,6
5					7,5	8,5	5,9	5,9	8,1	6,6	7,2	8,2	7,6	7,5
6						7	5,3	8	6,1	7,4	8,8	7,8	7,6	6,5
7							7,1	8,6	8,5	7,7	8,3	7,6	7,1	7,3
8								5,9	7,3	7,2	7,8	7,9	7,7	7,5
9									8,4	8	7	7,6	6,6	8,2
10										8,4	8,3	8	7,2	7,9
11											8,4	6,1	6,2	7,2
12												6,1	6,9	6,5
13													6,7	6,5
14														7,9
Pais	16	17	18	19	20	21	22	23	24	25	26	27	28	-
15	9	7	8,8	7,5	7,5	8,6	8,8	8,3	9	7,3	6,9	7,2	6,5	-
16		8,6	8,9	8,7	6,9	8,8	8	7,3	9	6,7	8,9	8,6	7,4	-
17			7,1	7,6	6,5	6,6	7,4	6,4	6,1	7,6	7	5,7	7,5	-
18				8,2	6,7	6,6	7,8	7	7,4	6,8	7,8	7,4	8,1	-
19					6,7	10	6,9	5,8	7,5	5,9	6,7	7,1	8,2	-
20						8	10	7,4	6,1	6,7	8,2	7,1	8,3	-
21							7,4	7,2	8,7	8,1	7	7,4	8,2	-
22								5,8	7,5	7,4	6,9	8,3	6,2	-
23									5,7	7	6,8	6,4	6,4	-
24										7,2	7,4	8,4	9	-
25											4,7	6,2	5,5	-
26												5,1	6,7	-
27													5,4	-


```

/*****
Programa: Simulação de dialelos circulantes
Módulo 1: Obtenção de valores fenotípicos, 2 repetições
Parâmetros: s/ dominância, herdabilidade: 10, 20, 50, 75
Ruben Delly Veiga – UFLA - 17/11/1998
Obs.: Permitido o uso para fins acadêmicos, conservando-se a autoria
*****/
options ps = 62 ls = 78 nodate nonumber;
title ;

/* Entrada de Matriz de linhagens */
data d;
infile 'c:\delly\simula1\linhas.txt';
input (gen001 - gen100) (100 * $2.);

proc iml;
use d;
read all var _char_ into genot ;

/* Valores Genotípicos das Linhagens */
v_linha = J(50,1,1000);
v_total = 0;

/* sem Dominância */
a = 2;
d = 0;
rep = 2;

do i = 1 to 50;
do j = 1 to 100;
if genot[i,j] = "11" then v_linha[i,1] = v_linha[i,1] + a;
if genot[i,j] = "00" then v_linha[i,1] = v_linha[i,1] - a;
end;
end;

nlin = 50*(50 + 1)/2;
f1 = J(nlin,1,1000);
pais = J(nlin,2,0);

ilin = 0;
do pai1 = 1 to 50;
do pai2 = pai1 to 50;
ilin = ilin + 1;
pais[ilin,1] = pai1;
pais[ilin,2] = pai2;
do j = 1 to 100;
if genot[pai1,j] = "11" then do;
if genot[pai2,j] = "11" then f1[ilin,1] = f1[ilin,1] + a;
if genot[pai2,j] = "00" then f1[ilin,1] = f1[ilin,1] + d;
end;
if genot[pai1,j] = "00" then do;

```

```

    if genot[pai2,j] = "11" then fl[i]in,1] = fl[i]in,1] + d;
    if genot[pai2,j] = "00" then fl[i]in,1] = fl[i]in,1] - a;
  end;
end;
end;
end;

```

```

/* Média de fl */
medf1 = 0;
do i = 1 to nlin;
  medf1 = medf1 + fl[i,1];
end;
medf1 = medf1 / nlin;

```

```

/* Variância de fl */
varf1 = 0;
do i = 1 to nlin;
  varf1 = varf1 + (fl[i,1] - medf1)*(fl[i,1] - medf1);
end;
varf1 = varf1 / nlin;

```

```

verr10 = rep * varf1*(1 - .1)/.1;
verr20 = rep * varf1*(1 - .2)/.2;
verr35 = rep * varf1*(1 - .35)/.35;
verr50 = rep * varf1*(1 - .5)/.5;
verr60 = rep * varf1*(1 - .6)/.6;
verr75 = rep * varf1*(1 - .75)/.75;
verr99 = rep * varf1*(1 - .99)/.99;

```

```

/* Arquivo herd 10% */
file 'c:\delly\simula00\simula10\ad\feno210.txt';
do i = 1 to nlin;
  p1 = pais[i,1];
  p2 = pais[i,2];
  do irep = 1 to rep;
    fenot = fl[i,1] + rannor(13)*sqrt(verr10);
    put p1 4. p2 4. i 6. irep 4. fenot 10.2;
  end;
end;

```

```

/* Arquivo herd 20% */
file 'c:\delly\simula00\simula20\ad\feno220.txt';
do i = 1 to nlin;
  p1 = pais[i,1];
  p2 = pais[i,2];
  do irep = 1 to rep;
    fenot = fl[i,1] + rannor(13)*sqrt(verr20);
    put p1 4. p2 4. i 6. irep 4. fenot 10.2;
  end;
end;

```

```

/* Arquivo herd 35 */
file 'c:\delly\simula00\simula35\ad\feno235.txt';
do i = 1 to nlin;
  p1 = pais[i,1];
  p2 = pais[i,2];
  do irep = 1 to rep;
    fenot = fl[i,1] + rannor(13)*sqrt(verr35);
    put p1 4. p2 4. i 6. irep 4. fenot 10.2;
  end;
end;

```

```

/* Arquivo herd 50% */
file 'c:\delly\simula00\simula50\ad\feno250.txt';
do i = 1 to nlin;
  p1 = pais[i,1];
  p2 = pais[i,2];
  do irep = 1 to rep;
    fenot = fl[i,1] + rannor(13)*sqrt(verr50);
    put p1 4. p2 4. i 6. irep 4. fenot 10.2;
  end;
end;

```

```

/* Arquivo herd 60% */
file 'c:\delly\simula00\simula60\ad\feno260.txt';
do i = 1 to nlin;
  p1 = pais[i,1];
  p2 = pais[i,2];
  do irep = 1 to rep;
    fenot = fl[i,1] + rannor(13)*sqrt(verr60);
    put p1 4. p2 4. i 6. irep 4. fenot 10.2;
  end;
end;

```

```

/* Arquivo herd 75% */
file 'c:\delly\simula00\simula75\ad\feno275.txt';
do i = 1 to nlin;
  p1 = pais[i,1];
  p2 = pais[i,2];
  do irep = 1 to rep;
    fenot = fl[i,1] + rannor(13)*sqrt(verr75);
    put p1 4. p2 4. i 6. irep 4. fenot 10.2;
  end;
end;

```

```

/* Arquivo herd 99% */
file 'c:\delly\simula00\simula99\ad\feno299.txt';
do i = 1 to nlin;
  p1 = pais[i,1];
  p2 = pais[i,2];
  do irep = 1 to rep;
    fenot = fl[i,1] + rannor(13)*sqrt(verr99);
  end;
end;

```

```

    put p1 4. p2 4. i 6. irep 4. fenot 10.2;
end;
end;

/* Arquivo médias herd 10% */
data m10;
infile 'c:\delly\simula00\simula10\ad\fenot210.txt';
input p1 p2 trat rep fenot ;
proc means data = m10 noprint;
var fenot;
output out = c10 mean = mean;
by p1 p2;
data _null_;
set c10;
file 'c:\delly\simula00\simula10\ad\fenm210.txt';
put p1 4. p2 4. mean 10.2;

/* Arquivo médias herd 20% */
data m20;
infile 'c:\delly\simula00\simula20\ad\fenot220.txt';
input p1 p2 trat rep fenot ;
proc means data = m20 noprint;
var fenot;
output out = c20 mean = mean;
by p1 p2;
data _null_;
set c20;
file 'c:\delly\simula00\simula20\ad\fenm220.txt';
put p1 4. p2 4. mean 10.2;

/* Arquivo médias herd 30% */
data m35;
infile 'c:\delly\simula00\simula35\ad\fenot235.txt';
input p1 p2 trat rep fenot ;
proc means data = m35 noprint;
var fenot;
output out = c35 mean = mean;
by p1 p2;
data _null_;
set c35;
file 'c:\delly\simula00\simula35\ad\fenm235.txt';
put p1 4. p2 4. mean 10.2;

/* Arquivo médias herd 50% */
data m50;
infile 'c:\delly\simula00\simula50\ad\fenot250.txt';
input p1 p2 trat rep fenot ;
proc means data = m50 noprint;
var fenot;
output out = c50 mean = mean;
by p1 p2;

```

```

data _null_;
  set c50;
  file 'c:\delly\simula00\simula50\ad\fenm250.txt';
  put p1 4. p2 4. mean 10.2;

/* Arquivo médias herd 60%      */
data m60;
  infile 'c:\delly\simula00\simula60\ad\feno260.txt';
  input p1 p2 trat rep fenot ;
  proc means data = m60 noprint;
  var fenot;
  output out = c60 mean = mean;
  by p1 p2;
data _null_;
  set c50;
  file 'c:\delly\simula00\simula60\ad\fenm260.txt';
  put p1 4. p2 4. mean 10.2;

/* Arquivo médias herd 75%      */
data m75;
  infile 'c:\delly\simula00\simula75\ad\feno275.txt';
  input p1 p2 trat rep fenot ;
  proc means data = m75 noprint;
  var fenot;
  output out = c75 mean = mean;
  by p1 p2;
data _null_;
  set c75;
  file 'c:\delly\simula00\simula75\ad\fenm275.txt';
  put p1 4. p2 4. mean 10.2;

/* Arquivo médias herd 99%      */
data m99;
  infile 'c:\delly\simula00\simula99\ad\feno299.txt';
  input p1 p2 trat rep fenot ;
  proc means data = m99 noprint;
  var fenot;
  output out = c99 mean = mean;
  by p1 p2;
data _null_;
  set c99;
  file 'c:\delly\simula00\simula99\ad\fenm299.txt';
  put p1 4. p2 4. mean 10.2;

*proc print data = c99;

run;
quit;
run;

```

```

/*****
Programa: Simulação de dialelos circulantes
Módulo 2: Aleatorização dos pais, 50 vezes
Parâmetros: s/ dominância, herdabilidade: 10, 20, 50, 75
Ruben Delly Veiga – UFLA - 17/11/1998
Obs.: Permitido o uso para fins acadêmicos, conservando-se a autoria
*****/
options ps = 64 nodate nonumber;
title ;
data d;
infile 'c:\delly\simula00\simula10\ad\fenm210.txt';
input x1-x3;

proc sort data = d; by x1 x2;
proc iml;
use d;
read all into mat;
nlin = nrow(mat);

do rep = 1 to 50;
pai1 = int(ranuni(rep)*50) + 1;
pai2 = int(ranuni(rep)*50) + 1;

pais = j(nlin,4,0);
do i = 1 to nlin;
pais[i,1] = mat[i,1];
pais[i,2] = mat[i,2];
pais[i,3] = mat[i,1];
pais[i,4] = mat[i,2];

if mat[i,1] = pai1 then pais[i,1] = pai2;
if mat[i,2] = pai1 then pais[i,2] = pai2;

if mat[i,1] = pai2 then pais[i,1] = pai1;
if mat[i,2] = pai2 then pais[i,2] = pai1;
end;

file 'c:\delly\simula00\simula10\ad\sihlrep2.txt';
do i = 1 to nlin;
pai1 = pais[i,1];
pai2 = pais[i,2];
paio1 = pais[i,3];
paio2 = pais[i,4];
y = mat[i,3];
put rep 5. pai1 5. pai2 5. y 10.4 paio1 4. paio2 4.;
end;

end;
quit;
run;

```

```

/*****
Programa: Simulação de dialelos circulantes
Módulo 3: Obtenção dos circulantes possíveis, estimativas de CGC e CEC
Parâmetros: s/ dominância, herdabilidade: 10, 20, 50, 75
Ruben Delly Veiga – UFLA - 17/11/1998
Obs.: Permitido o uso para fins acadêmicos, conservando-se a autoria
*****/

```

```

options ps = 64 nodate nonumber;
title ;

```

```

/* Entrada de Matriz de médias */

```

```

data d;
infile 'c:\delly\simula00\simula10\ad\sih1rep2.txt';
input rep pai1 pai2 prod;

```

```

data d (keep = rep pai1 pai2 prod);
set d;
if pai1 gt pai2 then do;
temp = pai1;
pai1 = pai2;
pai2 = temp;
end;

```

```

proc sort data = d;
by rep pai1 pai2;

```

```

proc iml;

```

```

/* Definição de matrizes do modelo */

```

```

use d;
read all into maty;
pais = 50;
nlin = pais*(pais-1)/2 + pais;

```

```

/* Definição de repetições */

```

```

do rep = 1 to 50;

```

```

inic = 1+(rep-1)*nlin;
ifim = inic + nlin - 1;
i = inic:ifim;
y = maty[i,4];
yq = j(pais,pais,0);
pvet = maty[i,2:3];

```

```

do i = 1 to nlin;
valor = y[i,1];
ind1 = pvet[i,1];
ind2 = pvet[i,2];
yq[ind1,ind2] = valor;
yq[ind2,ind1] = valor;

```

```

end;
/* CIRCULANTES */
do s = 3 to (pais - 1) by 2 ;
k = (pais + 1 - s)/2;
svet = j(pais,pais,0);
ycirc = j(pais,pais,0);
a = j(pais,pais,0);
do i = 1 to pais;
a[i,1] = s + 4;
svet[i,1] = 1;
ycirc[i,1] = yq[i,1];
do j = (k + 1) to (k + 1 - 1 + s);
jp = j;
if jp > pais then jp = j - pais;
svet[i,jp] = 1;
ycirc[i,jp] = yq[i,jp];
a[i,jp] = 1;
end; /* j */
end; /* i */
/* Média */
m = 0;
yip = j(pais,1,0);
do i = 1 to pais;
do j = 1 to pais;
yip[i,1] + ycirc[i,j];
end;
yip[i,1] = yip[i,1] + ycirc[i,1];
m = m + yip[i,1];
end;
m = m/(pais*(s+2));
/* O */
q = j(pais,1,0);
do i = 1 to pais;
q[i,1] = yip[i,1] - (s+2)*m;
end;
/* CG */
cg = inv(a)*q;
/* CE */
ice = 0;
nlincc = pais*s + pais;
ce = j(nlincc,1,0);
vce = j(nlincc,2,0);
do i = 1 to pais;

```

```
do j = 1 to pais;
```

```
  vce1 = svet[i,j];
```

```
  vce2 = svet[j,i];
```

```
  vces = vce1 + vce2;
```

```
  if vces < 0 then do;
```

```
    ice = ice + 1;
```

```
    ce[ice,1] = yq[i,j] - m - cg[i,1] - cg[j,1];
```

```
    vetce[ice,1] = i;
```

```
    vetce[ice,2] = j;
```

```
  end; /* if */
```

```
end; /* j */
```

```
end; /* i */
```

```
file 'c:\delly\simula00\simula10\ad\cgsh1rep.txt';
```

```
do i = 1 to pais;
```

```
  c = cg[i,1];
```

```
  pai1 = i;
```

```
  pai2 = i;
```

```
  put rep 5. s 4. pai1 4. pai2 4. c 10.3;
```

```
end;
```

```
file 'c:\delly\simula00\simula10\ad\cesh1rep.txt';
```

```
do i = 1 to nlnice;
```

```
  e = ce[i,1];
```

```
  pai1 = vetce[i,1];
```

```
  pai2 = vetce[i,2];
```

```
  if pai1 <= pai2 then do;
```

```
    put rep 5. s 4. pai1 4. pai2 4. e 10.3;
```

```
  end;
```

```
end;
```

```
end; /* s */
```

```
end; /* rep */
```

```
quit;
```

```
run; quit;
```

```

/*****
Programa: Simulação de dialelos circulantes
Módulo 4: Manutenção de pais originais e respectivas estimativas de CGC e CEC
Parâmetros: s/ dominância, herdabilidade: 10, 20, 50, 75
Ruben Delly Veiga – UFLA - 17/11/1998
Obs.: Permitido o uso para fins acadêmicos, conservando-se a autoria
*****/

```

```

options ps = 64 nodate nonumber;
title ;

```

```

data d2 (keep = rep pai1 pai2 pai1 pai2);
infile 'c:\delly\simula00\simula10\ad\sih1rep2.txt';
input rep pai1 pai2 prod pai1 pai2;

```

```

proc sort data = d2;
by rep pai1 pai2;

```

```

data cg;
infile 'c:\delly\simula00\simula10\ad\cgsh1rep.txt';
input rep 5. s 4. pai1 4. pai2 4. c 10.3;

```

```

proc sort data = cg;
by rep pai1 pai2;

```

```

data cgd2;
merge d2 cg;
by rep pai1 pai2;

```

```

data cgd2;
set cgd2 ;
if s eq . or c eq . then delete;

```

```

data ce;
infile 'c:\delly\simula00\simula10\ad\cesh1rep.txt';
input rep 5. s 4. pai1 4. pai2 4. e 10.3;

```

```

proc sort data = ce;
by rep pai1 pai2;

```

```

data ced2;
merge d2 ce;
by rep pai1 pai2;

```

```

data ced2;
set ced2 ;
if s eq . or e eq . or pai1 eq . or pai2 eq . then delete;

```

```

data _null_;
set cgd2;
file 'c:\delly\simula00\simula10\ad\cgsh1re2.txt';

```

```
put rep 5. s 4. pai1 4. pai2 4. c 10.3 pai01 4. paio2 4.;
```

```
data _null_;
```

```
set ced2;
```

```
file 'c:\delly\simula00\simula10\ad\cesh1re2.txt';
```

```
put rep 5. s 4. pai1 4. pai2 4. c 10.3 pai01 4. paio2 4.;
```

```
run;quit;
```

```

/*****
Programa: Simulação de dialetos circulantes
Módulo 5: Intervalos de Confiança para CGC e CEC e
          porcentagem de ocorrências
Parâmetros: s/ dominância, herdabilidade: 10, 20, 50, 75
Ruben Delly Veiga – UFLA - 17/11/1998
Obs.: Permitido o uso para fins acadêmicos, conservando-se a autoria
*****/

```

```

options ps = 56 nodate nonumber;
title ;

```

```

data cg;
  infile 'c:\delly\simula00\simula10\ad\cgsh1re2.txt';
  input rep 5. s 4. pai1 4. pai2 4. c 10.3 pai01 4. pai02 4.;

```

```

data cg26 (keep = pai01 pai02 icinf icsup);
  set cg;

```

```

v = ((50 - 1)/(50*(50 + 2))) * 192343.6/2;
erro = 1.96 * sqrt(v);

```

```

icinf = c - erro;
icsup = c + erro;

```

```

if icinf > icsup then do;
  temp = icinf;
  icinf = icsup;
  icsup = temp;
end;

```

```

if s = 49 and rep = 5 then output;

```

```

proc sort data = cg;
  by pai01 pai02;

```

```

proc sort data = cg26;
  by pai01 pai02;

```

```

data cg2;
  merge cg cg26;
  by pai01 pai02;

```

```

proc sort data = cg2;
  by s pai01 pai02 ;

```

```

data cgtab;
  set cg2;
  by s pai01 pai02 ;

```

```

if first.pai01 and first.pai02 then do;

```

```

nd = 0;
nb = 0;
na = 0;
percd = 0;
percb = 0;
perca = 0;
end;

if c <= icsup and c >= icinf then nd + 1;
if c < icinf then nb + 1;
if c > icsup then na + 1;

if last.paio1 and last.paio2 then do;
percd = (nd / 50) * 100;
percb = (nb / 50) * 100;
perca = (na / 50) * 100;
output cgtab;
end;

proc means data =cgtab noprint;
var percd ;
by s;
output out = c mean = mean;

proc print data = c;

data _null_;
set c;
file 'c:\delly\simula00\simula10\ad\cgsh1tab.txt';
put s 4. mean 7.2;

run; quit;

options ps = 64 nodate nonumber;
title ;

data ce;
infile 'c:\delly\simula00\simula10\ad\cesh1re2.txt';
input rep 5. s 4. pai1 4. pai2 4. e 10.3 paio1 4. paio2 4.;

data ce26 (keep = paio1 paio2 icinf icsup);
set ce;

vii = (50 * (50 - 1)/((50 + 1)*(50 + 2))) * 192343.6/2;
vij = (50 * 50 + 50 + 2)/((50 + 1)*(50 + 2)) * 192343.6/2;

if paio1 eq paio2 then erro = sqrt(vii);
if paio1 ne paio2 then erro = sqrt(vij);

icinf = e - erro;
icsup = e + erro;

```

```

if icinf > icsup then do;
    temp = icinf;
    icinf = icsup;
    icsup = temp;
end;

if s = 49 and rep = 5 then output;

proc sort data = ce;
by paio1 paio2;

proc sort data = ce26;
by paio1 paio2;

data ce2;
merge ce ce26;
by paio1 paio2;

proc sort data = ce2;
by s paio1 paio2 ;

data cetab;
set ce2;
by s paio1 paio2 ;

if /*first.paio1 and*/ first.paio2 then do;
    n = 0;
    nd = 0;
    nb = 0;
    na = 0;
    percd = 0;
    percb = 0;
    perca = 0;
end;

n + 1;

if e <= icsup and e >= icinf then nd + 1;
if e < icinf then nb + 1;
if e > icsup then na + 1;

if /*last.paio1 and*/ last.paio2 then do;
percd = (nd / n) * 100;
percb = (nb / n) * 100;
perca = (na / n) * 100;
output cetab;
end;

proc means data = cetab noprint;
var percd;

```

```
by s;  
output out = e mean = mean;  
  
proc print data = e;  
  
data _null_ ;  
set e;  
file 'c:\delly\simula00\simula10\ad\cesh1tab.txt';  
put s 4. mean 7.2;  
  
run; quit;
```

```

/*****
Programa: Simulação de dialelos circulantes
Módulo 6: Correlações entre estimativas de CGC e CEC em dialelos
completos e circulantes
Parâmetros: s/ dominância, herdabilidade: 10, 20, 50, 75
Ruben Delly Veiga – UFLA - 17/11/1998
Obs.: Permitido o uso para fins acadêmicos, conservando-se a autoria
*****/

```

```

options ps = 64 nodate nonumber;
title ;

```

```

data cg03 (keep = rep pai1 pai2 cg03 )
  cg05 (keep = rep pai1 pai2 cg05 )
  cg07 (keep = rep pai1 pai2 cg07 )
  cg09 (keep = rep pai1 pai2 cg09 )
  cg11 (keep = rep pai1 pai2 cg11 )
  cg13 (keep = rep pai1 pai2 cg13 )
  cg15 (keep = rep pai1 pai2 cg15 )
  cg17 (keep = rep pai1 pai2 cg17 )
  cg19 (keep = rep pai1 pai2 cg19 )
  cg21 (keep = rep pai1 pai2 cg21 )
  cg23 (keep = rep pai1 pai2 cg23 )
  cg25 (keep = rep pai1 pai2 cg25 )
  cg27 (keep = rep pai1 pai2 cg27 )
  cg29 (keep = rep pai1 pai2 cg29 )
  cg31 (keep = rep pai1 pai2 cg31 )
  cg33 (keep = rep pai1 pai2 cg33 )
  cg35 (keep = rep pai1 pai2 cg35 )
  cg37 (keep = rep pai1 pai2 cg37 )
  cg39 (keep = rep pai1 pai2 cg39 )
  cg41 (keep = rep pai1 pai2 cg41 )
  cg43 (keep = rep pai1 pai2 cg43 )
  cg45 (keep = rep pai1 pai2 cg45 )
  cg47 (keep = rep pai1 pai2 cg47 )
  cg49 (keep = rep pai1 pai2 cg49 ) ;

```

```

infile 'c:\delly\simula00\simula10\ad\cgsh1rep.txt';
input rep 5. s 4. pai1 4. pai2 4. c 10.3;
if s = 3 then do;
  cg03 = c;
  output cg03;
end;
if s = 5 then do;
  cg05 = c;
  output cg05;
end;
if s = 7 then do;
  cg07 = c;
  output cg07;
end;

```

```
if s = 9 then do;
  cg09 = c;
  output cg09;
end;
if s = 11 then do;
  cg11 = c;
  output cg11;
end;
if s = 13 then do;
  cg13 = c;
  output cg13;
end;
if s = 15 then do;
  cg15 = c;
  output cg15;
end;
if s = 17 then do;
  cg17 = c;
  output cg17;
end;
if s = 19 then do;
  cg19 = c;
  output cg19;
end;
if s = 21 then do;
  cg21 = c;
  output cg21;
end;
if s = 23 then do;
  cg23 = c;
  output cg23;
end;
if s = 25 then do;
  cg25 = c;
  output cg25;
end;
if s = 27 then do;
  cg27 = c;
  output cg27;
end;
if s = 29 then do;
  cg29 = c;
  output cg29;
end;
if s = 31 then do;
  cg31 = c;
  output cg31;
end;
if s = 33 then do;
  cg33 = c;
  output cg33;
```

```
end;
if s =35 then do;
  cg35= c;
  output cg35;
end;
if s =37 then do;
  cg37= c;
  output cg37;
end;
if s =39 then do;
  cg39= c;
  output cg39;
end;
if s =41 then do;
  cg41= c;
  output cg41;
end;
if s =43 then do;
  cg43= c;
  output cg43;
end;
if s =45 then do;
  cg45= c;
  output cg45;
end;
if s =47 then do;
  cg47= c;
  output cg47;
end;
if s =49 then do;
  cg49= c;
  output cg49;
end;
```

```
data ce03 (keep = rep pai1 pai2 ce03 )
  ce05 (keep = rep pai1 pai2 ce05 )
  ce07 (keep = rep pai1 pai2 ce07 )
  ce09 (keep = rep pai1 pai2 ce09 )
  ce11 (keep = rep pai1 pai2 ce11)
  ce13 (keep = rep pai1 pai2 ce13)
  ce15 (keep = rep pai1 pai2 ce15)
  ce17 (keep = rep pai1 pai2 ce17)
  ce19 (keep = rep pai1 pai2 ce19)
  ce21 (keep = rep pai1 pai2 ce21)
  ce23 (keep = rep pai1 pai2 ce23)
  ce25 (keep = rep pai1 pai2 ce25)
  ce27 (keep = rep pai1 pai2 ce27)
  ce29 (keep = rep pai1 pai2 ce29)
  ce31 (keep = rep pai1 pai2 ce31)
  ce33 (keep = rep pai1 pai2 ce33)
  ce35 (keep = rep pai1 pai2 ce35)
```

```
ce37 (keep = rep pai1 pai2 ce37)
ce39 (keep = rep pai1 pai2 ce39)
ce41 (keep = rep pai1 pai2 ce41)
ce43 (keep = rep pai1 pai2 ce43)
ce45 (keep = rep pai1 pai2 ce45)
ce47 (keep = rep pai1 pai2 ce47)
ce49 (keep = rep pai1 pai2 ce49) ;
```

```
infile 'c:\delly\simula00\simula10\ad\cesh1rep.txt';
input rep 5. s 4. pai1 4. pai2 4. e 10.3;
if s = 3 then do;
  ce03 = e;
  output ce03;
end;
if s = 5 then do;
  ce05 = e;
  output ce05;
end;
if s = 7 then do;
  ce07 = e;
  output ce07;
end;
if s = 9 then do;
  ce09 = e;
  output ce09;
end;
if s = 11 then do;
  ce11 = e;
  output ce11;
end;
if s = 13 then do;
  ce13 = e;
  output ce13;
end;
if s = 15 then do;
  ce15 = e;
  output ce15;
end;
if s = 17 then do;
  ce17 = e;
  output ce17;
end;
if s = 19 then do;
  ce19 = e;
  output ce19;
end;
if s = 21 then do;
  ce21 = e;
  output ce21;
end;
if s = 23 then do;
```

```
ce23= e;  
output ce23;  
end;  
if s =25 then do;  
ce25= e;  
output ce25;  
end;  
if s =27 then do;  
ce27= e;  
output ce27;  
end;  
if s =29 then do;  
ce29= e;  
output ce29;  
end;  
if s =31 then do;  
ce31= e;  
output ce31;  
end;  
if s =33 then do;  
ce33= e;  
output ce33;  
end;  
if s =35 then do;  
ce35= e;  
output ce35;  
end;  
if s =37 then do;  
ce37= e;  
output ce37;  
end;  
if s =39 then do;  
ce39= e;  
output ce39;  
end;  
if s =41 then do;  
ce41= e;  
output ce41;  
end;  
if s =43 then do;  
ce43= e;  
output ce43;  
end;  
if s =45 then do;  
ce45= e;  
output ce45;  
end;  
if s =47 then do;  
ce47= e;  
output ce47;  
end;
```

```
if s =49 then do;
  ce49= e;
  output ce49;
end;
```

```
proc sort data = cg03; by rep pai1 pai2;
proc sort data = cg05; by rep pai1 pai2;
proc sort data = cg07; by rep pai1 pai2;
proc sort data = cg09; by rep pai1 pai2;
proc sort data = cg11; by rep pai1 pai2;
proc sort data = cg13; by rep pai1 pai2;
proc sort data = cg15; by rep pai1 pai2;
proc sort data = cg17; by rep pai1 pai2;
proc sort data = cg19; by rep pai1 pai2;
proc sort data = cg21; by rep pai1 pai2;
proc sort data = cg23; by rep pai1 pai2;
proc sort data = cg25; by rep pai1 pai2;
proc sort data = cg27; by rep pai1 pai2;
proc sort data = cg29; by rep pai1 pai2;
proc sort data = cg31; by rep pai1 pai2;
proc sort data = cg33; by rep pai1 pai2;
proc sort data = cg35; by rep pai1 pai2;
proc sort data = cg37; by rep pai1 pai2;
proc sort data = cg39; by rep pai1 pai2;
proc sort data = cg41; by rep pai1 pai2;
proc sort data = cg43; by rep pai1 pai2;
proc sort data = cg45; by rep pai1 pai2;
proc sort data = cg47; by rep pai1 pai2;
proc sort data = cg49; by rep pai1 pai2;
```

```
proc sort data = ce03; by rep pai1 pai2;
proc sort data = ce05; by rep pai1 pai2;
proc sort data = ce07; by rep pai1 pai2;
proc sort data = ce09; by rep pai1 pai2;
proc sort data = ce11; by rep pai1 pai2;
proc sort data = ce13; by rep pai1 pai2;
proc sort data = ce15; by rep pai1 pai2;
proc sort data = ce17; by rep pai1 pai2;
proc sort data = ce19; by rep pai1 pai2;
proc sort data = ce21; by rep pai1 pai2;
proc sort data = ce23; by rep pai1 pai2;
proc sort data = ce25; by rep pai1 pai2;
proc sort data = ce27; by rep pai1 pai2;
proc sort data = ce29; by rep pai1 pai2;
proc sort data = ce31; by rep pai1 pai2;
proc sort data = ce33; by rep pai1 pai2;
proc sort data = ce35; by rep pai1 pai2;
proc sort data = ce37; by rep pai1 pai2;
proc sort data = ce39; by rep pai1 pai2;
proc sort data = ce41; by rep pai1 pai2;
proc sort data = ce43; by rep pai1 pai2;
```

```

proc sort data = ce45; by rep pai1 pai2;
proc sort data = ce47; by rep pai1 pai2;
proc sort data = ce49; by rep pai1 pai2;

data cg;
merge   cg03 cg05 cg07 cg09 cg11 cg13 cg15 cg17 cg19
        cg21 cg23 cg25 cg27 cg29 cg31 cg33 cg35 cg37 cg39
        cg41 cg43 cg45 cg47 cg49 ;
by rep pai1 pai2;

data ce;
merge   ce03 ce05 ce07 ce09 ce11 ce13 ce15 ce17 ce19
        ce21 ce23 ce25 ce27 ce29 ce31 ce33 ce35 ce37 ce39
        ce41 ce43 ce45 ce47 ce49 ;
by rep pai1 pai2;

proc corr data = cg nosimple noprint outs = spearcg ;
var cg49;
with   cg03 cg05 cg07 cg09 cg11 cg13 cg15 cg17 cg19
        cg21 cg23 cg25 cg27 cg29 cg31 cg33 cg35 cg37 cg39
        cg41 cg43 cg45 cg47   ;

title4 'Capacidade Geral - Correlações';
by rep;

proc corr data = ce nosimple noprint outs = spearce ;
var ce49;
with   ce03 ce05 ce07 ce09 ce11 ce13 ce15 ce17 ce19
        ce21 ce23 ce25 ce27 ce29 ce31 ce33 ce35 ce37 ce39
        ce41 ce43 ce45 ce47   ;
title4 'Capacidade Especifica - Correlações';
by rep;

proc sort data = spearcg; by _name_;
proc sort data = spearce; by _name_;

data cs;
set spearcg;
if _name_ ne " ";

proc means data = cs noprint;
var cg49;
by _name_;
output out = cgs mean = med min = min max = max n = n;

data es;
set spearce;
if _name_ ne " ";

proc means data = es noprint ;
var ce49;

```

```

by _name_;
output out = ces mean = med min = min max = max n = n;

proc print data = cgs;
proc print data = ces;

data _null_;
  set cgs;
  file 'c:\dely\simula00\simula10\ad\sh1\cgs.txt';
  minsig = 1 - probf((min*min*(n-2)/(1-min*min)), 1, (n-2));
  medsig = 1 - probf((med*med*(n-2)/(1-med*med)), 1, (n-2));
  maxsig = 1 - probf((max*max*(n-2)/(1-max*max)), 1, (n-2));
  put _name_ 1-5 min 10.7 minsig 10.7 med 10.7 medsig 10.7 max 10.7 maxsig 10.7;

data _null_;
  set ces;
  file 'c:\dely\simula00\simula10\ad\sh1\ces.txt';
  minsig = 1 - probf((min*min*(n-2)/(1-min*min)), 1, (n-2));
  medsig = 1 - probf((med*med*(n-2)/(1-med*med)), 1, (n-2));
  maxsig = 1 - probf((max*max*(n-2)/(1-max*max)), 1, (n-2));
  put _name_ 1-5 min 10.7 minsig 10.7 med 10.7 medsig 10.7 max 10.7 maxsig 10.7;

run;quit;

```