

DIVERSIDADE GENÉTICA E HERANÇA DE CARACTERÍSTICAS FÍSICAS DE FRUTOS E SEMENTES DO CACAUEIRO

José Luis Pires, Wilson Reis Monteiro

Comissão Executiva do Plano da Lavoura Cacaueira, Centro de Pesquisa do Cacau, Ilhéus, Bahia, Brazil.
joseluiszoegapires@gmail.com; wrmwr2@gmail.com

Para o cacau, as características: peso da semente, peso de sementes por fruto, peso total de fruto e proporção de mucilagem são importantes para o melhoramento genético, pois são fatores de produção e estão relacionados à distribuição de fotoassimilados. Além disso, são relevantes devido às exigências da indústria e ao impacto que exercem sobre os processos de colheita e pós-colheita. Dessa forma, 536 acessos preservados na Coleção de Germoplasma do Centro de Pesquisa do Cacau, em Ilhéus, Bahia, foram avaliados quanto a essas características. Para investigar a importância da avaliação de germoplasma para estes atributos, também foram avaliados dez ensaios de campo para avaliação de progênies, abrangendo dezenas de combinações. Os resultados mostraram que a maioria dos grupos considerados promissores para fornecer genes de resistência à vassoura de bruxa e à podridão parda, bem como para o aumento do teor de gordura, apresentam frutos e sementes pequenas. Os caracteres analisados podem ser usados como elementos de caracterização de acessos, e permitem uma distinção clara entre séries de diferentes origens. Foi possível identificar diferenças na capacidade geral de combinação para características físicas dos frutos e sementes, mesmo entre grupos de clones com alta uniformidade. Também foram observados efeitos significativos da capacidade específica de combinação para o peso de uma semente e para o peso de sementes por fruto. Por fim, as distinções entre genótipos *per se* mostraram-se eficientes como indicadores de diferenças na capacidade geral de combinação para o peso de sementes e o peso de sementes por fruto, evidenciando a importância dessa avaliação na seleção de progenitores e mostrando a possibilidade de melhoramento genético para estes caracteres.

Palavras-chave: *Theobroma cacao*, melhoramento genético, características físicas de fruto e semente.

Genetic diversity and inheritance of physical characteristics of fruits and seeds of the cocoa tree.

For cocoa, the characteristics: seed weight, seed weight per fruit, total fruit weight, and mucilage proportion are crucial for genetic improvement, as they are production factors and are associated with the distribution of photoassimilates. These traits are also important due to industry requirements and their impact on harvesting and post-harvest processes. To assess these characteristics, 536 accessions for the germplasm collection of the Cocoa Research Center in Ilhéus, Bahia, were evaluated. Additionally, ten field trials for progeny evaluation, including dozens of combinations, were analyzed to investigate the importance of germplasm assessment for these traits. The study revealed several key findings: most of the groups identified as potential sources of genes for resistance to Witch's broom and Black pod diseases, as well as for increased fat content, have small fruits and seeds. The evaluated traits serve as valuable descriptors for characterizing accessions and effectively distinguishing different origin series. Furthermore, differences in general combining ability for fruit and seed physical traits can be detected even among highly uniform clone groups. Significant effects of specific combining ability were also observed for the weight of a seed and for the weight of seeds per fruit. Additionally, differences between genotypes *per se* serve as effective indicators of differences in general combining ability for seed weight and seed weight per fruit, reinforcing the importance of these evaluations for selecting progenitors and showing the possibility of genetic improvement for these characters

Key words: *Theobroma cacao*, genetic improvement, physical characteristics of fruit and seed.

Introdução

Algumas características de frutos e sementes do cacaueiro são importantes para o melhoramento, por constituírem fatores de produção, por serem relativos à distribuição de fotoassimilados e aproveitamento de produto, por exigências da indústria ou por facilitarem ou dificultarem processos de colheita e práticas de pós colheita, como é o caso do peso da semente, do peso da semente por fruto, do peso total de fruto e da proporção de mucilagem.

Estes fatores têm, tradicionalmente, sido considerados como caracterizadores (Engels et al., 1980; ICGD, 2024) e, para tal, a obtenção de seus valores é minuciosa, demandando grandes esforços, o que resulta na impossibilidade de se considerar um grande número de acessos a cada período de tomada de dados. Nesta estrutura, se houver efeito de época na determinação dos caracteres, a comparação de genótipos só será apropriada entre aqueles caracterizados nos mesmos períodos, e a comparação entre um grande número de genótipos é inviabilizada. Assim, são abordados, aqui, procedimentos de avaliação destes caracteres, conduzidos de forma simplificada, de modo que um grande número de genótipos possa ser considerado simultaneamente e possam ser adequadamente comparados.

Também, não foram, até então, explicitadas definições claras sobre as formas de se conduzir processos de melhoramento genético para estes caracteres, e assim, com o objetivo de estruturar elementos básicos sobre este tema, foi estudada a importância da realização da avaliação de germoplasma para a seleção de genótipos para a introdução em programas de melhoramento, ou a eficiência da seleção de clones a serem progenitores em processos de seleção recorrente, pela comparação do desempenho ‘per se’ de acessos com suas capacidades gerais de combinação, analisada a relação entre estes caracteres, para a seleção conjunta, estudada a distribuição destas características na espécie, para a definição de regiões mais promissoras para prospecções, identificados genótipos de destaque, para a introdução nos programas de melhoramento, analisada a possibilidade de se utilizar estes caracteres avaliados como caracterizadores, e estudados os efeitos da capacidade geral e específica de combinação entre genótipos.

Materiais e Métodos

Material

Coleção de germoplasma

Foram avaliados 536 acessos, preservados na coleção de germoplasma do CEPEC. Esta coleção, que possui aproximadamente 1.300 acessos, tem os de introdução mais antiga (grupo a que pertencem os considerados neste estudo) cultivados em fileiras de 10 plantas por acesso, em espaçamento de 3 x 3 metros, com os acessos distribuídos paralelamente, formando blocos com aproximadamente 60 elementos, que se sucedem, intercalados por ruas para manejo e escoamento.

Os caracteres considerados no processo de avaliação foram o peso de uma semente úmida (PSU), o peso de uma semente seca (PSS), o peso de sementes úmidas por fruto (PSUF), o peso total de fruto (PF), a relação semente seca/semente úmida ($RSSSU = PSS/PSU$) e o peso de semente seca por fruto ($PSSF = PSUF \times RSSSU$) (pesos em gramas).

A avaliação das características consideradas foi realizada a partir de amostras obtidas de plantas individuais, em três períodos de safra (final do ano) e um de temporão (meio do ano). A amostragem consistiu-se da colheita de todos os frutos maduros disponíveis de cada uma de 5 plantas tomadas ao acaso nas fileiras e marcadas para a avaliação, pesagem dos frutos inteiros (todos juntos) e pesagem do total de sementes dos frutos colhidos. Deste total de sementes de cada planta foram tomadas 40 e pesadas, em conjunto, ainda com a mucilagem. Essa mucilagem foi removida para a secagem da semente, em estufa a 105 graus centígrados por 24 horas. As sementes, agora com teor de umidade de aproximadamente 0% foram, novamente, pesadas e os pesos unitários obtidos pela divisão do valor observado pelo número de sementes da amostra.

Ensaio de avaliação de progênes híbridas

Para o estudo da importância da avaliação de germoplasma em respeito à seleção de genótipos para a introdução em programas de melhoramento (avaliação das relações entre o desempenho “per se” de clones e sua capacidade geral de combinação) foram analisados, para as variáveis PSU, PSS e RSSSU, dois ensaios de avaliação de cultivares

híbridas, denominados no CEPEC de ensaios 4 e 32, ambos cultivados em delineamento em blocos ao acaso e estruturados como fatoriais (pais e mães), com 5 repetições por progênie e 16 plantas por parcela (Monteiro et al, 1995). Do ensaio 4, foram consideradas as combinações entre os clones DR 2, ICS 1 e UF 613 com os clones SIC 2, 4, 17, 23 e, ainda, dos clones ICS 1 e UF 613 com os clones SIC 5, 19, 20, 21, 22, 24, 813 e 823, e do ensaio 32 as combinações de Ca 4, Ma 15, ICS 1 e UF 613 com IMC 67, Pound 12 e Pa 150.

Foram tomadas amostras de 30 sementes de cada parcela em duas épocas para o ensaio 4 (safra e temporão) e uma época para o ensaio 32. Cada semente foi pesada individualmente, inicialmente com mucilagem, em seguida e ainda úmida, sem mucilagem com e sem testa, e por fim, novamente sem testa, após secagem a 105 graus centígrados por 24 horas. As 30 testas secas foram pesadas conjuntamente. Considerou-se como PSU o peso da semente sem testa e a RSSSU foi obtida a partir da relação PSS/PSU, enquanto que na avaliação de germoplasma o PSS incluía o peso de testa. O peso proporcional da testa, no entanto, tem pouca variação entre progênies (a proporção do peso de testa para o de semente completa seca variou de 4,75% a 6,93% nos dois ensaios - Pires et al 1995, 1995a) de modo que as comparações entre os valores obtidos na avaliação de germoplasma e na de progênies não sofre limitações expressivas.

A relação entre o desempenho “per se” de clones e o comportamento destes como progenitores também foi avaliada para a variável PSUF. Para tal, foram considerados dados de peso de sementes úmidas por planta e número de frutos sadios, que geraram o PSUF, agora obtido de todos os frutos colhidos por árvore, em colheitas mensais, nos ensaios 1, 33, 34, 35, 37, 40, 41 e 42 (para o ensaio 1 o PSUF foi obtido a partir da relação entre os totais anuais de parcela para PSU e NFS; para os ensaios 33, 34, 35 e 37 a partir dos totais anuais por planta e para o conjunto 40, 41 e 42 a partir dos totais por planta/colheita). Estes experimentos foram conduzidos no CEPEC e dispostos em blocos ao acaso com 5 repetições de 16 plantas (ensaios 1, 33, 34, 35, 40, 41 e 42) ou 8 de 10 plantas (ensaio 37) e foram instalados, entre os anos 1975 e 1983 (Monteiro et al, 1995).

O ensaio 1 envolve todas a combinações entre os clones CC 41, CEPEC 1, SIAL 169 e SIC 19; o ensaio

33 é um fatorial que tem como mães os clones IMC 67, Pound 1 e Scavina 6 e como pais o CEPEC 11, 12 e 14; o ensaio 34 é formado pelas mães SIAL 169, 325, 327, 88, SIC 22, 23, 24 e 329 e pelos pais ICS 1 e PA 150; o 35 pelas mães ICS 1, 6, IMC 67, UF 613 e 667 e pais SIAL 169, SIC 250, 831 e 891; o 37 pelas mães ICS 6, SPA 17, UF 296 e pais EEG 65, SIAL 244 e SIC 18; e os ensaios 40, 41 e 42 são constituídos, conjuntamente, pelos cruzamentos entre as mães AB 1, Be 2, 3, 4 e 5, Ca 2, Ma 12 e 14, RB 40 e 48 e os pais ICS 1, 8, SIAL 325 e SIC 19.

Foram consideradas as médias anuais de parcela para o ensaio 1, as médias anuais por planta para o 33, 34, 35 e 37 e o valor por colheita/planta para os ensaios 40, 41, 42 (6 colheitas), sendo analisados dados dos anos 1986/87 e 89/90 para o ensaio 1, 82 a 88 para o 33, 85 a 87 para o 34, 84 a 88 para o 35, 85 a 87 para o 37 e 89 para o 40, 41 e 42.

Análises

Qualidade dos dados da avaliação de germoplasma

A avaliação das características consideradas foi realizada a partir de amostras obtidas de plantas que não estão casualizadas em campo e, para dimensionar a magnitude das possíveis distorções produzidas por esta situação, foram realizadas análises independentes para um grupo de 21 genótipos, que são representados na coleção por mais de uma fileira. Estas análises foram efetuadas tomando-se, ao acaso, ora uma ora outra de duas fileiras de cada genótipo considerando-se como fontes de variação o período de amostragem, o acesso e a interação acesso x período de amostragem; e a partir destas foram observadas as correlações entre as médias ajustadas de fileiras de genótipos repetidos, e os coeficientes de determinação relativos a repetibilidade das médias conforme os métodos de análise de variância, componente principal e análise estrutural (Cruz e Regazzi, 1977; Cruz, 1997).

Distinções entre acessos e grupos de acessos

As distribuições de frequência das variáveis consideradas foram analisadas através do uso do Teste de Kolmogorov-Smirnov (PROC UNIVARIATE - SAS Institute, 1988) e a homogeneidade de variâncias através do teste de Bartlett (Martin, 1995). As diferenças entre acessos foram estudadas através de

análises de variância, segundo modelo em que foram considerados como fontes de variação o período de amostragem, os acessos e a interação entre eles, todos admitidos como fixos (PROC GLM - SAS Institute, 1988), e as médias ajustadas foram obtidas reduzindo-se o modelo para efeitos de período e acessos, de modo a tornar todas elas e todos os contrastes entre acessos e entre períodos estimáveis. Para a análise das diferenças entre grupos de genótipos, grupos estes constituídos conforme relação em respeito à origem (de acordo com ICGD, 2024) (Tabela 1), o modelo usado contemplou como fontes de variação a época de amostragem, o grupo de acessos, os acessos dentro de grupo, a interação entre época e grupo e a interação entre época e acessos dentro de grupo, todos considerados como fixos (PROC GLM - SAS Institute, 1988).

Nas duas situações, cada variável foi, também, analisada como medidas repetidas em modelo multivariado (cada qual seccionada conforme as 4 épocas e analisadas como 4 variáveis usando PROC GLM – MANOVA, para a interpretação dos efeitos de acesso ou grupo e acesso e PROC GLM – REPEATED, para a consideração dos efeitos de época e interações).

O estudo dos grupos incluiu, ainda, a consideração dos caracteres simultaneamente, através de análise multivariada (PROC GLM/MANOVA - SAS Institute, 1988) e análise de variáveis canônicas (PROC CANDISC – SAS Institute, 1988), conduzidas a partir das médias ajustadas obtidas nas análises univariadas, com os grupos como fonte de variação e os diversos clones dentro de grupos como repetições destes.

Relações entre o desempenho “per se” de clones e a capacidade geral de combinação

Para a análise desta relação foram correlacionados valores obtidos da avaliação ‘per se’ de clones na coleção de germoplasma e médias de progenitores observados em ensaios de avaliação (já explicitados).

O ensaio 4 foi analisado tendo-se como fontes de variação a época, o bloco, o grupo 1 (clones: DR2, ICS3 e UF 613), o grupo 2 (progenitores da série SIC), a interação grupo 1 x grupo 2, semente dentro de grupo 1 x grupo 2, e as interações: época x grupo 1, época x grupo 2, época x grupo 1 x grupo 2 e época x semente dentro de grupo 1 x grupo 2 (PROC GLM – SAS

Tabela 1 - Relação dos grupos estudados

Procedência	Grupo	Descrição
Amazônia brasileira	ABR	Materiais da Amazônia brasileira - séries: Ob, CAB, Mocorongo, IQ, SE, AB, AM
Bahia	BAH	Seleções em variedades da Bahia e Espírito Santo - séries: SIC, SIAL, EEG, CEPEC
Pará	BE	Série BE
Amazonas	CA	Série CA
Pará- Est. exp.	CAS	Série CAS
Costa Rica	CC	Série CC
Equador - híbridos	CCN CEQ	Série CCN Série CEPEC - família de origem desconhecida
Bahia - híbridos	CHC	Série CEPEC – híbridos de Criollo
Amapá	CJ	Série CJ
Acre	CSU	Série Cruzeiro do Sul
Bahia - híbridos	CTF	Série CEPEC - híbridos Trinitário x Forastero
Equador	EET	Série EET
Granada	GS	Série GS
Trinidade	ICS	Série ICS
Peru - Amazônia	IMC	Série IMC
Amazonas	MA	Série MA
México	MEX	Série P
Equador	MOQ	Série Moq
Peru - Amazônia	NA	Série NA
Venezuela	OC	Série OC
Peru - Amazônia	PA	Série PA
Peru - Amazônia	PER	Genótipos do Peru - séries: Amazon, Mo
Peru - Amazônia	POU	Série Pound
Acre	RB	Série RB
México	RIM	Série RIM
Colômbia	SC	Série SC
Peru - Amazônia	SCA	Série Scavina
Varias	SDE	Genótipos descendentes do clone Scavina 6 - Séries: TSA, TSH, CEPEC, EET
Guatemala	SGU	Série Sgu
Colômbia	SPA	Série SPA
Costa Rica	UF	Série UF
Venezuela	VE	Genótipos da Venezuela - séries: Chuao, Choroni, PQTO, Playa Alta.

System –SAS Institute 1988); e as médias ajustadas de progenitores foram obtidas com a redução do modelo, pela desconsideração das interações e efeito de planta. O ensaio 32 foi analisado em modelo com blocos, grupo 1 (Ca 4, Ma 15, ICS 1 e UF 613), grupo 2 (IMC 67, Pound 12 e Pa 150) e interação grupo 1 x grupo 2 como fontes de variação.

Os ensaios 33, 34, 35 e 37, cujos tratamentos considerados formam estruturas fatoriais, foram analisados tendo como fontes de variação o ano, bloco, mãe, pai, interação mãe x pai, planta dentro de mãe x pai e as interações ano x pai, ano x mãe, ano x pai x mãe e ano x planta dentro de pai x mãe (PROC GLM – SAS System – SAS Institute, 1988) e as médias ajustadas de tratamentos (mãe x pai), os quadrado médios e números de graus de liberdade dos resíduos, e as médias harmônicas para os números de repetições que geraram as médias de tratamentos, foram usadas para análises de dialelos pelo modelo de Griffing (CRUZ, 1997). Analogamente, o ensaio 1, que é um dialelo desbalanceado completo e inclui auto-fecundações e recíprocos, foi analisado, primariamente, tendo como fontes de variação o ano, bloco, pai, mãe e interações ano x pai, ano x mãe, ano x pai x mãe; e, em seguida, conforme metodologia de Griffing (Cruz e Regazzi 1997; Cruz, 1997).

Os ensaios 40 41 e 42 formam, conjuntamente, um esquema fatorial e têm duas testemunhas comuns não participantes deste fatorial. Estes ensaios foram analisados em conjunto, tendo como fontes de variação o ensaio, bloco dentro de ensaio, tratamento, planta dentro do tratamento e as interações ensaio x tratamentos comuns e ensaio x planta dentro de tratamentos comuns. A partir desta análise foram obtidos o quadrado médio e o grau de liberdade do resíduo, a média harmônica do número de repetições para tratamento, e as médias ajustadas de tratamento (obtidas com a redução do modelo pela desconsideração das interações) (PROC GLM – SAS System – SAS Institute, 1988), que formaram os elementos para a análise de dialelo parcial pelo modelo de Griffing (Cruz, 1997). Neste caso, para a comparação entre médias de progenitores e médias de clones “per se”, as médias dos progenitores foram obtidas diretamente a partir das médias ajustadas de tratamentos.

Resultados e Discussão

Qualidade dos dados da avaliação de germoplasma

Para dimensionar a magnitude das possíveis distorções produzidas pela não casualização das plantas avaliadas, foram realizadas análises independentes para um grupo de 21 genótipos (séries: Be, Ca, CEPEC, Cruzeiro do Sul, Iq, Mocarongo, Pa, Pound, RB, RIM, SIC e SPA), que são representados na coleção por mais de uma fileira. As análises foram efetuadas tomando-se, ao acaso, ora uma ora outra de duas fileiras de cada genótipo, nos moldes utilizados para o estudo do conjunto geral de acessos (fontes de variação: períodos de amostragem, acesso e interação acesso x período de amostragem). Em ambas as situações, com o primeiro ou segundo e distinto grupo de 21 fileiras representantes dos 21 acessos, foram encontradas diferenças significativas entre os genótipos ao nível de 0,01% de probabilidade de erro para esta afirmação, para todos os caracteres considerados: peso de uma semente úmida (PSU), peso de uma semente seca (PSS), peso de sementes úmidas por fruto (PSUF), peso total de fruto (PF), relação semente seca/semente úmida ($RSSSU = PSS/PSU$) e peso de semente seca por fruto ($PSSF = PSUF \times RSSSU$) (para as variáveis PSU, PSS, PSUF e PSSF os valores foram transformados através da obtenção de suas raízes quadradas, para aproximar as curvas de distribuição de frequências de curvas de distribuição normal, e este aspecto será discutido com detalhes, posteriormente, nas análises das diferenças entre acessos e grupos de acessos).

As correlações entre as médias de fileiras ajustadas para efeitos do período de amostragem (correção efetuada devido ao número diferente de medições e obtida com a redução do modelo para duas fontes de variação: período e genótipo - LSMEANS – SAS Institute, 1988), para cada uma das situações, foram (Métodos de Spearman e Pearson, respectivamente): 0,98 e 0,99 para o PSU; 0,97 e 0,99 para o PSS; 0,94 e 0,95 para PSUF e 0,95 e 0,95 para o PF, todas significativas a 0,01% de probabilidade de erro, e as variáveis PSU e PSS alcançaram, com uma única fileira, coeficientes de determinação (R^2), relativos á repetibilidade, entre 0,95 e 0,99 conforme os métodos de análise de variância, componente principal e análise

estrutural (Cruz e Regazzi, 1977; Cruz, 1997), enquanto as variáveis PSUF e PF apresentam R^2 entre 0,90 e 0,99 para avaliações de uma única fileira. Estes valores elevados mostram ampla concordância entre médias das diferentes fileiras dos mesmos genótipos, reduzida troca de posição hierárquica entre as médias e confiabilidade satisfatória na análise das diferenças entre materiais.

A variável mais importante dentre as constituídas a partir de relações entre as anteriores, o peso de semente seca por fruto, preservou elevados coeficientes de correlação: 0,92 e 0,95 para os métodos de Spearman e Pearson, respectivamente, e a avaliação de uma única fileira resultou em coeficientes de determinação entre 0,90 e 0,99. Já a variável RSSU mostrou correlações de 0,87 e 0,88 (Spearman e Pearson, respectivamente) e um R^2 para a avaliação de uma única fileira entre 0,85 e 0,90. Assim, esta deve ser interpretada com maior cautela, mesmo tendo em conta que as distinções mais acentuadas entre suas médias resultam não somente de possíveis maiores efeitos de posição, mas, também, da eficiência em sua medição. A razão é que há grandes e rápidas alterações no teor de umidade da polpa durante a fase final de amadurecimento do fruto, de modo a gerar grandes diferenças entre frutos colhidos em uma mesma época, conforme seu grau de amadurecimento. Esta menor eficiência de medição é constatada por menores valores de R^2 (relação entre a soma de quadrados do modelo e a soma de quadrados totais) e, portanto, resíduos, proporcionalmente, mais amplos: $R^2 = 0,64$ e $0,69$ para RSSU nas duas análises, enquanto PMSU e PMSS tiveram R^2 superiores a 0,94, PSF e PF superiores a 0,76 e PSSF superior a 0,74.

Resíduos proporcionalmente maiores geram diferenças mínimas significativas proporcionalmente mais amplas, de modo que a aceitação da diferença entre médias requer distinções maiores e isto se contrapõe aos baixos coeficientes de correlação e repetibilidade observados para a RSSSU, ampliando a confiabilidade na interpretação das distinções (maior variação dentro de fileira servindo como contrabalanço dos possíveis maiores efeitos de posição).

A variável PSSF inclui na sua formação a RSSSU, mas preserva elevada repetibilidade por ser determinada principalmente pelo PSUF: coeficientes de correlação e determinação iguais a, respectivamente, 0,94 e 0,88; enquanto que a RSSSU explica apenas

0,07 da variação de PSSF. Estes valores já são provenientes do conjunto geral de acessos avaliados (médias ajustadas), e a determinação do PSUF na variável PSSF é relacionada à elevada correlação existente entre as variáveis PSU e PSS, diretamente mensuradas: $r = 0,95$; o que mostra que a primeira é um indicador eficiente da segunda. Outras associações importantes, todas, assim como as anteriores, significativas a 0,01% de probabilidade, são: frutos maiores e com maiores pesos de sementes tendem a ter sementes maiores (coeficientes de correlação entre PSU e PSS com PSUF superior a 0,80, com PF superiores a 0,71 e com PSSF superiores a 0,77), o que simplifica o melhoramento para estes caracteres em conjunto; a RSSSU tem coeficientes de correlação baixos com os variáveis primárias que a originam: – 0,10 com PSU e 0,18 com PSS.

Ainda analisando a qualidade dos dados, foi avaliada a repetibilidade entre amostras obtidas em um único período (foi considerado o período de safra do 2º ano de avaliação) e entre as médias dos diferentes períodos para 100 acessos tomados ao acaso. Com cinco amostras dentro de um mesmo período foram obtidos coeficientes de determinação (pelos métodos de análise de variância, primeiro componente principal e análise estrutural) entre 96% e 97% para PSU e PSS e, neste caso, duas repetições seriam o suficiente para superar o nível de 90% de determinação; entre 91% e 92% para PSUF; entre 92 e 93% para PF; entre 88 e 89% para PSSF e entre 85 e 87% para RSSSU; mostrando-se a variável RSSSU menos confiável (a média de frutos por amostra foi de 14,4 para a safra do 2º ano e de 10,6 para os quatro períodos, considerando-se todos os acessos avaliados).

Os quatro períodos de avaliação levaram a coeficientes de determinação entre 98 e 99% para as variáveis PSU e PSS e, com a exclusão da única avaliação feita na entre-safra (temporão), as três restantes resultam nos mesmos valores. Um único ano seria suficiente para prever o valor real com coeficiente de determinação superior a 90%, de modo que duas amostras/ano de 40 sementes em um único período de safra podem representar um método rápido e de baixa demanda de esforços para a avaliação de peso da semente em coleções de germoplasma.

As mensurações diretas de fruto (PSUF e PF) alcançaram coeficientes de determinações entre 94 e

95%, com os quatro períodos, e entre 94 e 96%, com os três períodos de safra. Dois períodos de safra resultariam em R^2 superior a 90%, de modo que, para a análise do peso de frutos e sementes úmidas por fruto, poderia ser sugerida a utilização de quatro ou cinco repetições em dois períodos de safra. Dois períodos de safra seriam também suficientes para atingir níveis de R^2 superiores a 0,90 para a variável PSSF, para a qual foram observados coeficientes entre 92 e 94% para os quatro períodos e entre 93 e 94% para os três períodos de safra. Já a RSSSU, necessitou dos quatro períodos para atingir R^2 entre 89 e 90%, alcançando apenas 84 a 86% com os três períodos de safra.

A repetibilidade alcançada é, obviamente, relativa às condições específicas do experimento, o que inclui o número médio de frutos amostrados, de modo que as sugestões, quando da avaliação de outras coleções, devem ser encaradas apenas como um elemento de referência, e ajustadas para cada condição, por exemplo através de análises da repetibilidade, a medida que os dados vão sendo obtidos. As amostras foram obtidas de frutos formados por polinização não controlada, mas houve o cuidado de se escolher épocas em que houvesse um número elevado de frutos e as amostras de sementes foram obtidas de misturas de frutos, para se evitar possíveis efeitos da fonte de pólen, embora, quando analisados, estes efeitos não foram significativos para as variáveis consideradas (Pereira et al, 1994).

Distinções entre acessos e grupos de acessos

As distribuições das variáveis consideradas foram constadas como distintas da distribuição normal pelo Teste de Kolmogorov-Smirnov (PROC UNIVARIATE - SAS Institute, 1988) ao nível de 1% de probabilidade de erro, com as curvas de frequência mostrando, de modo geral (com a exceção da RSSSU), um desvio relativamente pronunciado no sentido de valores mais altos (D para PSU, PSS, RSSSU, PSUF, PF e PSSF, respectivamente 0,080; 0,098; 0,034; 0,086; 0,072 e 0,070). Optou-se, então, pela transformação das variáveis PSU, PSS, PSUF, PF e PSSF através da obtenção de suas raízes quadradas. Este tipo de transformação tem a vantagem de permitir fácil interpretação das médias, pelo dimensionamento aproximado do que seriam os valores não transformados e, embora seja, na verdade, indicado

para dados de contagem (Johnson e Wichern, 1999), resultou, aqui, em ampla redução dos desvios (viés) unilaterais das curvas de frequências (D para PSU, PSS, PSUF, PF e PSSF, respectivamente 0,052; 0,067; 0,049; 0,041 e 0,034), apesar dos testes de Kolmogorov-Smirnov terem continuado a mostrar diferenças significativas (isto em razão, também, do número elevado de dados, o que torna o teste muito rigoroso).

Mais delicado é o aspecto de que as variâncias dos diversos acessos foram constatadas como não homogêneas pelo teste de Bartlett (Martin, 1995). Tal fato, no entanto, é intrínseco às análises desejadas, estando relacionado às dimensões do estudo. A princípio, aumentando-se o número de tratamentos amplia-se a probabilidade de se alcançar valores mais extremos de variância, e esta pode sofrer efeito de escala, tendo relação com a média dos genótipos. Para dois genótipos com variação proporcionalmente igual (mesmo CV) a relação entre as duas variâncias (maior sobre a menor) é igual ao quadrado da relação entre as médias. Assim, se há médias que se distinguem na razão de mais que 2,65 vezes, e mesmo CV, a relação entre as variâncias suplantará 7 vezes (valor considerado como limite para adequação por Gomes, 1990).

Na verdade, com maiores médias, esperam-se menores CVs – efeito Roginskii-Yablokov (Lynch e Walsh, 1998), mas, no trabalho em questão, que, pela grande diversidade armazenada na coleção, engloba freqüentes distinções superiores ao valor discutido, há a preservação de algum efeito de escala. Como exemplo, temos que, para PSU, cujos valores extremos (não ajustados) são 1,49g (RB 37) e 6,66g (UF 667) e a média 3,36, as médias para os 40 genótipos de maior variância e os 40 de menor foram, respectivamente: 4,26g e 2,37g (as outras variáveis também mostraram amplas diferenças entre os dois grupos: PSS – 1,52 e 0,87g; PSUF – 146,02 e 87,98g; PF – 668,322 e 486,99g; PSSF – 53,07 e 32,13g, respectivamente). Desta forma, as características do material estudado impedem que todos os acessos sejam analisados simultaneamente e estes foram divididos, para cada variável, em grupos, de forma a que os maiores valores de variância dentro de acesso não suplantassem em 7 vezes os menores.

Dois grupos, um com maiores variâncias e outro com menores, comportaram, respectivamente, 442 e 94 acessos para a variável PSU (5635 e 1338 observações), 408 e 132 para o PSU (5192 e 1779

observações), 417 e 117 para a RSSSU (5429 e 1439 observações), 482 e 61 para o PSUF (6466 e 861 observações), 475 e 67 para o PF (6403 e 942 observações), 442 e 98 para o PSSF (5584 e 1327 observações).

As análises de variância, segundo o modelo em que foram considerados como fontes de variação o período de amostragem, acessos ajustados para efeito de período e a interação entre eles, ajustada para os anteriores, todos considerados como fixos (PROC GLM - SAS Institute, 1988), mostraram diferenças significativas a, no máximo, 5% de probabilidade de erro para todos os efeitos com todas as variáveis e grupos (mais precisamente, para o grupo de menores variâncias a interação foi significativa a 0,38% para PMSUF, a 1,11% para PMF e a 0,35% para PSSF, enquanto que para todas as demais variáveis, grupos e efeitos as diferenças foram significativas a 0,01%). Confirmando estas informações tem-se que, com cada variável analisada como medidas repetidas em modelo multivariado (cada qual seccionada conforme as quatro épocas e analisadas como quatro variáveis usando PROC GLM – MANOVA, para a interpretação dos efeitos de acesso e PROC GLM – REPEATED, para a consideração dos efeitos de época e interação época x acesso), os resultados obtidos foram semelhantes: todos os efeitos para todas as variáveis e grupos foram significativos a 0.01% pelo teste de Wilks. Esta metodologia, não requer homogeneidade para as matrizes de covariância dos diversos tratamentos (em medidas repetidas, as correlações entre medidas podem diferir para diferentes épocas) e uniformidade da matriz comum de covariâncias; e deve ser a escolhida quando da ocorrência de influência destes fatores na diferenciação dos resultados do método univariado em relação a esta (Danford et al., 1960; Littell et al., 1996). Ressalta-se, no entanto, que, no estudo em questão, diferenças podem ser geradas, também, pelo fato de que, com a utilização da metodologia multivariada há uma considerável redução do número de observações, visto que são eliminadas da análise todas aquelas para as quais há um dado perdido em qualquer época (os dados das outras três épocas são eliminados).

Reduzindo-se o modelo univariado para efeitos de período e acessos ajustados para período, de modo a tomar todas as médias e contrastes possíveis entre acessos e entre períodos estimáveis, verifica-se para

o período de entre safra frutos maiores (PF), com maior peso de semente (PSUF e PSSF) e com sementes maiores (PSU e PSS), o que é coerente com o menor número de frutos no período e, portanto, menor competição por fotoassimilados (considerando o limite de 5% de significância o período de entressafra não foi significativamente diferente para PSU e PSS dos dois últimos períodos de safra com o grupo de variâncias menores e para PSS do último período de safra com o grupo de maiores variâncias – não apresentado). Já para a variável RSSSU, os maiores valores foram observados com o primeiro período de safra (não distinto do último para o grupo 2 – menores variâncias).

As distinções entre períodos, mesmo de safra em diferentes anos, associadas à significância dos efeitos da interação entre períodos e acessos ilustra a existência de limitações nos processos tradicionais de caracterização de germoplasma de cacau para os descritores usuais (ICGD, 2024), análogos às características aqui avaliadas: pesos e dimensões de fruto e semente. Esta caracterização é minuciosa e demanda grandes esforços, o que resulta na impossibilidade de avaliar um grande número de acessos a cada período e na necessidade de coleta de frutos por períodos longos em proporções diferentes para cada genótipo.

A redução do modelo foi repetida para a estimação das médias de clones e é, aqui, necessária, porque, com a consideração da interação acessos x época, os contrastes entre médias, ou as médias obtidas por LSMEANS/PROC GLM - SAS Institute (1988), que gera os valores de classes e sub-classes que seriam esperados para um desenho balanceado, não são estimáveis quando da presença de células vazias: ausência de dados sobre um clone em um período. Com a redução do modelo, as médias de clones são ajustadas apenas para período, o que mantém a coerência entre a forma de estimação das médias e as análises de variância com o modelo completo, no qual foram consideradas as correções de cada efeito apenas pelos que o precedem (Soma de quadrados Type 1 - SAS Institute, 1988).

Em respeito às distinções entre acessos, as Tabelas 2 e 3 apresentam os 40 materiais com os valores mais favoráveis (médias ajustadas) para cada uma das variáveis, assim como valores aproximados de DMS

Tabela 2 - Médias ajustadas de acessos de cacaueiros para os caracteres: peso de uma semente úmida (PSU), peso de uma semente seca (PSS), relação semente seca / semente úmida (RSSSU); e os respectivos grupos de variâncias residuais.

CLONE	PSU	GR	CLONE	PSS	GR	CLONE	RSSSU	GR
UF 667	2,59	1	RIM 76	1,52	1	IAC 1	0,44	1
UF 650	2,53	1	UF 677	1,52	1	CC 34	0,43	1
UF 221	2,51	1	ICS 60	1,52	1	BE 5	0,43	1
UF 677	2,50	1	UF 667	1,51	1	SIC 812	0,42	1
CEPEC 532	2,50	1	SGU 50	1,51	1	SIC 801	0,42	1
CC 11	2,49	1	UF 650	1,51	1	SIC 891	0,42	1
ICS 40	2,47	1	CC 11	1,50	2	SIC 842	0,42	2
UF 654	2,46	1	UF 221	1,50	1	CEPEC 10	0,42	1
ICS 39	2,46	1	ICS 39	1,49	1	SIAL 512	0,42	1
ICS 60	2,45	1	UF 10	1,48	1	ICS 16	0,42	1
RIM 76	2,45	1	UF 654	1,48	1	LARANJA 2	0,41	1
UF 668	2,44	1	ICS 40	1,48	1	SIC 813	0,41	1
UF 12	2,42	1	ICS 78	1,48	1	CEPEC 538	0,41	1
SGU 50	2,41	1	UF 168	1,47	1	SIAL 325S1	0,41	1
UF 10	2,40	1	UF 11	1,47	1	CC 10	0,41	2
UF 11	2,39	1	UF 12	1,47	1	SIAL 2	0,41	2
ICS 47	2,35	1	CC 10	1,47	1	CEPEC 88	0,41	1
UF 168	2,32	1	UF 668	1,46	1	CEPEC 41	0,41	1
ICS 6	2,32	1	DES 267	1,46	1	SIC 961	0,41	2
CEPEC 514	2,32	1	CEPEC 532	1,46	1	SIC 864	0,41	2
ICS 78	2,32	2	CEPEC 514	1,45	1	CJ 8	0,41	1
ICS 75	2,31	1	ICS 48	1,44	1	DES 267	0,41	1
CC 10	2,30	1	ICS 47	1,38	1	SIC 872	0,41	2
DES 267	2,29	1	ICS 75	1,38	1	ICS 78	0,41	1
ICS 48	2,28	1	SC 49	1,38	1	SIC 848	0,41	2
OC 66	2,27	1	EET 59	1,36	2	SIAL 910	0,41	1
UF 708	2,27	2	UF 708	1,36	2	SIC 823	0,40	1
CEPEC 67	2,25	1	ICS 16	1,35	1	SIC 806	0,40	2
RIM 8	2,23	1	RIM 8	1,35	1	UF 168	0,40	1
ICS 9	2,23	1	GS 29	1,35	1	ICS 48	0,40	1
SCR 4	2,22	1	CC 18	1,35	1	SIAL 698	0,40	1
CEPEC 501	2,22	1	OC 66	1,34	2	CEPEC 534	0,40	1
ICS 89	2,21	1	OC 61	1,34	1	SIC 831	0,40	2
GS 29	2,21	1	RIM 117	1,34	1	SIC 439	0,40	1
SC 49	2,21	1	SCR 4	1,33	1	RB 41	0,40	1
CEPEC 68	2,20	1	GW 1	1,33	2	CEPEC 28	0,40	1
22 P	2,20	1	ICS 9	1,33	1	CJ 9	0,40	1
EET 59	2,18	1	10 P	1,33	2	IQ 1	0,40	2
10 P	2,17	2	ICS 6	1,32	1	EEG 50	0,40	1
RIM 9	2,17	2	RIM 41	1,32	1	SIAL 581	0,40	1
MÉDIA*	1,81	1	MÉDIA*	1,09	1	MÉDIA*	0,37	1
MÉDIA*	1,69	2	MÉDIA*	1,05	2	MÉDIA*	0,37	2
DMS-TUKEY 5%	0,21	1	DMS-TUKEY 5%	0,13	1	DMS-TUKEY 5%	0,06	1
DMS-TUKEY 5%	0,13	2	DMS-TUKEY 5%	0,08	2	DMS-TUKEY 5%	0,04	2

* Média geral de todos os genótipos avaliados do grupo.

(diferença mínima significativa a 5% de probabilidade) obtidos, para cada grupo, a partir da média harmônica do número de repetições dos diversos genótipos (no modelo não reduzido), uma vez que as comparações dois a dois requeridas para esta situação, em que os números de

repetições são diferentes, geraria matrizes de probabilidade com dimensões excessivamente grandes, o que dificultaria muito a discussão dos resultados.

Há nítido destaque de genótipos procedentes da América Central, México, Caribe e Venezuela

Tabela 3 - Médias ajustadas de acessos de cacauzeiros para os caracteres: peso de sementes úmidas por fruto (PSUF), peso total de fruto (PF), e peso de semente seca por fruto (PSSF); e os respectivos grupos de variâncias residuais

CLONE	PSUF	GR	CLONE	PF	GR	CLONE	PSSF	GR
UF 677	15,32	1	CCN 34	38,95	1	UF 677	9,38	1
CCN 51	15,04	1	LAFI 7	36,55	1	ICS 60	9,25	1
ICS 6	14,87	1	TSH 1188	34,03	1	RIM 76	9,06	1
ICS 40	14,83	1	UF 677	32,71	1	DES 267	9,03	1
ICS 39	14,79	1	UF 613	32,51	1	CEPEC 31	8,99	1
UF 650	14,72	1	SNK 12	32,01	2	ICS 39	8,95	2
RIM 76	14,69	1	RIM 76	31,99	1	UF 650	8,94	1
UF 668	14,67	1	OC 66	31,81	1	SGU 50	8,90	1
TSH 1188	14,62	1	UF 710	31,32	1	ICS 40	8,87	1
CCN 34	14,61	2	UF 708	31,05	1	UF 668	8,83	1
ICS 60	14,55	1	IMC 30	31,05	1	CCN 34	8,81	2
UF 710	14,49	1	MOQ 417	30,71	1	CC 18	8,73	1
CEPEC 532	14,47	1	UF 668	30,62	1	UF 654	8,69	1
LAFI 7	14,46	1	CEPEC 522	30,57	1	UF 12	8,66	1
UF 654	14,41	1	ICS 39	30,55	1	UF 10	8,64	1
SGU 50	14,36	1	UF 650	30,44	1	UF 168	8,57	1
UF 667	14,29	1	CEPEC 31	30,36	1	CC 11	8,55	1
UF 12	14,23	1	BE 9	30,24	1	CEPEC 49	8,54	1
DES 267	14,22	1	SGU 60	29,98	1	CCN 51	8,52	1
CC 11	14,12	1	SC 49	29,96	1	ICS 78	8,50	1
TSH 565	14,07	1	CCN 51	29,94	1	ICS 6	8,49	1
CC 18	14,00	1	UF 20	29,92	2	UF 11	8,47	1
ICS 47	13,92	1	ICS 6	29,80	1	ICS 48	8,45	1
CEPEC 49	13,84	1	CEPEC 514	29,73	1	UF 710	8,44	1
IMC 30	13,84	1	IMC 67	29,73	2	UF 667	8,35	1
CEPEC 31	13,82	1	ICS 40	29,72	1	CEPEC 514	8,28	1
CEPEC 50	13,75	1	MOQ 647	29,70	1	CEPEC 532	8,26	1
UF 11	13,75	1	CC 11	29,45	1	CEPEC 50	8,24	1
RIM 8	13,71	1	CEPEC 5	29,39	1	TSH 1188	8,23	1
UF 613	13,66	1	ICS 60	29,34	1	UF 613	8,23	1
CEPEC 29	13,60	1	CEPEC 531	29,33	1	ICS 47	8,19	1
UF 221	13,55	1	CEPEC 27	29,26	1	UF 708	8,19	1
UF 10	13,54	1	IMC 76	29,23	1	CC 10	8,18	1
UF 168	13,52	1	COJON TORO	29,12	1	22 P	8,14	1
ICS 48	13,47	2	SGU 50	29,11	1	RIM 113	8,12	1
TSA 792	13,43	1	RIM 44	29,06	1	UF 221	8,09	1
RIM 113	13,42	1	UF 168	29,04	1	RIM 44	8,08	1
CEPEC 59	13,40	1	UF 667	29,04	1	CEPEC 29	8,08	1
CEPEC 514	13,35	1	CEPEC 86	29,03	2	ICS 16	8,07	2
ICS 78	13,35	1	CEPEC 33	29,00	1	RIM 8	8,00	1
MÉDIA*	10,80	1	MÉDIA*	23,63	1	MÉDIA*	6,55	1
MÉDIA*	10,00	2	MÉDIA*	22,03	2	MÉDIA*	6,07	2
DMS-TUKEY 5%	2,50	1	DMS-TUKEY 5%	5,40	1	DMS-TUKEY 5%	1,62	1
DMS-TUKEY 5%	1,55	2	DMS-TUKEY 5%	3,03	1	DMS-TUKEY 5%	1,07	1

* Média geral de todos os genótipos avaliados do grupo.

(Criollos, Trinitários e híbridos de Criollo) para as características PSU e PSS (nota: o código DES indica materiais de procedência desconhecida e também o são os CEPECs com numeração superior a 500). Dentre os maiores valores de PSU inclui-se ainda

uma seleção da Colômbia (SC 49), uma do Equador (EET 59) e uma seleção do CEPEC de genealogia desconhecida (CEPEC 68) e dentre os maiores valores de PSS uma seleção do Equador (EET 59) e uma da Indonésia (GW 1).

Dentre os 40 maiores valores de RSSU, 21 foram observados com seleções feitas dentro das variedades locais da Bahia (CEPEC 10, SIC, SIAL, EEG), sendo a proporção muito superior à frequência relativa destes na coleção; e valores elevados podem ser considerados como de interesse, quando o objetivo é utilizar comercialmente somente as amêndoas, pois significam menores consumos de fotoassimilados para a produção de polpa.

Com as características de fruto, voltam a se destacar materiais procedentes da América Central, México e Caribe, agora acompanhados de híbridos (CCN, TSA, TSH, CEPECs 27, 29, 31, 33, 49, 50 e 59).

Incluem-se, ainda, nas listas de maiores valores para PSUF, uma seleção de Samoa (LAFI 7) e um clone coletado no Peru (IMC 30); para PF, além dos dois anteriores, uma seleção da República dos Camarões (SNK 12), um Criollo da Venezuela (OC 66), dois clones coletados no Equador (MOQ 417 e MOQ 647), dois coletados no Peru (IMC 67 e IMC 76), um coletado no Pará (Be 9), uma seleção da Colômbia (SC 49), um híbrido de Criollo obtido na Bahia (CEPEC 5), um material de origem desconhecida (Cojon del Toro) e uma seleção do Sul da Bahia, de material conhecido popularmente como Pará manteiga, cuja possível ascendência será posteriormente discutida (CEPEC 86), enquanto que para PSSF a lista dos 40 maiores valores fica restrita aos materiais procedentes da América Central, México e Caribe, e híbridos, tendo como possíveis exceções, apenas, os clones de origem desconhecida: DES e CEPEC com numeração superior a 500.

Para uma melhor compreensão das diferenças entre materiais de acordo com o tipo, procedência, grau de domesticação e variedade, foram analisadas as diferenças entre grupos ou séries de genótipos, constituídos conforme disposto na Tabela 1. Novamente, as análises foram realizadas agrupando-se materiais com maior homogeneidade de variância (limite máximo de 7 vezes entre a maior e a menor) e as séries e grupos listados foram incluídas nas análises quando da ocorrência de, no mínimo, dois acessos, do grupo ou série, dentro do conjunto definido pelas variâncias.

Com o grupo de maiores variâncias, foram constatadas diferenças significativas a 0,01% para todos os efeitos (época de amostragem, grupo de acessos, acessos dentro de grupo, interação entre época

e grupo e interação entre época e acessos dentro de grupo) e variáveis (PSU, PSS, RSSU, PSUF, PF e PSSF) e com o de menores só não foram significativos, a 5% de probabilidade, as interações período x acesso dentro de grupo para as variáveis PSUF e PF (em todos os casos, época, grupo e acesso dentro de grupo foram significativos a 0,01% de probabilidade).

Com cada variável analisada como medidas repetidas em modelo multivariado (PROC GLM – MANOVA, para a interpretação dos efeitos de grupo e acesso dentro de grupo e PROC GLM – REPEATED, para a consideração dos efeitos de época e interações – SAS System, SAS Institute, 1988), foi constatada significância a 0,01% de probabilidade pelo teste de Wilks para todos os efeitos, variáveis e grupos. Estes resultados permitem a conclusão de que há ocorrência generalizada de efeitos de interações.

De modo coerente ao já relatado, os maiores valores para PSU e PSS foram observados com grupos e séries procedentes do México, América Central, Venezuela e Caribe: UF, ICS, RIM, MEX, OC, GS, CC, SGU e VE; alguns grupos de materiais híbridos: CCN e CTF e com a série SC da Colômbia, que foi gerada por seleções do ICA (Instituto Colombiano de Agropecuária) e engloba materiais do tipo Tinitário (ICGD, 2024) (Tabela 4 e Tabela 5).

A série CAS, procedente de seleções feitas no Campo Agrícola de Santarém, onde havia clones introduzidos e seleções locais (Almeida, et al., 1987), e que tem caracterização de tipo dúbia, está entre os grupos de maiores PSU, mas aí se coloca mais pela sua elevada proporção de polpa (baixo RSSU). A série EET, obtida de seleções dentro de populações formadas pela mistura da variedade nacional do Equador com introduções, principalmente de Trinitários e Criollos (Bartley, 2001), tem, também, valores elevados de PSU e PSS.

Os maiores valores para estes caracteres são observados em seleção efetuadas em material domesticado, sendo exceções as seleções das variedades locais da Bahia, que mostram valores reduzidos, e neste aspecto estão estas acompanhadas de grupos também forasteiros, porém selvagens, resultantes de prospeção em florestas ou, as vezes, em fazendas, mas dentro da área de dispersão natural da espécie: SCA, ABR, CEQ (na verdade de origem desconhecida, mas com características que o tipificam

Tabela 4- Médias ajustadas de grupos com origem comum para os caracteres: peso de uma semente úmida (PSU), peso de uma semente seca (PSS), peso de sementes úmidas por fruto (PSUF), peso total de fruto (PF), peso de semente seca por fruto (PSSF) e relação semente seca / semente úmida (RSSSU); e número de representantes considerados (NR)

GR	PSU	NR	GR	PSS	NR	GR	RSSSU	NR	GR	PSUF	NR	GR	PF	NR	GR	PSSF	NR
Grupo 1 - Maiores Variâncias																	
UF	2,27	19	UF	1,36	20	BAH	0,38	78	UF	12,95	24	CCN	33,63	2	UF	7,92	24
ICS	2,24	15	ICS	1,33	16	MOQ	0,38	2	RIM	12,60	18	MOQ	29,22	3	RIM	7,63	18
RIM	2,15	14	RIM	1,32	12	CTF	0,38	14	ICS	12,52	16	UF	27,66	20	MEX	7,45	4
MEX	2,15	4	MEX	1,28	2	SPA	0,38	6	IMC	12,29	9	RIM	27,63	18	ICS	7,42	15
OC	2,13	4	OC	1,26	3	BE	0,38	10	CAS	12,17	2	CAS	27,51	2	CTF	7,38	14
GS	2,04	2	SC	1,24	4	EET	0,37	8	CTF	12,03	14	IMC	27,47	8	SC	7,38	3
CC	1,99	13	CCN	1,18	2	VE	0,37	2	MEX	12,01	5	ICS	27,43	14	MOQ	7,31	2
SC	1,99	3	CC	1,18	13	RIM	0,37	14	MA	11,61	4	OC	27,31	4	OC	6,99	3
SGU	1,98	5	SGU	1,18	5	CC	0,37	11	CC	11,57	12	SC	27,08	4	CC	6,99	13
CCN	1,94	2	GS	1,17	2	UF	0,37	17	OC	11,49	4	MEX	26,61	5	IMC	6,95	9
VE	1,90	2	CTF	1,15	13	CA	0,37	3	SC	11,41	4	MA	26,46	4	SGU	6,92	6
CAS	1,90	3	VE	1,15	4	SC	0,37	3	CA	11,37	6	SPA	26,29	5	GS	6,77	2
EET	1,90	11	EET	1,14	9	SGU	0,36	7	SGU	11,36	6	GS	25,91	3	MA	6,77	5
CTF	1,88	14	CJ	1,08	4	MEX	0,36	4	MOQ	11,09	2	CA	25,35	6	CA	6,76	6
CHC	1,87	6	BE	1,07	7	ICS	0,36	16	BE	11,08	9	BE	25,18	10	BE	6,72	8
BE	1,86	7	CHC	1,07	7	CJ	0,36	7	CJ	10,86	6	SGU	25,08	6	EET	6,60	10
PA	1,86	25	CA	1,07	3	CSU	0,36	6	SDE	10,83	34	VE	25,02	4	SPA	6,54	5
IMC	1,82	5	MOQ	1,07	2	MA	0,36	5	EET	10,82	11	CTF	24,37	14	CJ	6,43	6
MA	1,80	4	MA	1,06	5	SDE	0,36	31	GS	10,81	2	EET	23,75	10	SDE	6,43	34
MOQ	1,79	3	IMC	1,06	4	POU	0,36	11	SPA	10,75	7	CHC	23,59	8	VE	6,35	3
CA	1,79	6	PA	1,05	19	ABR	0,35	10	VE	10,72	3	CC	23,57	11	BAH	6,24	90
SDE	1,76	37	SDE	1,03	30	GS	0,35	3	PA	10,47	23	SDE	23,31	36	CHC	6,11	9
CJ	1,73	7	CAS	1,01	2	RB	0,35	12	PER	10,35	4	PER	22,69	5	PA	5,97	21
SPA	1,70	6	BAH	1,01	95	PER	0,35	3	CEQ	10,17	5	CJ	22,64	6	POU	5,91	10
ABR	1,69	9	ABR	1,00	7	NA	0,35	4	CHC	10,13	8	RB	22,34	9	ABR	5,79	9
NA	1,65	5	SPA	0,98	3	CHC	0,34	8	BAH	10,07	95	PA	22,21	25	CEQ	5,71	3
PER	1,64	4	POU	0,96	10	CAS	0,33	3	ABR	9,97	10	POU	22,05	8	NA	5,51	4
POU	1,63	7	NA	0,96	5	CEQ	0,33	2	POU	9,85	11	ABR	21,95	10	PER	5,22	3
BAH	1,62	94	RB	0,91	6	IMC	0,32	7	NA	9,43	3	CEQ	21,63	4	CSU	5,12	4
RB	1,57	6	PER	0,89	2	PA	0,32	23	SCA	9,01	2	BAH	21,50	100	RB	5,10	9
CSU	1,38	3	CSU	0,84	4	SCA	0,31	2	RB	8,71	12	NA	19,56	5	SCA	5,09	2
									CSU	8,33	6	CSU	19,20	7			
DMS	0,07			0,05			0,02			0,78			1,89			0,56	
TUKEY 5%																	

como Forastero Amazônico), CJ, IMC, CSU, PER, RB, NA, POU, etc.

Este padrão se repete para peso de semente por fruto e peso de fruto, mas a separação não é tão nítida, havendo grupos de seleções de materiais amazônicos inseridos dentre àqueles com maiores valores: MA, IMC, o que mostra a viabilidade de se buscar genes para o melhoramento destes caracteres também neste tipo (o grupo MOQ do Equador, incluído entre os de maiores valores, não pode ser caracterizado como Forastero e sua constituição é análoga à dos EETs – ICGD, 2024).

Por outro lado, alguns grupos importantes quanto à possibilidade de fornecer genes de resistência à vassoura-de-bruxa e à podridão parda e para o incremento do teor de gordura entre eles: CSU, POU,

NA, CEQ, PER, além dos Scavinas (Pires et al., 1997; 1998; 2021; 2022;), possuem frutos e sementes pequenas. O grupo SCA se destaca, ainda, por apresentar um dos menores valores de RSSSU, o que indica um amplo consumo de fotoassimilados na produção de polpa (isto, se analisado por outro ângulo, e tendo-se, também, em conta o sabor adocicado característico de sua polpa, dominante em cruzamentos, notadamente em respeito ao clone Scavina 6, leva a constatação de que o grupo é promissor em respeito a este caráter para utilização no melhoramento, se houver interesse no aproveitamento da polpa). Para a RSSSU não há tendência clara para origem ou tipo.

Quando da consideração dos caracteres simultaneamente, a análise multivariada, feita a partir das médias ajustadas, obtidas pelas análises anteriores,

Tabela 5- Médias ajustadas de grupos com origem comum para os caracteres: peso de uma semente úmida (PSU), peso de uma semente seca (PSS), peso de sementes úmidas por fruto (PSUF), peso total de fruto (PF), peso de semente seca por fruto (PSSF) e relação semente seca / semente úmida (RSSSU); e número de representantes considerados (NR)

GR	PSU	NR	GR	PSS	NR	GR	RSSSU	NR	GR	PSUF	NR	GR	PF	NR	GR	PSSF	NR
Grupo 2 - Menores Variâncias																	
RIM	2,14	5	MEX	1,30	3	BAH	0,39	27	ICS	13,26	2	UF	27,84	4	ICS	8,10	3
ICS	2,05	3	RIM	1,29	7	CC	0,39	2	SDE	11,81	2	CC	25,70	3	SPA	6,95	2
UF	1,94	5	SGU	1,23	2	ABR	0,38	2	GS	11,13	2	SPA	24,01	2	PER	6,91	2
SGU	1,94	2	UF	1,21	4	RIM	0,37	5	CC	10,96	2	CHC	23,09	3	CFF	6,80	2
VE	1,94	2	EET	1,18	2	OC	0,37	3	PA	10,80	4	CAS	22,95	2	SDE	6,75	3
PA	1,79	3	CAS	1,13	2	EET	0,37	3	BE	10,43	2	BAH	22,05	9	GS	6,60	2
GS	1,78	2	CHC	1,10	3	UF	0,37	7	CAS	10,26	2	ICS	21,81	4	BE	6,48	3
CHC	1,78	4	SDE	1,09	7	ICS	0,36	2	CHC	9,93	3	PA	21,61	3	CAS	6,44	3
MA	1,70	2	SPA	1,08	4	CEQ	0,35	2	BAH	9,92	13	RB	20,28	7	ABR	6,28	2
BAH	1,66	15	GS	1,06	2	IMC	0,35	2	NA	9,10	2	POU	17,99	3	BAH	6,16	19
IMC	1,62	2	BE	1,06	4	SDE	0,35	5	CSU	8,46	2	CSU	17,94	2	CHC	5,90	2
SCA	1,62	2	CA	1,06	3	CHC	0,35	3	RB	8,19	4				PA	5,82	7
BE	1,60	3	PA	1,04	9	CA	0,34	3							RB	5,04	7
POU	1,58	2	PER	1,03	2	RB	0,34	3							CSU	4,85	4
ABR	1,51	3	BAH	1,00	14	PER	0,34	2									
CEQ	1,51	3	IMC	1,00	4	VE	0,34	2									
RB	1,45	9	CJ	0,99	3	CSU	0,32	3									
CSU	1,36	5	ABR	0,96	5	PA	0,32	5									
			SCA	0,89	2												
			CEQ	0,88	3												
			RB	0,88	8												
			CSU	0,79	4												
DMS	0,06			0,03			0,02			0,60			1,05			0,40	
TUKEY	5%																

com os grupos como fonte de variação e os diversos clones dentro de grupos como repetições destes, mostrou diferenças significativas entre grupos, a 0,01% de probabilidade, pelo teste de Wilks.

As probabilidades para a aceitação ou rejeição das hipóteses de igualdade para os contrastes entre as médias dos diversos grupos são apresentados nas Tabela 6 e Tabela 7, e as amplas distinções entre grupos, notadamente os de diferentes origens, mostra que os caracteres considerados podem ser usados como elementos de caracterização de materiais. Ilustrando este aspecto tem-se que, por exemplo, o grupo de seleções feitas nas variedades tradicionais da Bahia só não difere significativamente a 5% de probabilidade dos grupos CJ e Pound, o grupo CCN difere de todos, o grupo CSU só não difere do SCA e o ICS dos grupos MEX, OC e RIM.

A análise de variáveis canônicas a partir das seis variáveis discutidas, utilizando-se as médias ajustadas e tendo-se os acessos como repetições de grupos (PROC CANDISC – SAS INSTITUTE, 1988) resultam em 6 variáveis com diferenças significativas (5%) entre os grupos, sendo que três delas acumulam

83,2% da variação total (primeira variável 49,5%, segunda 20,3%, terceira 13,5%, Quarta 9,6%, Quinta 4,1%, Sexta 3,0%). A primeira é determinada, principalmente (maiores coeficientes de ponderação), pelo PSS, RSSSU e PSUF, mas mantém alta correlação com PSU (0,96) e PSS (0,91), média com PSUF (0,67), PMF (0,59) e PSSF (0,63) e tem baixa correlação com RSSSU. A segunda tem para as variáveis PSS, RSSSU e PSSF os maiores coeficientes de ponderação e as correlações mais altas ocorrem com a RSSSU (0,85) e o PSSF (0,66). Já a terceira é determinada principalmente pelo PMF, única variável com quem preserva correlação superior a 50% (0,69), enquanto as demais variáveis canônicas não têm correlação superior a 50% em valor absoluto com nenhuma das variáveis originais.

Estes resultados mostram uma clara diferenciação entre materiais domesticados procedentes do México, América Central, Caribe e Venezuela e materiais procedentes da América do Sul, notadamente os selvagens, conforme apresentado na Figura 1, onde estão dispostas as médias de grupos para as três primeiras variáveis canônicas. Os primeiros são

representados por estrelas, verdes para os procedentes da Venezuela e vermelhas para os demais, e ocupam posições relativas a valores elevados para a primeira e segunda variável canônica e de intermediários a reduzidos para a terceira. Os materiais procedentes da América do Sul são representados pelo símbolo de paus, a exceção da série EET, do Equador, mais próxima dos Trinitários e Criollos, que é representada por uma flâmula negra, e do grupo CEQ representado por um losango negro, sendo que este, de origem desconhecida, posiciona-se entre os amazônicos.

Os materiais da Colômbia (SPA e SC) estão destacados em amarelo e o grupo que se coloca entre os Trinitários e Criollos é o SC. Este último, conforme já descrito, tem como origem seleções realizadas pelo ICA (Instituto Colombiano de Agropecuária) e engloba materiais do tipo Trinitário, enquanto o grupo SPA é formado, também por seleções conduzida por este Instituto, mas em uma progênie de uma planta procedente do Peru, provavelmente da região do rio Napo (ICGD, 2024), o que condiz com seus posicionamentos. Os grupos de acesso procedentes do Brasil são apresentados em verde (a maioria), com o de seleções da Bahia destacado em vermelho. Este último tem valor reduzido para a terceira variável (alta RSSSU e PMF baixo), sendo esta posição oposta aos materiais do Equador MOQ (realçados em negro), selecionados na Fazenda Moquique, por Pound (ICGD, 2024), cujos representantes aqui considerados provêm de três árvores diferentes que podem envolver genes da variedade nacional do Equador e de Trinitários. É também oposta aos híbridos do Equador CCN, (destacados em verde - os híbridos são representadas por pirâmides), que descendem de um material desconhecido de denominação Canellos e cruzamentos feito por Homero Castro envolvendo Trinitários (informação não

Tabela 6 - Probabilidades de erro para a rejeição da hipótese de nulidade de contrastes entre médias de grupos (GR) - análise multivariada para o peso de uma semente úmida, de uma semente seca, de sementes úmidas por fruto, de semente seca por fruto, peso total de fruto e relação semente seca semente úmida - teste de Wilks

GR	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
ABR	1	0,0107	0,1641	0,2464	0,1898	0,0005	0,0001	0,0007	0,0003	0,9032	0,0024	0,0001	0,0044	0,1415	0,0001	0,0008	0,1833	0,0001	0,0001	0,0001	0,0001	0,5388
BAH	2		0,0020	0,0008	0,0003	0,0001	0,0001	0,0001	0,0001	0,2699	0,0001	0,0001	0,0001	0,0051	0,0001	0,0001	0,0014	0,0001	0,0001	0,0182	0,0001	0,0026
BE	3			0,9377	0,0790	0,0001	0,0001	0,0001	0,0001	0,3111	0,0001	0,0095	0,0032	0,3671	0,0001	0,0098	0,9477	0,0001	0,0001	0,0036	0,0001	0,0237
CA	4				0,2651	0,0001	0,0001	0,0002	0,0002	0,2669	0,0001	0,0081	0,0036	0,3000	0,0001	0,3240	0,9988	0,0001	0,0003	0,0054	0,0001	0,0675
CAS	5					0,1256	0,0001	0,0002	0,4713	0,1759	0,0002	0,0031	0,2854	0,6128	0,0027	0,0146	0,2604	0,0528	0,0001	0,1033	0,0544	0,5553
CC	6						0,0001	0,0001	0,0055	0,0200	0,0001	0,0001	0,6936	0,1725	0,0001	0,0001	0,0002	0,0034	0,0001	0,0066	0,0611	0,0001
CCN	7							0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0004	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
CEQ	8								0,0001	0,0007	0,0090	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0006
CHC	9									0,0018	0,0001	0,0001	0,0595	0,5924	0,0001	0,0001	0,0011	0,0079	0,0001	0,0561	0,0381	0,0001
CJ	10										0,0009	0,0550	0,0346	0,2582	0,0001	0,0012	0,1702	0,0001	0,0001	0,2874	0,0001	0,2481
CSU	11											0,0001	0,0001	0,0001	0,0001	0,0002	0,0001	0,0001	0,0268	0,0001	0,0001	
CTF	12												0,0001	0,0425	0,0001	0,0001	0,0068	0,0001	0,0001	0,0002	0,0001	0,0001
EET	13													0,7052	0,0001	0,0001	0,0089	0,0303	0,0001	0,0096	0,1673	0,0001
GS	14														0,0011	0,0004	0,4285	0,1217	0,0006	0,0877	0,0762	0,0005
ICS	15															0,0001	0,0001	0,1488	0,0001	0,0001	0,0866	0,0001
IMC	16																0,2150	0,0001	0,0001	0,0001	0,0001	0,0035
MA	17																	0,0001	0,0030	0,0053	0,0002	0,0001
MEX	18																		0,0001	0,0001	0,4691	0,0001
MOQ	19																		0,0001	0,0001	0,0001	0,0001
NA	20																			0,0002	0,0009	0,1592
OC	21																				0,0027	0,0001
PA	22																					0,0387

Tabela 7- Probabilidades de erro para a rejeição da hipótese de nulidade de contrastes entre médias de grupos (GR) - análise multivariada para o peso de uma semente úmida, de uma semente seca, de sementes úmidas por fruto, de semente seca por fruto, peso total de fruto e relação semente seca semente úmida - teste de Wilks.

GR		24	25	26	27	28	29	30	31	32	33
ABR	1	0,7195	0,0017	0,0001	0,0001	0,0266	0,5842	0,0001	0,0001	0,0001	0,0001
BAH	2	0,0717	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0003	0,0001	0,0001
BE	3	0,0165	0,0001	0,0001	0,0014	0,0001	0,2453	0,0009	0,6684	0,0001	0,0012
CA	4	0,0264	0,0005	0,0001	0,0016	0,0067	0,4869	0,0017	0,5051	0,0001	0,0002
CAS	5	0,0407	0,0054	0,0078	0,1253	0,2637	0,2437	0,3048	0,0143	0,0062	0,3475
CC	6	0,0001	0,0001	0,0002	0,0100	0,0009	0,0001	0,6148	0,0001	0,0001	0,1589
CCN	7	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
CEQ	8	0,0011	0,0001	0,0001	0,0001	0,0057	0,0001	0,0001	0,0001	0,0001	0,0001
CHC	9	0,0002	0,0001	0,0001	0,2294	0,1559	0,0001	0,1580	0,0001	0,0001	0,2580
CJ	10	0,8160	0,0001	0,0001	0,0006	0,0121	0,9195	0,0080	0,0166	0,0001	0,0027
CSU	11	0,0389	0,0080	0,0001	0,0001	0,0610	0,0001	0,0001	0,0001	0,0001	0,0001
CTF	12	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
EET	13	0,0003	0,0001	0,0019	0,1828	0,0019	0,0003	0,8164	0,0002	0,0001	0,6092
GS	14	0,0500	0,0009	0,0310	0,6385	0,0411	0,2116	0,5867	0,1469	0,0557	0,3730
ICS	15	0,0001	0,0001	0,0578	0,0022	0,0001	0,0001	0,0101	0,0001	0,0066	0,0037
IMC	16	0,0001	0,0001	0,0001	0,0001	0,0004	0,0004	0,0001	0,0029	0,0001	0,0001
MA	17	0,0199	0,0027	0,0001	0,0078	0,0085	0,2985	0,0043	0,8075	0,0001	0,0060
MEX	18	0,0001	0,0001	0,0957	0,6634	0,0001	0,0001	0,0723	0,0001	0,4333	0,2298
MOQ	19	0,0001	0,0001	0,0001	0,0005	0,0001	0,0001	0,0001	0,0075	0,0001	0,0001
NA	20	0,6538	0,0024	0,0001	0,0021	0,1609	0,0272	0,0019	0,0008	0,0001	0,0067
OC	21	0,0001	0,0001	0,1117	0,4616	0,0006	0,0001	0,3448	0,0001	0,0249	0,8685
PA	22	0,0001	0,0001	0,0001	0,0001	0,1061	0,0001	0,0001	0,0001	0,0001	0,0005
PER	23	0,1191	0,0585	0,0001	0,0011	0,0536	0,1000	0,0165	0,0040	0,0001	0,0793
POU	24		0,0002	0,0001	0,0001	0,0379	0,1011	0,0001	0,0018	0,0001	0,0001
RB	25			0,0001	0,0001	0,0382	0,0001	0,0001	0,0011	0,0001	0,0001
RIM	26				0,0273	0,0001	0,0001	0,1409	0,0001	0,0859	0,0167
SC	27					0,0031	0,0001	0,2199	0,0017	0,0991	0,5982
SCA	28						0,0057	0,0043	0,0007	0,0001	0,0090
SDE	29							0,0002	0,0027	0,0001	0,0003
SGU	30								0,0001	0,0115	0,5445
SPA	31									0,0001	0,0006
UF	32										0,0029

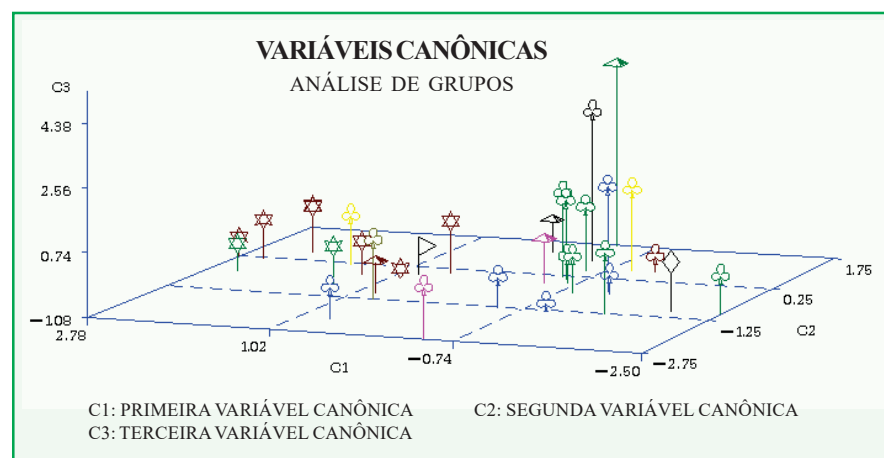


Figura 1 - Dispersão gráfica de séries segundo variáveis canônicas formadas a partir das variáveis: peso de uma semente seca, peso de sementes úmidas por fruto, peso total de fruto, peso de semente seca por fruto e relação semente seca / semente úmida

publicada). Outro dos grupos procedentes da Amazônia está destacado (em marrom) e coloca-se próximo aos Trinitários e Criollos. Este grupo é o CAS cuja origem genética não é definida, conforme já citado (na Estação Experimental de Santarém, onde havia materiais procedentes do Peru, Trinidad, Costa Rica, Equador e da própria região, foram selecionados quatro clones CAS 1, 2, 3, 4, com base em características de fruto e resistência - Almeida et. al., 1987).

Os grupos procedentes do Peru estão representados em azul, com exceção do SCA, em rosa, cor também usada para o grupo formado por seus descendentes (pirâmide), este último próximo, ou mesmo inserido entre os grupos amazônicos (dentre os descendentes de Scavina há cruzamentos com Trinitários, mas não são predominantes), enquanto que o grupo Scavina realça-se na figura pelo menor valor para segunda variável canônica em razão da conjunção de baixo PSSF e baixa RSSSU.

O grupo CEPEC híbrido de Criollo, CHC - pirâmide vermelha, insere-se, coerentemente, próximo aos grupos de Trinitários e Criollos e o grupo CEPEC Trinitários x Forasteros, CTF - pirâmide preta, aproxima-se dos Forasteros (os demais grupos podem ser identificados pela consideração conjunta da

Figura 1 e da Tabela 8, em que são apresentados os valores das coordenadas relativas a cada grupo para as três variáveis canônicas; e uma visualização geral de todos os acessos considerados pode ser obtida através da Figura 2, que preserva as mesmas codificações).

Tem-se, então que este grupo reduzido de variáveis, notadamente as relativas ao peso de sementes e frutos, distingue com eficiência materiais com origens e formação diferentes, inclusive posicionando híbridos de forma intermediária, conforme o maior ou menor peso de cada tipo de ascendência.

Tabela 8 - Valores para grupos de origem comum de variáveis canônicas (CAN) formadas a partir das variáveis: peso de uma semente seca, peso de sementes úmidas por fruto, peso total de fruto, peso de semente seca por fruto e relação semente seca/ semente úmida

GRUPO	CAN1	CAN2	CAN3	GRUPO	CAN1	CAN2	CAN3
ABR	-0,855	-0,461	0,008	MEX	2,528	0,227	0,044
BAH	-1,185	0,652	-0,570	MOQ	-0,417	0,990	3,323
BE	-0,544	0,543	0,818	NA	-0,793	-0,909	-1,079
CA	-0,538	0,111	1,242	OC	2,470	-0,429	-0,283
CAS	0,726	-1,259	0,653	PA	0,709	-2,181	-0,147
CC	0,945	-0,214	-0,857	PER	-0,494	-1,276	-0,176
CCN	-0,329	1,747	4,379	POU	-1,188	-0,402	-0,499
CEQ	-2,021	-0,957	0,310	RB	-1,526	-1,283	0,706
CHC	0,834	-0,983	-0,103	RIM	2,225	0,625	0,254
CJ	-0,614	-0,028	-0,280	SC	1,614	0,149	0,398
CSU	-2,500	-0,977	0,035	SCA	-0,458	-2,754	0,397
CTF	0,114	1,292	-0,080	SDE	-0,434	-0,138	0,265
EET	0,842	-0,063	-0,280	SGU	1,319	-0,244	-0,117
GS	0,574	0,061	0,427	SPA	-0,943	0,684	1,299
ICS	2,779	0,234	-0,509	UF	2,278	0,742	0,229
IMC	-1,125	-0,296	1,960	VE	1,382	-0,706	-0,011
MA	-0,476	0,186	1,417				

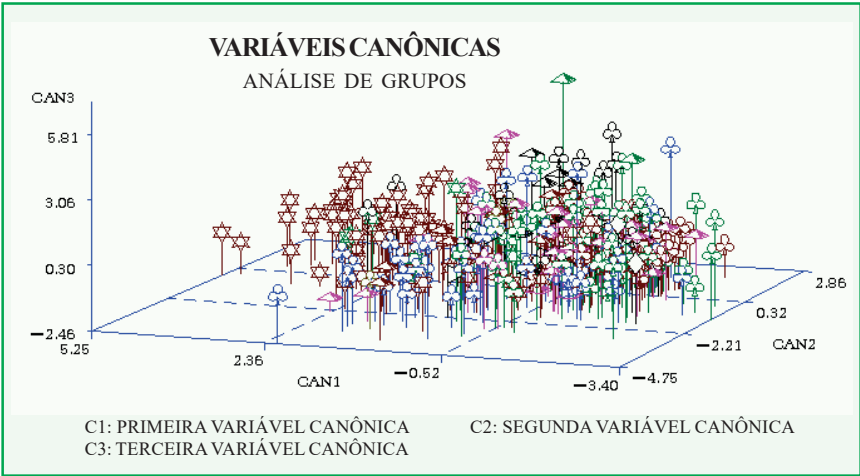


Figura 2 -Dispersão gráfica de acessos segundo variáveis canônicas formadas a partir das variáveis: peso de uma semente seca, peso de sementes úmidas por fruto, peso total de fruto, peso de semente seca por fruto e relação semente seca / semente úmida.

Relações entre o desempenho “per se” de clones e a capacidade geral de combinação

Para o estudo da importância da avaliação de germoplasma em respeito à seleção de genótipos para a introdução em programas de melhoramento, ou da eficiência da seleção de clones para processos recorrentes de melhoramento, foram analisados, a princípio, para as variáveis PSU, PSS e RSSSU, dois ensaios de avaliação de cultivares híbridas, denominados no CEPEC de ensaios 4 e 32. Para estes as distribuições de frequências seguiram padrão de curvas normais sem desvio importante (respectivamente, D = 0,027 e 0,032), embora tenham sido identificadas como distintas da normal pelo teste de Kolmogorov-Smirnov e, apesar da significância do teste de Bartlett para a homogeneidade das variâncias, não houve distinções maiores que 5,1 vezes entre a menor e maior variância dentro de tratamento para as variáveis consideradas, nos dois ensaios.

O ensaio 4 foi analisado tendo-se como fontes de variação a época, o bloco, o grupo 1 (clones: DR2, ICS3 e UF 613), o grupo 2 (progenitores da série SIC), a interação grupo 1 x grupo 2, semente dentro de grupo 1 x grupo 2, e as interações: época x grupo 1, época x grupo 2, época x grupo 1 x grupo 2 e época x semente dentro de grupo 1 x grupo 2 (PROC GLM – SAS System –SAS Institute, 1988).

Considerados os efeitos ajustados apenas para aqueles que os precedem, foram observados efeitos significativos de época, grupo 1 e grupo 2 (a 0,01%) para PSU, PSS e RSSSU, mostrando a possibilidade de identificação de diferenças para a capacidade geral de combinação mesmo em grupos de clones com grande uniformidade, como é o caso das seleções locais. Foi, ainda, constatada significância (a 0,01% de probabilidade) para as interações grupo 1 x grupo 2, ou capacidade específica de combinação, para as três variáveis; e todas as interações foram significativas (a 5% para época x grupo 1, com a variável PSS e a 0,01% para as demais interações e efeitos), a exceção daquelas que incluíram o efeito de semente dentro de grupo 1 x grupo 2. Não foram identificadas diferenças significativas entre sementes dentro de progênie (grupo 1 x grupo 2).

Tomando-se as médias ajustadas de progenitores, obtidas com a redução do modelo, pela desconSIDERAÇÃO das interações e efeito de planta, verificou-se significância (a 5%) para a correlação entre estes valores e os obtidos com as avaliações dos clones “per se” apenas para o caso da variável RSSSU com o grupo 2 (0,68). Assim pequenas diferenças observadas na avaliação de germoplasma, como as entre os clones SIC para PSU e PSS não permitiram a previsão das diferenças quanto à capacidade geral de combinação, ao contrário do que ocorre quando de distinções mais nítidas, como será mostrado com o ensaio 32.

Este foi analisado em modelo com blocos, grupo 1 (Ca 4, Ma 15, ICS 1 e UF 613), grupo 2 (IMC 67, Pound 12 e Pa 150) e interação grupo 1 x grupo 2 como fontes de variação. Houve diferenças significativas entre os progenitores do grupo 2 para PSU, PSS e RSSSU (0,01% de probabilidade) e entre os do grupo 1 para PSU e PSS (0,01%). As interações foram significativas a 0,01% para PSU e PSS e a 5% para RSSSU. Neste caso, as correlações de Pearson entre as médias de progenitores (obtidas sem a necessidade de redução do modelo) e as da avaliação “per se” foram significativas (a 10%), para o grupo 1 com as variáveis PSU (0,94) e PSS (0,93) (para o grupo 2, cujos clones tem menores diferenças de PSU e PSS as correlações foram não significativas).

Desta forma foi possível observar que distinções expressivas observadas entre genótipos avaliados na coleção de germoplasma podem ser indicadores eficientes de diferenças quanto à capacidade geral de combinação, comprovando a importância deste processo de avaliação.

A relação entre o desempenho “per se” de clones e o comportamento destes como progenitores também foi avaliada para a variável PSUF. Para tal, foram considerados dados dos ensaios 1, 33, 34, 35, 37, 40, 41 e 42 (Monteiro et al, 1995). Todos mostraram distribuição diferente da normal pelos testes de Shapiro-Wilk e Kolmogorov-Smirnov, mas as curvas de distribuição de frequência não mostraram qualquer desvio importante (na ordem, $W = 0,948$; $D = 0,020$; $0,051$; $0,050$; $0,057$ e para o conjunto 40-41-42, $D = 0,055$). Foi constatada heterogeneidade de variância dentro de tratamentos pelo teste de Bartlett, a 5%, para todos os ensaios a exceção do 1, mas os tratamentos considerados (listados anteriormente) apresentaram variâncias cujas diferenças não superam o valor 7. As relações entre maiores e menores variâncias foram de 6,38 para o ensaio 1 (na verdade, neste caso, combinações com um quinto progenitor, o ICS 1, pertencente ao desenho original, que conduziam a não homogeneidade pelo teste de Bartlett, tiveram que ser desconSIDERADOS para manter as diferenças entre variáveis dentro do limite de 7 vezes); de 1,8 para o ensaio 33; de 2,86 para o ensaio 34; de 2,87 para o ensaio 35; de 2,13 para o ensaio 37 e de 3,32 para a análise conjunta dos ensaios 40, 41 e 42.

Houve efeito significativo de ano (a 0,01%) para os ensaios com esta fonte de variação (ensaios 1, 33, 34, 35 e 37); as interações ano x mãe foram significativas (a 0,01%) nos ensaios 33, 35 e 37 e as interações ano x pai e ano x pai x mãe foram não significativas (a 5% de probabilidade) nos 5 ensaios. Os ensaios com dados por planta mostraram efeito significativo a 0,01% para planta dentro de pai x mãe (33, 34, 35 e 37) ou planta dentro de tratamento (conjunto 40, 41, 42). Com o ensaio 1, foram observados efeitos significativos a 0,01% para mãe e pai (neste caso o conjunto de mães e pais é o mesmo, pois temos aqui um dialelo completo) e interação mãe x pai. O mesmo foi observado para os ensaios 34, 35 e 37. Com o ensaio 33 verificou-

se significância para as diferenças entre mães (a 0,01% de probabilidade), ausência de efeitos significativos de pais (a 5%) e efeito significativo da interação mãe x pai (a 1%).

Para os ensaios com dados tomados durante mais de um ano (ensaios 1, 33, 34, 35 e 37), a análise do PSUF como medidas repetidas (PROC GLM - MANOVA, para a interpretação dos efeitos de pai, mãe e interação pai x mãe e PROC GLM REPEATED para os efeitos de ano e interação que incluem este fator - SAS System, SAS Institute, 1988) levou a resultados análogos. As diferenças dignas de nota foram: para o ensaio 33 a interação ano x pai x mãe foi significativa a 5% pelos teste de Wilks; e para o 34 foi significativa, a 5%, a interação ano x pai.

Para o grupo de ensaios 40-41-42 foram observados efeitos significativos de ensaio (a 0,01%), ausência de efeitos significativos da interação entre tratamentos comuns e ensaios (a 1% de probabilidade) e diferenças significativas entre tratamentos (a 0,01%). Analisado pela metodologia de Griffing este conjunto de ensaios mostrou diferenças amplamente significativas para a capacidade geral de combinação das mães e pais (valor de F superior a 20.000) e para a capacidade específica de combinação (valor de F superior a 15.000, com 12.693 graus de liberdade do resíduo).

Desta forma, constata-se a possibilidade de identificação de diferenças quanto à capacidade geral de combinação mesmo para grupos ou populações mais uniformes como a representada pelas séries SIC, SIAL e EEG (exemplo ensaios 34, 35 e 37) e há, de modo geral, amplos efeitos em relação às combinações específicas.

Quanto à predição da capacidade geral de combinação pela avaliação “per se” dos clones, foram observadas correlações altas entre médias geradas pela avaliação de germoplasma e médias de progenitores dos diversos ensaios. Estas correlações tiveram valores de 0,99 (Pearson, significativa a 1% de probabilidade) para o grupo de pais dos ensaios 40, 41 e 42, que incluem dois clones Trinitários com maiores valores e dois locais, e de 0,87 (significativa a 1%) para o grupo de mães que contempla apenas materiais da Amazônia brasileira. Para os demais ensaios, cujos grupos de progenitores, em sua maioria, têm número menor

de representantes, o que torna as comparações menos precisas, foram observados os valores: 0,85 para o grupo de mães do ensaio 37 (não significativo – 3 representantes, todos do tipo comum) e 1,0 para o grupo de pais (significativo a 1% - 3 representantes); 0,75 para o grupo de mães do ensaio 34 (significativo a 10% - 8 representantes, sendo todos seleções obtidas dentro das variedades tradicionais da Bahia) e 1,0 para pais (não significativo - apenas dois representantes); 0,84 para o grupo de mães do ensaio 33 (não significativo – 3 representantes), para o qual não houve diferenças significativas entre médias de pais; e 0,88 para o grupo de progenitores do ensaio 1 que é um dialelo completo (não significativo - 4 representantes). Apenas para o ensaio 35 os valores das correlações foram inferiores a 0,7: 0,58 para as 5 mães e 0,50 para os 3 pais (estes últimos são seleções locais e não há dados de avaliação “per se” de um quarto pai: o SIAL 169).

Assim, considerando estes resultados conjuntamente, e tendo-se em conta a tendência geral de valores superiores a 0,7 para os coeficientes de correlação, podemos considerar que a avaliação de germoplasma para o caráter é um processo importante na seleção de genótipos para serem introduzidos em programas de melhoramento.

Conclusões

Foi possível constatar que: há limitações nos processos tradicionais de caracterização de germoplasma de cacau para os descritores de fruto e semente; a consideração de duas amostras de 40 sementes em um único período de safra pode ser sugerida como um método rápido e de baixa demanda de esforços para a avaliação do peso da semente em coleções de germoplasma, e para a análise do peso de frutos e semente úmidas por fruto, pode ser sugerida a utilização de 4 ou 5 amostras em dois períodos de safra; frutos maiores e com maiores pesos de sementes tendem a ter sementes maiores, o que simplifica o melhoramento para estes caracteres em conjunto; a maior parte dos grupos importantes quanto a possibilidade de fornecer genes de resistência à vassoura-de-bruxa e à podridão parda e para o incremento do teor de gordura possuem frutos e

sementes pequenas, que são maiores em populações domesticadas e híbridos; os caracteres considerados podem ser usados como elementos de caracterização de materiais e em conjunto propiciam clara separação de séries de diferentes origens; são identificáveis diferenças para a capacidade geral de combinação para características físicas frutos e sementes mesmo dentro de grupos de clones com elevada uniformidade; há efeitos expressivos da capacidade específica de combinação para o peso de uma semente e peso de sementes por fruto; e distinções entre genótipos ‘per se’ podem ser indicadores eficientes de diferenças quanto à capacidade geral de combinação, para o peso de sementes e o peso de sementes por fruto, comprovando a importância desta avaliação para a seleção de progenitores e comprovando a possibilidade de melhoramento genético para estes caracteres.

Literatura Citada

- ALMEIDA, C. M. V. C. de et al. 1987. Evolução do programa de conservação de recursos genéticos de cacau na Amazônia brasileira. CEPLAC/DEPEA/ Belém, PA. Boletim Técnico n. 5. 108 p.
- BARTLEY, B. G. D. 2001. Refractorio – An explanation of the meaning of the term and its relationship to the introductions from Ecuador in 1937. INGENIC Newsletter (UK) 6: 10-15.
- CRUZ, C. D. 1997. Programa GENES; aplicativo computacional em genética e estatística. Viçosa, MG, UFV. 442p.
- CRUZ, C. D.; REGAZZI, A. J. 1997. Modelos biométricos aplicados ao melhoramento genético. Viçosa, MG, UFV. 390p.
- DANFORD, M. B. et al. 1960. On the analysis of repeated measurements experiments. Biometrics (UK) 16:547-565.
- ENGELS, J. M. M.; Bartley, B. G.; Enriquez, G. A. 1980. Cacao descriptors, their states and modus operandi. Turrialba 30 (2): 209-218.
- INTERNATIONAL COCOA GERMPLASM DATABASE - ICGD. 2024. Reading: University of Reading. Available at: <<http://www.icgd.rdg.ac.uk>>. Accessed on: July 2024.
- JOHNSON, R. A.; WICHERN, D. W. 1999. Applied Multivariate Statistical Analysis. New Jersey: Prentice Hall. 816p.
- LYNCH, M.; WALSH, B. 1988. Genetics and Analysis of Quantitative Traits. Sinauer Associates, Massachusetts Inc. 980p.
- LITTELL, R. C. et al. 1996. SAS System for Mixed Models. Cary, NC: SAS Institute Inc. 633p.
- MARTIN, F. G. 1995. Statistical design and analysis. Gainesville, FL, University Copy Center.
- MONTEIRO, W. R. et al. 1995. Desenvolvimento e avaliação de cultivares híbridas de cacau. Informe de Pesquisas, 1987/90, CEPLAC, Ilhéus, BA, Brasil. p.19-23.
- PEREIRA, M. G. et al. 1994. Influência do progenitor paterno sobre caracteres físicos e químicos da semente, relacionados com a qualidade do cacau. Agrotrópica (Brasil) 6(2): 31-40, 1994.
- PIRES, J. L. et al. 1995a. Avaliação de caracteres relacionados com a qualidade de amêndoas de cacau – Ensaio 04. Informe de Pesquisas, 1987/90, CEPLAC, Ilhéus, BA, Brasil. p.11-12.
- PIRES, J. L. et al. 1995b. Avaliação de caracteres relacionados com a qualidade de amêndoas de cacau – Ensaio 32. Informe de Pesquisas, 1987/90, CEPLAC, Ilhéus, BA, Brasil. p.13-15.
- PIRES, J. L. et al. 1997. Resistência em campo de clones de cacauceiro à podridão parda, causada por *Phytophthora spp.* na Bahia, Brasil. Fitopatologia Brasileira 22: 375-380.
- PIRES, J. L. et al. 1998. Increasing cocoa butter yield through genetic improvement of *Theobroma cacao* L.: Seed fat content variability, inheritance and association with seed yield. Euphytica (USA) 103, 115-121.
- PIRES, J. L.; LUZ, E. D. M. N.; MELO, G. R. P. 2021. New sources of resistance - Cacao germplasm evaluation for natural incidence of witches' broom. Agrotrópica (Brasil) 33(2):91-100.

PIRES, J. L. et al. 2022. Identification of resistance sources to Witches' broom disease – Cacao germplasm evaluation. *Agrotropica (Brasil)* 34(3): 169-180.

SAS INSTITUTE INC. 1988. SAS/STAT User's Guide. Release 6.03. Cary, NC: SAS Institute Inc, 1028p.

