

Estratégias de amostragem e estabelecimento de coleções nucleares

Edmar Soares de Vasconcelos⁽¹⁾, Cosme Damião Cruz⁽¹⁾, Leonardo Lopes Bhering⁽¹⁾ e Adésio Ferreira⁽¹⁾

⁽¹⁾Universidade Federal de Viçosa, Av. P.H. Rolfs, s/nº, CEP 36571-000 Viçosa, MG. E-mail: vasconceloses@vicosa.ufv.br, cdcruz@ufv.br, leobhering@yahoo.com.br, adesio@vicosa.ufv.br

Resumo – O objetivo deste trabalho foi avaliar a influência da intensidade de amostragem, do tamanho da coleção de germoplasma inicial e da variância da amostragem sobre a qualidade das respectivas coleções nucleares, quanto à representatividade das coleções iniciais. Foram simulados sete tamanhos de coleções iniciais e utilizadas seis intensidades de amostragem para estabelecimento de coleções nucleares, utilizando caracteres morfoagronômicos. Determinaram-se o número de grupos formados, o coeficiente de coincidência entre a coleção nuclear e a coleção inicial e o coeficiente de determinação dos acessos amostrados para comporem a coleção nuclear. Também foi proposto o uso de uma estratégia alternativa para estabelecer coleções nucleares, de forma a maximizar a diversidade entre os acessos. O tamanho da coleção inicial influencia a intensidade de amostragem empregada na obtenção da coleção nuclear. A amostragem de acessos pelo método de Tocher, com critério de aglomeração inverso, mostrou-se eficiente na obtenção de coleções nucleares. As diferentes magnitudes de variância das coleções iniciais não influenciaram os coeficientes de determinação (R^2) nem os coeficientes de coincidência entre a coleção inicial e as respectivas coleções nucleares.

Termos para indexação: coleção de germoplasma, coleção nuclear, intensidade de amostragem, coeficiente de coincidência, método de Tocher.

Strategies for sampling and establishment of core collections

Abstract – The objective of this work was to evaluate the effect of sampling intensity, initial collection's size and its variance over the quality of the respective core collections, in terms of representation of the initial collections. Therefore, simulations were performed for seven sizes of the initial collections and six sampling intensities for the establishment of core collections were used, comprising morpho-agronomic traits. Groups formed were determined, as well as the coincidence coefficient between the core and the initial collection, and the determination coefficient of the accessions sampling to compose the core collection. The use of an alternative strategy was proposed to establish core collections, this strategy maximizes the diversity among the accessions. The size of initial collections affects the sampling intensity to be used to obtain a core collection. The selection of accessions with the method of Tocher, comprising criteria of inverse grouping, was efficient for the formation of the core collections. The different magnitudes of variance of the initial collections did not affect either the coefficient of determination (R^2) or the coincidence rate between initial collections and respective core collection.

Index terms: germplasm collection, core collection, sampling intensity, coincidence rate, Tocher method.

Introdução

Recursos genéticos vegetais se perdem em virtude da utilização inadequada e da destruição de habitats. Porém, nos últimos anos, observa-se notável progresso na obtenção e conservação de recursos genéticos (Hintum et al., 2000). Cerca de 6,1 milhões de acessos de plantas são conservados em todo o mundo e esses acessos encontram-se dispostos em 1.320 bancos de germoplasma. Esses bancos encontram-se distribuídos pelos 157 países que compõem a Comissão de Recursos Fitogenéticos da FAO. No Brasil, existem mais de 200 mil acessos de plantas conservados

ex situ, em cerca de 160 bancos de germoplasma implantados em mais de 50 locais.

Na realização dessa tarefa o desafio é o incremento do uso deste germoplasma conservado, pois, atualmente, apenas cerca de 4% de todos os genótipos armazenados vêm sendo utilizados no mundo. A fim de aumentar o uso deste germoplasma, é necessário priorizar a caracterização, avaliação, documentação e informação sobre esses acessos. Um dos maiores problemas para a implementação da caracterização, avaliação, documentação e informação do germoplasma é o excessivo tamanho das coleções de germoplasma, o que

difícil as tarefas de organização e manejo. Grandes coleções de germoplasma apresentam baixa qualidade, considerando que não se tem acesso às informações de variabilidade para características de importância econômica. Tem-se, por fim, a baixa utilização dos seus acessos de germoplasma.

A fim de representar a variabilidade em uma coleção de germoplasma com número reduzido de acessos, vem sendo elaborada o que convencionalmente se denomina de coleção nuclear. Uma coleção nuclear pode ser definida por um grupo limitado de acessos derivados de uma coleção de germoplasma, escolhido para representar o espectro genético da coleção inteira (Frankel & Brown, 1984; Brown, 1989). Tradicionalmente, as coleções nucleares são estabelecidas com tamanho em torno de 10% dos acessos de toda a coleção original e incluem ($p = 95\%$) aproximadamente 70% da riqueza alélica (alelos comuns e raros) disponível no acervo gênico inicial (Brown & Spillane, 1999).

Grande importância tem sido dada, atualmente, à obtenção de coleções nucleares, existindo coleções para arroz (Abadie et al., 2005), batata (Chandra et al., 2002), cevada (Liu et al., 1999), entre outras.

A seleção de genótipos, para a obtenção de uma coleção nuclear, utiliza estratégias que envolvem diversas intensidades e métodos de amostragem (Thompson, 2002). A intensidade de amostragem refere-se à porcentagem de acessos da coleção de germoplasma que serão selecionados para compor a coleção nuclear. Alguns autores estudaram intensidades de amostragem que corresponderam, normalmente, entre 5 a 20% do total do número de acessos (Brown, 1989; Brown & Spillane, 1999; Van Hintun et al., 2000). Upadhyaya & Ortiz (2001) utilizaram baixa intensidade de amostragem, em que apenas 1% dos acessos da coleção inicial de grão-de-bico foi suficiente para compor a coleção nuclear. Idealmente, a coleção nuclear deveria ser criada com base na caracterização e avaliação completa dos materiais, porém é geralmente criada com base em informações morfogeográficas, validadas por análise genética (Rodrigues et al., 2005).

Os métodos de alocação possuem critérios para determinar o número de acessos a serem selecionados em cada grupo. Esses grupos são caracterizados por apresentarem genótipos com características similares, que são utilizadas para diferenciar os acessos de uma coleção inicial. Os métodos de alocação procuram estabelecer o número e tamanho dos grupos, além dos acessos que irão compor os grupos, de acordo com a similaridade desses acessos.

Franco et al. (2005) avaliaram diferentes estratégias de determinação do número de indivíduos de cada grupo. Os métodos avaliados por Franco et al. (2005) foram baseados na intensidade de amostragem, no tamanho do grupo e na diversidade existente entre os acessos, em relação à coleção inicial como um todo. Porém, praticamente, não há estudos relacionando o tamanho da coleção inicial e sua variabilidade com o número de grupos obtidos pela distância entre os acessos.

O objetivo deste trabalho foi avaliar o efeito que os seguintes aspectos exercem sobre a eficácia com a qual as coleções nucleares representam as coleções de germoplasma iniciais: tamanho da coleção de germoplasma inicial, variância de suas características constituintes e intensidades de amostragem utilizadas no processo de obtenção das coleções nucleares. O trabalho também propõe o uso de uma estratégia alternativa para estabelecer coleções nucleares, que maximize a diversidade entre os acessos.

Material e Métodos

Foram simulados sete bancos de germoplasmas (ou coleções iniciais) com tamanhos de 250, 500, 1.000, 1.500, 2.000, 2.500 e 3.000 acessos. Os acessos dos bancos de germoplasmas foram descritos em relação a 21 características morfoagronômicas (Tabela 1), cujos

Tabela 1. Médias e variâncias das 21 características simuladas nas coleções iniciais.

Característica	Média	Variância
Entrelaçamento	0,6600	0,0049
Hábito de crescimento	6,8200	0,0025
Diâmetro do internódio da rama	3,8000	0,0016
Comprimento dos internódios da rama	2,8300	0,0036
Cor predominante da rama	4,6100	0,0064
Cor da rama secundária	2,0000	0,0081
Esboço geral da folha	4,5600	0,0025
Tipo de lóbulo da folha	3,7000	0,0100
Número de lóbulos da folha	3,7200	0,0036
Forma do lóbulo da folha central	2,9700	0,0049
Tamanho da folha madura	5,2500	0,0009
Pigmentação das nervuras abaxiais da folha	4,9400	0,0064
Cor da folha madura	3,0300	0,0036
Cor da folha imatura	3,4100	0,0036
Pigmentação do pecíolo	5,1000	0,0121
Comprimento do pecíolo	4,2900	0,0025
Cor predominante da película externa da raiz	13,1600	0,0900
Cor secundária da película externa da raiz	1,5700	0,0121
Cor predominante da raiz	3,7600	0,0036
Cor secundária da raiz	2,6500	0,0144
Distribuição da cor secundária da raiz	1,4800	0,0081

parâmetros de média e variância foram simulados e fixados, independentemente dos tamanhos das coleções iniciais, com base nos dados do estudo de Huaman et al. (1999), no qual foram avaliados acessos de batata-doce (*Ipomoea batatas* L.).

A fim de verificar a influência das diferentes intensidades de amostragem em coleções iniciais de diferentes tamanhos sobre a eficácia de representação das coleções nucleares resultantes, foram aplicadas as intensidades de amostragens de 1, 5, 10, 20, 40 e 80% do número de acessos das diferentes coleções iniciais com diferentes tamanhos iniciais. Desta forma, obtiveram-se 42 diferentes coleções nucleares: seis diferentes intensidades de amostragem e sete diferentes tamanhos de coleção inicial.

Para avaliar o efeito de diferentes variâncias de coleção de germoplasma inicial sobre a eficácia da representação de suas respectivas coleções nucleares, obtidas com diferentes intensidades de amostragens, foram simuladas outras seis coleções de germoplasma iniciais com 1.000 acessos cada, diferindo entre si, em magnitude de variância das características morfológicas dos seus acessos. Nessas simulações, partiu-se dos valores de médias e variâncias do estudo de Huaman et al. (1999), e uma das coleções simuladas tinha exatamente os mesmos valores

de médias e variâncias que este estudo (coleção 3, Tabela 2). As cinco demais coleções de germoplasma apresentaram valores de variância diferentes e fixos em relação às variâncias da coleção 3, dentro das seguintes faixas: 1/100, 1/50, 50, 100 e 200 vezes os valores das variâncias da coleção 3 (Tabela 2).

Com base nos dados de caracterização disponíveis (simulados), foram calculadas distâncias euclidianas entre os acessos de germoplasma das coleções iniciais. A partir dessas distâncias, foram realizados os agrupamentos dos acessos para avaliar o efeito do tamanho da coleção inicial sobre o número de grupos formados pelo método de agrupamento de Tocher, conforme descrito em Cruz & Carneiro (2003).

Quanto à seleção dos acessos componentes das coleções nucleares, utilizou-se a técnica de agrupamento de Tocher (Rao, 1952), porém com uma modificação, que consiste em agrupar acessos com maior dissimilaridade, ou seja, método de Tocher com critério de aglomeração inverso. Por meio dessa alteração, os acessos selecionados formam um único grupo, que corresponde à coleção nuclear.

Pelo método proposto, encontra-se o par de indivíduos mais dissimilares e esses formaram o grupo inicial. Para a inclusão de um outro acesso, são calculadas as

Tabela 2. Médias e variâncias das 21 características simuladas para seis populações de 1.000 acessos com mesma média e diferentes variâncias.

Características	Média	Var 1 ^{a(1)}	Var 2 ^{a(2)}	Var 3 ^{a(3)}	Var 4 ^{a(4)}	Var 5 ^{a(5)}	Var 6 ^{a(6)}
Entrelaçamento	0,6600	0,00005	0,00010	0,00490	0,24500	0,49000	0,98000
Hábito de crescimento	6,8200	0,00003	0,00005	0,00250	0,12500	0,25000	0,50000
Diâmetro do internódio da rama	3,8000	0,00002	0,00003	0,00160	0,08000	0,16000	0,32000
Comprimento dos internódios da rama	2,8300	0,00004	0,00007	0,00360	0,18000	0,36000	0,72000
Cor predominante da rama	4,6100	0,00006	0,00013	0,00640	0,32000	0,64000	1,28000
Cor da rama secundária	2,0000	0,00008	0,00016	0,00810	0,40500	0,81000	1,62000
Esboço geral da folha	4,5600	0,00003	0,00005	0,00250	0,12500	0,25000	0,50000
Tipo de lóbulo da folha	3,7000	0,00010	0,00020	0,01000	0,50000	1,00000	2,00000
Número de lóbulos da folha	3,7200	0,00004	0,00007	0,00360	0,18000	0,36000	0,72000
Forma do lóbulo da folha central	2,9700	0,00005	0,00010	0,00490	0,24500	0,49000	0,98000
Tamanho da folha madura	5,2500	0,00001	0,00002	0,00090	0,04500	0,09000	0,18000
Pigmentação das nervuras abaxiais da folha	4,9400	0,00006	0,00013	0,00640	0,32000	0,64000	1,28000
Cor da folha madura	3,0300	0,00004	0,00007	0,00360	0,18000	0,36000	0,72000
Cor da folha imatura	3,4100	0,00004	0,00007	0,00360	0,18000	0,36000	0,72000
Pigmentação do pecíolo	5,1000	0,00012	0,00024	0,01210	0,60500	1,21000	2,42000
Comprimento do pecíolo	4,2900	0,00003	0,00005	0,00250	0,12500	0,25000	0,50000
Cor predominante da película externa da raiz	13,1600	0,00090	0,00180	0,09000	4,50000	9,00000	18,00000
Cor secundária da película externa da raiz	1,5700	0,00012	0,00024	0,01210	0,60500	1,21000	2,42000
Cor predominante da raiz	3,7600	0,00004	0,00007	0,00360	0,18000	0,36000	0,72000
Cor secundária da raiz	2,6500	0,00014	0,00029	0,01440	0,72000	1,44000	2,88000
Distribuição da cor secundária da raiz	1,4800	0,00008	0,00016	0,00810	0,40500	0,81000	1,62000

(¹)Variância = (Var 3^a)/100. (²)Variância = (Var 3^a)/50. (³)Variância = (Var 3^a). (⁴)Variância = (Var 3^a) x 50. (⁵)Variância = (Var 3^a) x 100.

(⁶)Variância = (Var 3^a) x 200.

distâncias entre o grupo inicial e os demais genótipos: o acesso que apresentar maior distância desse grupo, será o genótipo incluído. Assim se mantém a regra, até que se obtenha o número de acessos necessários para a formação da coleção nuclear.

Depois da obtenção da coleção nuclear, estabeleceu-se a determinação do coeficiente de similaridade entre essa coleção e sua respectiva coleção inicial. De acordo com Hu et al. (2000), o coeficiente de similaridade pode ser obtido por:

$$CR\% = \left(\frac{1}{m} \sum_{j=1}^m \frac{R_c}{R_I} \right) 100$$

em que: CR% é o coeficiente de similaridade entre a coleção inicial e a coleção nuclear; m é o número de características avaliadas; R_c é a diferença entre o número máximo e o mínimo de cada variável observado na coleção nuclear; R_I é a diferença entre o número máximo e o mínimo de cada variável observado na coleção inicial.

Um outro critério utilizado para verificar a eficácia de representatividade das coleções nucleares foi o coeficiente de determinação (R^2), que foi determinado pela razão entre a soma de quadrados dos desvios, entre acessos selecionados para compor a coleção nuclear e a soma de quadrados total dos desvios da coleção de germoplasma (Cruz, 1997). Os acessos foram selecionados pelo método de Tocher com critério de aglomeração inverso.

Ainda com o intuito de verificar a eficácia das coleções nucleares em representar as coleções de germoplasma, empregou-se um teste de homogeneidade entre as variâncias (teste F) e um teste entre as médias (teste t) da coleção inicial e das respectivas coleções nucleares. Isso foi realizado com a finalidade de verificar diferenças nas variáveis avaliadas (simuladas) entre coleções nucleares e coleções iniciais.

As simulações das coleções iniciais, a obtenção das distâncias euclidianas, a formação dos grupos, amostragem dos acessos pelo método de Tocher com critério de aglomeração inverso e a comparação de médias e variâncias entre as coleções foram efetuadas com auxílio do aplicativo computacional GENES (Cruz, 1997).

Resultados e Discussão

Não se observaram diferenças significativas entre as médias das coleções iniciais e suas respectivas coleções,

coleções de germoplasma simuladas com diferentes números de acessos, obtidas por diferentes intensidades de amostragem, pelo teste t. Também não se verificaram diferenças significativas entre as variâncias das coleções de germoplasma e das respectivas coleções nucleares pelo teste F.

Os valores dos coeficientes de determinação (R^2) obtidos com a estratégia de amostragem do método de Tocher com critério de aglomeração inverso foram afetados pela intensidade de amostragem empregada (Figura 1). Observou-se que, quanto maior a intensidade de amostragem, maior os valores de R^2 obtidos. Os valores desse coeficiente variaram de 3,7 a 15,5%, quando se empregou a intensidade de amostragem de 1% dos acessos de cada banco de germoplasma, e variaram entre 86,7 a 88,7% na intensidade de amostragem de 80%. O coeficiente (R^2) apresentou, nesta última intensidade de amostragem, valores muito superiores aos obtidos com o emprego de 1% de intensidade de amostragem. Esse resultado era esperado, pois quanto maior o número de acessos da coleção nuclear, maior será a relação entre a soma de quadrado dos desvios dos acessos amostrados e a soma de quadrado total dos desvios.

Em intensidades de amostragem de 1, 5 e 10%, verificou-se comportamento logarítmico dos valores de

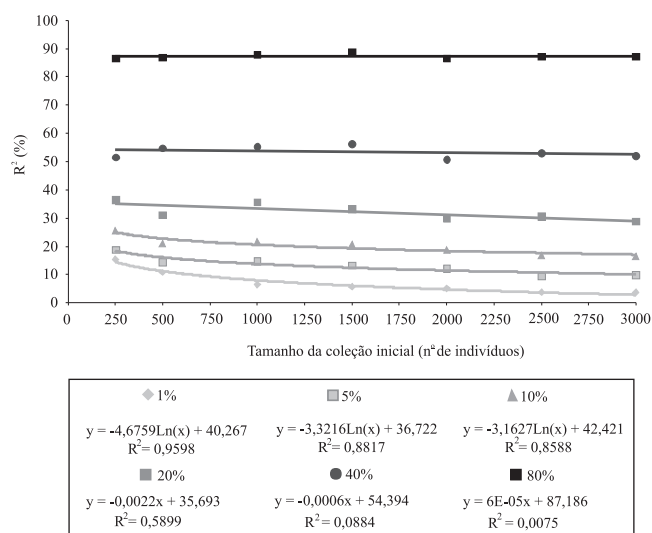


Figura 1. Valores de R^2 calculados pelo método de seleção de Tocher com critério de aglomeração inverso em função do tamanho da coleção inicial, para diferentes intensidades de amostragem.

R^2 de acordo com diferentes tamanhos de coleções de germoplasma iniciais. Tal comportamento, provavelmente, se deva ao fato do reduzido número de indivíduos selecionados para coleções nucleares a partir das coleções iniciais menores, uma vez que, nas coleções iniciais de 250 indivíduos, foram selecionados apenas 3 e 13 acessos na intensidade de amostragem de 1 e 5%, respectivamente. Na população inicial de 500 indivíduos foram selecionados apenas 5 e 25 acessos, nas intensidade de 1 e 5%, respectivamente. Com relação à intensidade de amostragem de 10%, observou-se esse desvio logarítmico causado pelo baixo número de acessos amostrados, apenas na coleção inicial menor, com tamanho de 250 acessos.

Outro fator, concomitante, que pode explicar esse comportamento logarítmico dos valores de R^2 em função dos diferentes tamanhos das coleções iniciais é o maior valor da soma de quadrados de tratamentos das menores coleções iniciais. Considerando que os valores de média e variância de cada coleção inicial (avaliadas) são fixos, a heterogeneidade entre os acessos da coleção inicial tende a ser maior quanto menor for seu número de acessos. O mesmo é verificado para a coleção nuclear.

Pelos resultados dos coeficientes de coincidência obtidos entre as coleções iniciais e suas respectivas coleções nucleares, observou-se que, na intensidade de amostragem de 1% apenas nas coleções iniciais de 2.500 e de 3.000 acessos, os coeficientes de coincidência foram superiores a 80% (Figura 2). O limite de 80% para o coeficiente de coincidência é considerado por

Hu et al. (2000) como o limite mínimo para que a respectiva coleção nuclear seja considerada representativa da coleção inicial.

Quando se empregou uma intensidade de amostragem de 5%, ou maior, observou-se que, em geral, as coleções nucleares resultaram em coeficientes de coincidência maiores que 80%, com exceção da obtida para a menor das coleções iniciais, 250 acessos (Figura 2).

Em intensidades de amostragens iguais ou maiores que 5% e, para coleções nucleares obtidas das maiores coleções iniciais (com número de acessos maior que 1.500), os coeficientes de coincidência foram superiores a 90%. Para essas coleções iniciais maiores, uma intensidade de amostragem igual ou superior a 5% foi suficiente para que o coeficiente de coincidência fosse superior a 90%. Entretanto, em coleções iniciais menores, com número de acessos inferior a 1.500, observou-se variação nos valores dos coeficientes de coincidência entre diferentes coleções nucleares e coleções iniciais, para intensidade de amostragem menores ou iguais a 10% (Figura 3).

Ao se considerar o limite mínimo de 80% para o coeficiente de coincidência entre coleção nuclear e coleção inicial (Hu et al., 2000), verifica-se a necessidade de selecionar um número próximo a 10% dos acessos da coleção inicial pelo método de Tocher com critério de aglomeração inverso, no caso de coleção inicial composta por aproximadamente 250 acessos.

Numa coleção inicial com número de acessos entre 500 e 1.000, é necessária uma seleção de apenas 5%

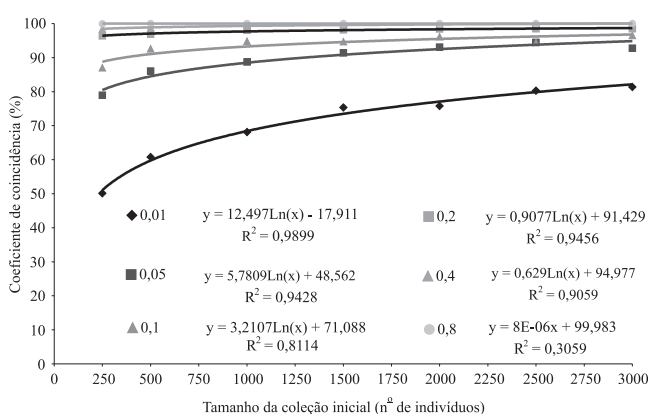


Figura 2. Coeficiente de coincidência entre coleções nucleares e respectivas coleções iniciais em função de diferentes tamanhos de coleções iniciais, em diferentes intensidades de amostragens.

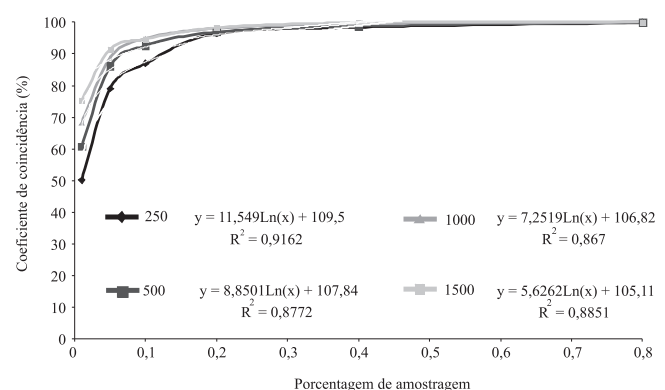


Figura 3. Coeficientes de coincidência entre coleções iniciais e coleções nucleares, obtidas pelo método de Tocher com critério de aglomeração inverso, em função da intensidade de amostragem, em coleções iniciais de diferentes tamanhos.

para compor sua respectiva coleção nuclear, que apresente aceitáveis coeficientes de coincidência. Para coleções iniciais ainda maiores, com 1.500 a 2.000 acessos, a intensidade de amostragem de 5% resulta em coleções nucleares com coeficiente de coincidência em torno de 90%. Nessas coleções iniciais maiores (com 1.500 e 2.000 acessos), até mesmo a intensidade de amostragem de 1% fornece coleções nucleares com valores do coeficiente de coincidência aceitáveis, o que concorda com o trabalho de Upadhyaya & Ortiz (2001), cuja coleção nuclear correspondeu a 1% dos acessos da coleção inicial.

Os resultados apresentados na Figura 4 indicam ausência do efeito de variabilidade das coleções iniciais (diferentes variâncias das características avaliadas) sobre os valores de R^2 obtidos entre essas coleções iniciais e suas respectivas coleções nucleares. Tal comportamento é observado em todas as diferentes intensidades de amostragem testadas, quando se utiliza a amostragem pelo método de Tocher com critério de aglomeração inverso. Os resultados confirmam que a intensidade de amostragem empregada exerce influência nos valores de R^2 obtidos, quando se realiza a seleção de acessos pelo método de Tocher com critério de aglomeração inverso (Figuras 1 e 4).

O coeficiente de coincidência não foi influenciado pelas diferentes variabilidades das coleções iniciais

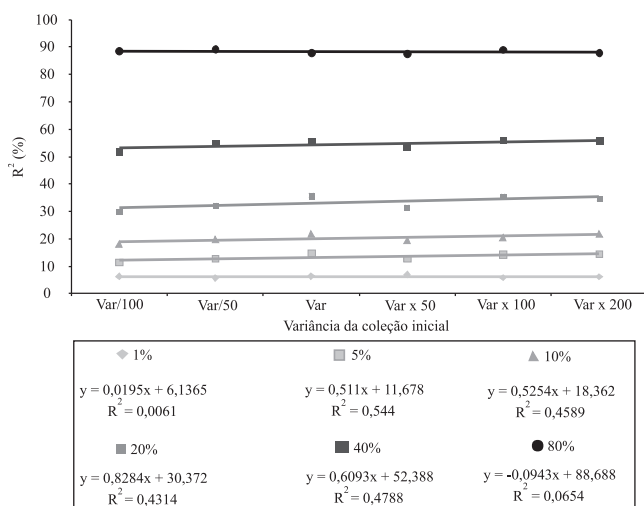


Figura 4. Valores de R^2 calculados pelo método de seleção de Tocher com critério aglomerativo inverso, em função de diferentes magnitudes de variância das coleções iniciais, em diferentes intensidades de amostragem.

(Figura 5). Dessa forma, diferentes coleções iniciais, em termos de suas variâncias, produzem, em geral, coleções nucleares com as mesmas qualidades (valores de R^2 e coeficientes de coincidência) quando obtidas pelo método de Tocher com critério de seleção inverso.

O número de grupos formados pelo uso do método de agrupamento de Tocher, considerando as distâncias euclidianas, apresentou relação linear com o tamanho da coleção inicial de germoplasma (Figura 6 A). Tal resultado era esperado, pois o método de agrupamento empregado baseia-se na formação dos grupos pela distância existente entre os acessos. Ao se fixar os valores de média e de variância das coleções iniciais, a homogeneidade entre os acessos passou a aumentar com o aumento do tamanho dessas coleções. Assim os acessos passaram a possuir menores distâncias entre si nas coleções iniciais maiores. Como a formação dos grupos respeita mais o critério de menores distâncias dentro dos grupos do que entre os grupos, o número de grupos formados aumentou, acompanhando, o aumento da homogeneidade dos acessos, o que explica a influência do tamanho da coleção inicial sobre o número de grupos formados a partir do método de agrupamento de Tocher.

Os procedimentos utilizados para formação de coleções nucleares, em sua maioria, fazem uso do número de grupos formados, e, por isso, quanto maior o número de grupos, maior, também, será o número de acessos para representar todos os grupos da

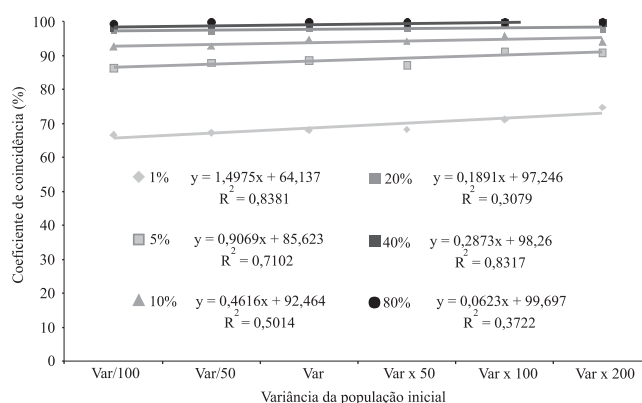


Figura 5. Coeficiente de coincidência entre coleções iniciais e coleções nucleares, obtidas pelo método de Tocher com critério de aglomeração inverso, em função de diferentes magnitudes de variância das coleções iniciais, em diferentes intensidades de amostragem.

coleção inicial em uma coleção nuclear. O processo de amostragem pelo método de Tocher com critério de aglomeração inverso não faz uso do número de grupos formados, e, por isso, esse número de grupos não influencia o tamanho da coleção nuclear, obtida pelo método de Tocher com critério de aglomeração inverso.

Não houve efeito de diferentes variâncias de características das coleções iniciais sobre o número de grupos obtidos pelo agrupamento de Tocher, utilizando-se distâncias euclidianas entre os acessos (Figura 6 B). Isso evidencia que, quando se utiliza o método de Tocher para estabelecer o número de grupos (dentro do processo de obtenção de uma coleção nuclear), a variabilidade da característica avaliada, nos acessos do banco de germoplasma, não influencia o número de grupos formados.

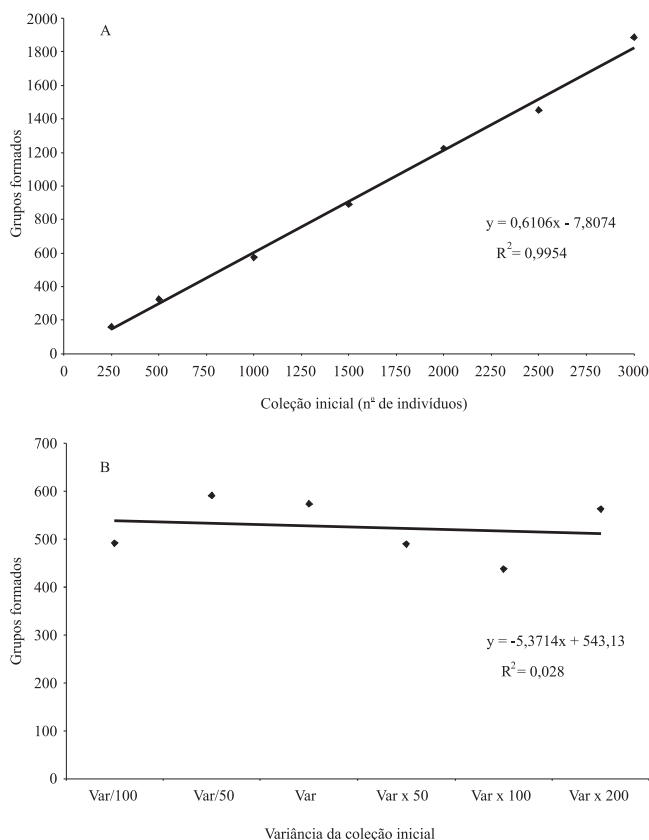


Figura 6. Grupos formados em função dos diferentes tamanhos de coleções iniciais (A) e em função de diferentes magnitudes de variâncias das coleções iniciais (B).

A intensidade de amostragem, pelo método de Tocher com critério de aglomeração inverso, igual ou superior a 5%, em coleções iniciais entre 250 e 3.000 acessos, independentemente da variabilidade das características morfofisiológicas, é suficiente para representar a coleção inicial, pelos seguintes parâmetros: comparações de médias e variâncias, do coeficiente de determinação e do coeficiente de coincidência entre coleção inicial e coleção nuclear.

Esta intensidade de amostragem (5%) fica próxima à empregada por Abadie et al. (2005) para a formação de uma coleção nuclear de arroz (5,6%). Contudo, é inferior à apresentada por Brown (1989), que corresponde a 10% dos acessos da coleção inicial. Sendo menor também à utilizada por Huaman et al. (2000), que formaram uma coleção nuclear de batatas com intensidade de amostragem de 13%. Chandra et al. (2002) afirmaram que essa coleção inicial de batatas seria representada apenas por 25% dos seus acessos.

Conclusões

1. Coleções de germoplasma iniciais com diferentes tamanhos possuem distintas intensidades de amostragem mínimas (pelo método de Tocher com critério de agrupamento inverso), para gerar coleções nucleares adequadas em suas representatividades.

2. Não existe diferença entre as médias e as variâncias das coleções iniciais (simuladas) e suas respectivas coleções nucleares, obtidas pelo método de Tocher com critério de aglomeração inverso, quando a variação do tamanho da coleção inicial está entre 250 e 3.000 acessos.

3. O método de Tocher com critério de agrupamento inverso é eficiente na obtenção de coleção nuclear representativa da coleção inicial, em relação a recuperação da média, variância e coeficiente de determinação, gerando valores do coeficiente de coincidência superior a 80%.

4. Não existe efeito das magnitudes de variâncias das coleções iniciais sobre a qualidade das coleções nucleares em representar as respectivas coleções iniciais (qualidade avaliadas pelos valores de R^2 e pelos coeficientes de coincidência), fixando-se as intensidades de amostragem e os tamanhos de coleções iniciais.

Referências

- ABADIE, T.; CORDEIRO, C.M.T.; FONSECA, J.R.; ALVES, R. de B. das N.; BURLE, M.L.; BRONDANI, C.; RANGEL, P.H.N.; CASTRO, E. da M. de; SILVA, H.T. da; FREIRE, M.S.; ZIMMERMANN, F.J.P.; MAGALHÃES, J.R. Construção de uma coleção nuclear de arroz para o Brasil. **Pesquisa Agropecuária Brasileira**, v.40, p.129-136, 2005.
- BROWN, A.H.D. The case for core collections. In: BROWN, A.H.D.; FRANKEL, O.H.; MARSHALL, D.R.; WILLIAMS, J.T. (Ed.). **The use of plant genetic resources**. New York: Cambridge University Press, 1989. p.136-156.
- BROWN, A.H.D.; SPILLANE, C. Implementing core collections principles procedures, progress, problems and promise. In: JOHNSON, R.C.; HODGKIN, T. (Ed.). **Core collections for today and tomorrow**. Rome: IPGRI, 1999. p.1-9.
- CHANDRA, S.; HUAMÁN, Z.; KRISHNA, S.H.; ORTIZ, R. Optimal sampling strategy and core collection size of Andean tetraploid potato based on isozyme data - a simulation study. **Theoretical and Applied Genetics**, v.104, p.1325-1334, 2002.
- CRUZ, C.D. **Programa Genes**: aplicativo computacional em genética e estatística. Viçosa: UFV, 1997. 442p.
- CRUZ, C.D.; CARNEIRO, P.C.S. **Modelos biométricos aplicados ao melhoramento genético**. 2.ed. Viçosa: UFV, 2003. v.2, 585p.
- FRANCO, J.; CROSSA, J.; TABA, S.; SHANDS, H. A sampling strategy for conserving genetic diversity when forming core subsets. **Crop Science**, v.45, p.1035-1044, 2005.
- FRANKEL, O.H.; BROWN, A.H.D. Plant genetic resources today: a critical appraisal. In: HOLDEN, J.H.W.; WILLIAMS, J.T. (Ed.). **Crop genetic resources: conservation & evaluation**. London: Allen & Unwin, 1984. p.249-257.
- HINTUM, T.J.L. van; BROWN, A.H.D.; SPILLANE, C.; HODGKIN, T. **Core collections of plant genetic resources**. Rome: International Plant Genetic Resources Institute, 2000. 48p. (IPGRI. Technical bulletin, 3).
- HU, J.; ZHU, J.; XU, H.M. Methods of constructing core collections by stepwise clustering with three sampling strategies based on the genotypic values of crops. **Theoretical and Applied Genetics**, v.101, p.264-268, 2000.
- HUAMÁN, Z.; AGUILAR, C.; ORTIZ, R. Selecting a Peruvian sweetpotato core collection on the basis of morphological, eco-geographical and disease and pest reaction data. **Theoretical and Applied Genetics**, v.98, p.840-844, 1999.
- HUAMÁN, Z.; ORTIZ, R.; ZHANG, D.; RODRÍGUEZ, F. Isozyme analysis of entire and core collections of *Solanum tuberosum* subsp. *andigena* potato cultivars. **Crop Science**, v.40, p.273-276, 2000.
- LIU, F.; BOTHMER, R. von; SALOMON, B. Genetic diversity among East Asian accessions of the barley core collection as revealed by six isozyme loci. **Theoretical and Applied Genetics**, v.98, p.1226-1233, 1999.
- RAO, C.R. **Advanced statistical methods in biometric research**. New York: Wiley, 1952. 389p.
- RODRIGUES, D.P.; ASTOLFI FILHO, S.; CLEMENT, C.R. Molecular marker-mediated validation of morphologically defined landraces of Pejibaye (*Bactris gasipaes*) and their phylogenetic relationships. **Genetic Resources and Crop Evolution**, v.51, p.871-882, 2005.
- THOMPSON, S.K. **Sampling**. 2nd ed. New York: Wiley, 2002. 367p.
- UPADHYAYA, H.D.; ORTIZ, R. A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. **Theoretical and Applied Genetics**, v.102, p.1292-1298, 2001.

Recebido em 28 de julho de 2006 e aprovado em 19 de fevereiro de 2007