

## Aceleração do melhoramento do cafeeiro via seleção genômica: agilidade e eficácia no lançamento de novas cultivares



**Empresa Brasileira de Pesquisa Agropecuária  
Embrapa Café  
Ministério da Agricultura, Pecuária e Abastecimento**

## **DOCUMENTOS 17**

# Aceleração do melhoramento do cafeeiro via seleção genômica: agilidade e eficácia no lançamento de novas cultivares

*Eveline Teixeira Caixeta  
Marcos Deon Vilela Resende  
Emilly Ruas Alkimim  
Tiago Vieira Sousa  
Antonio Carlos Baião de Oliveira  
Antonio Alves Pereira  
Rodrigo Silva Alves*

**Embrapa Café**  
Brasília, DF  
2022

Exemplares desta publicação podem ser adquiridos na:

**Embrapa Café**  
Parque Estação Biológica (PqEB)  
Av. W3 Norte (final), Ed. Sede  
CEP: 70770-901, Brasília, DF  
Fone: +55 (61) 3448-4378 / 4010  
Fax: +55 (61) 3448-1797  
www.embrapa.br  
www.embrapa.br/fale-conosco/sac

Comitê Local de Publicações da Embrapa Café

Presidente  
*Lucas Tadeu Ferreira*  
Vice-Presidente  
*Jamilsen de Freitas Santos*  
Secretária-Executiva  
*Adriana Maria Silva Macedo*

Membros  
*Anísio José Diniz, Carlos Henrique Siqueira de Carvalho, Helena Maria Ramos Alves, Lucilene Maria de Andrade, Mauricio Sergio Zacarias, Milene Alves de Figueiredo Carvalho, Omar Cruz Rocha, Rogério Novais Teixeira, Roseane Pereira Villela.*

Revisão de texto  
*Francisca Eljani do Nascimento*

Normalização bibliográfica  
*Maria de Fátima da Cunha*

Tratamento das ilustrações  
*Thiago Farah Cavaton*

Projeto gráfico da coleção  
*Carlos Eduardo Felice Barbeiro*

Editoração eletrônica  
*Thiago Farah Cavaton*

Fotos da capa  
*Eveline Teixeira Caixeta*  
*Antonio Carlos Baião de Oliveira*

**1ª edição**  
Publicação digital (2022): PDF

**Todos os direitos reservados.**

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).

**Dados Internacionais de Catalogação na Publicação (CIP)**

Embrapa, Superintendência de Serviços Compartilhados

---

Aceleração do melhoramento do cafeeiro via seleção genômica : agilidade e eficácia no lançamento de novas cultivares / Eveline Teixeira Caixeta ... [et al.].  
– Brasília, DF : Embrapa Café, 2022.  
PDF (54 p.). – (Documentos / Embrapa Café, ISSN 1678-1694 ; 17).

1. Café. 2. *Coffea arabica*. 3. *Coffea canephora*. 4. Hemileia vastatrix. 5. Marcador molecular. 6. Melhoramento genético vegetal. I. Caixeta, Eveline Teixeira. II. Resende, Marcos Deon Vilela. III. Alkimim, Emilly Ruas. IV. Sousa, Tiago Vieira. V. Oliveira, Antonio Carlos Baião de. VI. Pereira, Antonio Alves. VII. Alves, Rodrigo Silva. VIII. Série.

CDD 633.73

## Autores

### **Eveline Teixeira Caixeta**

Engenheira-agrônoma, doutora em Genética e Melhoramento, pesquisadora da Embrapa Café, Brasília, DF.

### **Marcos Deon Vilela Resende**

Engenheiro-agrônomo, doutor em Genética, pesquisador da Embrapa Café, Brasília, DF.

### **Emilly Ruas Alkimim**

Engenheira-agrônoma, doutora em Genética e Melhoramento, professora da Universidade Federal do Triângulo Mineiro, Iturama, MG.

### **Tiago Vieira Sousa**

Engenheiro-agrônomo, doutor em Genética e Melhoramento, professor do Instituto Federal do Triângulo Mineiro, Campina Verde, MG.

### **Antonio Carlos Baião de Oliveira**

Engenheiro-agrônomo, doutor em Genética e Melhoramento, pesquisador da Embrapa Café, Brasília, DF.

### **Antonio Alves Pereira**

Engenheiro-agrônomo, doutor em Fitopatologia, pesquisador da Empresa de Pesquisa Agropecuária de Minas Gerais (Epamig), Viçosa, MG.

### **Rodrigo Silva Alves**

Engenheiro florestal, doutor em Genética e Melhoramento, pós-doutorado na Universidade Federal de Viçosa, MG.



## Agradecimentos

Agradecemos ao Consórcio Pesquisa Café, à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (Fapemig), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e ao Instituto Nacional de Ciência e Tecnologia do Café (INCT-Café) pelo suporte financeiro.



## Apresentação

Os programas de melhoramento genético de plantas têm tido impacto significativo na produção agrícola. Apesar do sucesso acumulado desses programas, ganhos adicionais na eficiência do melhoramento podem ser obtidos por meio de aplicação de tecnologias moleculares. Inicialmente, o melhoramento genético vegetal se baseava exclusivamente nas características morfoagronômicas para selecionar indivíduos nas populações segregantes. Nos anos 80, marcadores moleculares foram desenvolvidos e passaram a ser utilizados como ferramenta auxiliar às informações fenotípicas, permitindo avanços significativos em diversos programas de melhoramento.

Atualmente, com o desenvolvimento de metodologias de genotipagem e sequenciamento em larga escala, os programas de melhoramento estão incorporando gradualmente essas novas tecnologias. Destaque tem sido dado à Seleção Genômica Ampla. Por meio dessa técnica, é possível estimar o valor genético dos indivíduos, o que possibilita aumentar os ganhos com seleção por unidade de tempo, sendo os seus benefícios demonstrados em vários estudos, inclusive em caracteres poligênicos, com baixa herdabilidade, de difícil mensuração e avaliação dispendiosa.

Em plantas perenes, de ciclos longos, como o cafeeiro, a aplicabilidade da Seleção Genômica Ampla assume importância acentuada. Essa estratégia permite aumentar os ganhos genéticos, para vários caracteres agrônômicos, por unidade de tempo, o que possibilita reduzir o tempo de lançamento de cultivares de cafés.

Neste documento, foram demonstrados os usos da Seleção Assistida por Marcadores Moleculares e da Seleção Genômica nas espécies *Coffea arabica* e *Coffea canephora*. O uso dessas técnicas seletivas possibilitou o aumento na eficiência nos programas de melhoramento dessas espécies. Além disso, evidencia a importância da utilização de diferentes metodologias de seleção, sendo explorado os benefícios de cada técnica seletiva, tornando os programas de melhoramento mais eficientes e competitivos.

*Antonio Fernando Guerra*  
Chefe-Geral da Embrapa Café



## Sumário

Introdução .....	nº 11
Melhoramento genético do cafeeiro .....	nº 14
Seleção assistida por marcadores moleculares .....	nº 17
Seleção genômica ampla .....	nº 22
Métodos de seleção genômica.....	nº 27
Capacidade preditiva, viés da predição e acurácia da GWS .....	nº 30
Aumento da eficiência seletiva do melhoramento de plantas .....	nº 31
GWS em <i>Coffea arabica</i> .....	nº 32
Material genético .....	nº 32
Avaliações fenotípicas.....	nº 32
Análise de qualidade de marcadores moleculares.....	nº 34
Eficiência da GWS .....	nº 34
Resultados e discussão .....	nº 34
GWS em <i>Coffea canephora</i> .....	nº 39
Material genético .....	nº 39
Avaliações fenotípicas .....	nº 39
Resultados e discussão .....	nº 40
Conclusões .....	nº 45
Referências .....	nº 46



## Introdução

No histórico do melhoramento do cafeeiro arábica (*Coffea arabica*) no Brasil, os primeiros trabalhos de seleção foram iniciados em 1932 e, até a década de 1960, os objetivos eram direcionados ao desenvolvimento de cultivares com alta produção, vigor, longevidade e adaptadas às diferentes regiões do país. Após 1970, com o aparecimento da ferrugem do cafeeiro e sua rápida dispersão nas lavouras brasileiras, enorme ênfase foi dada ao melhoramento para resistência a *Hemileia vastatrix*, agente causal dessa doença, e novos programas de melhoramento foram iniciados nas principais regiões produtoras do Brasil. A partir de então, resistência à ferrugem e a outras doenças passou a ser um dos principais objetivos dos programas de melhoramento dessa espécie.

Além de cultivares resistentes, os avanços alcançados pelo melhoramento genético do cafeeiro também têm permitido a obtenção e recomendação de materiais genéticos adaptados a diferentes regiões do país. Essas cultivares apresentam importantes características como: produtividade elevada, estável e com precocidade na primeira colheita; maturação uniforme; tolerância à geada, à seca e ao alumínio tóxico; porte baixo e formato de copa adequado para colheita mecanizada; arquitetura adequada para adensamento; e elevado tamanho de grão e qualidade de bebida (Medina Filho et al., 2008). Atualmente, além de características morfoagronômicas, o melhoramento do cafeeiro tem focado na qualidade de bebida, para atender um mercado e consumidores cada vez mais exigentes.

Apesar de a maioria dos programas de melhoramento brasileiro ter foco principal no desenvolvimento de cultivares de *C. arabica*, a crescente importância da espécie *C. canephora* (café canéfora) tem levado ao interesse em também obter cultivares dessa espécie. Dentro do gênero *Coffea*, as duas espécies de maior importância comercial são *C. arabica* e *C. canephora*. O café arábica é responsável pelo aroma e sabor adocicado, proporcionando bebida de melhor qualidade. O café canéfora apresenta maiores quantidades de cafeína, sólidos solúveis e oferece corpo à bebida (Babova et al., 2016). O cultivo comercial do café canéfora no Brasil foi impulsionado a partir dos anos 1950, com o advento dos cafés solúveis e, posteriormente, com o seu uso no café torrado e moído, em misturas (*blend*) com o café arábica. Esse aumento na

demanda pelos grãos de *C. canephora* estimulou a produção brasileira e, conseqüentemente, a expansão do parque cafeeiro nacional e o desenvolvimento de novas cultivares.

Apesar do sucesso alcançado, a obtenção de uma cultivar, de *C. arabica* ou *C. canephora*, não é uma atividade simples e demanda acúmulo de amplo conhecimento sobre a cultura e sobre as características que nela se deseja melhorar. Além disso, o processo de melhoramento genético consiste em um trabalho de longo prazo e, muitas vezes, quando se desenvolve uma cultivar, outros problemas aparecem, fazendo com que se estabeleça um constante desafio aos melhoristas de sempre buscar novas cultivares. Deve-se considerar também que, dependendo da espécie a ser trabalhada, especialmente as perenes como o cafeeiro, muitas vezes o surgimento de novos desafios agrônômicos é mais dinâmico do que a capacidade de resposta dos programas de melhoramento. Diante disso, faz-se então necessário adicionar novas tecnologias capazes de incrementar a dinâmica e a capacidade de resposta dos programas de melhoramento.

Os avanços da biotecnologia ocorridos nos últimos anos certamente concorrerão para dinamizar esse processo (Sousa et al., 2017). O potencial da biotecnologia no melhoramento está não só na redução do tempo de condução do programa, mas também na base científica sólida que pode explicar a genética e a bioquímica das mudanças que ocorreram ou que poderão ocorrer no processo de melhoramento genético. Mesmo não conhecendo todos os componentes genéticos envolvidos em uma característica, marcadores moleculares podem ser importantes para manipular genes ou blocos gênicos desejáveis, com maior precisão e rapidez. As novas tecnologias, com destaque ao uso de marcadores moleculares, vêm então se associar aos procedimentos de melhoramento denominados tradicionais, dotando os mesmos de maior versatilidade na solução dos mais variados problemas agrônômicos.

Além dos marcadores moleculares, diferentes ferramentas moleculares estão sendo continuamente desenvolvidas e incorporadas no melhoramento genético de plantas (Savadi et al., 2018). Os avanços das técnicas moleculares, da bioinformática, do sequenciamento, bem como da genotipagem e fenotipagem de alto rendimento resultaram em uma nova era do melhoramento genético. A utilização dessas ferramentas tem permitido a obtenção mais rápida de cultivares que atendam às necessidades dos produtores, das indústrias

e do mercado tanto nacional quanto internacional (Sousa et al., 2019). As abordagens mais comumente usadas são a seleção assistida por marcadores moleculares (SAM) e seleção genômica ampla (GWS – *Genome Wide Selection*).

Dentro desse contexto, atenção especial deve ser dada ao melhoramento do cafeeiro, que, por ser espécie perene e de ciclo produtivo longo, demanda agilidade e eficácia no lançamento de novas cultivares. Para alcançar esse objetivo, a otimização do melhoramento e a maximização da eficiência seletiva são essenciais na rápida e acurada criação de genótipos superiores. A seleção genômica ampla (GWS – *Genome Wide Selection*) ou seleção genômica (GS – *Genome Selection*) foi proposta em 2001 como uma forma de aumentar a eficiência e acelerar o melhoramento genético. A GWS enfatiza a predição simultânea, sem o uso de testes de significância para marcas individuais, dos efeitos genéticos de milhares de marcadores genéticos de DNA dispersos em todo o genoma de um organismo. Essa estratégia permite capturar os efeitos de todos os locos, tanto de pequenos quanto de grandes efeitos, e explicar toda a variação genética de um caráter quantitativo. Dessa forma, a GWS, aplicada via genotipagem em larga escala de plântulas jovens, permite uma eficiente seleção precoce, conduzindo a um melhoramento rápido e efetivo. No caso, a soma dos efeitos genéticos estimados dos marcadores presentes em um indivíduo fornece o valor genético dos indivíduos para fins de seleção.

Os estudos de seleção genômica em cafeeiro são ainda incipientes e o acúmulo de informações nessa área são importantes para obtenção de mais informações e para subsidiar ações mais efetivas no melhoramento para produção cafeeira no país e maior sustentabilidade do sistema produtivo associado à cadeia do café. Os resultados preliminares obtidos têm sido animadores, conduzindo a uma boa acurácia seletiva e maior ganho genético por unidade de tempo.

O presente trabalho refere-se a uma atualização da avaliação do potencial da GWS no melhoramento dos cafeeiros arábica e canéfora. O objetivo geral é avaliar a eficiência da GWS em caracteres produtivos, agrônômicos e de resistência a doenças e pragas do cafeeiro. Foram avaliadas as acurácias seletivas, ganho genético por ciclo de seleção e por unidade de tempo (anos). Os objetivos específicos são discorrer sobre a mudança estrutural nos pro-

gramas de melhoramento de cafeeiro, decorrente da seleção genômica sobre o desenvolvimento de métodos otimizados de predição genômica no cafeeiro e sobre como realizar a prática da seleção genômica.

Na avaliação da metodologia, foram utilizados dados de experimentos conduzidos em várias safras, pela Empresa de Pesquisa Agropecuária de Minas Gerais (Epamig) – Embrapa Café, instalados e avaliados em Viçosa, MG e Oratórios, MG, e dados de genotipagem realizada para milhares de marcadores SNP. Também foram utilizados resultados da literatura no contexto da aceleração do melhoramento do cafeeiro via seleção genômica visando subsidiar a seleção de genótipos superiores portadores de genes que governam as principais características agrônômicas, usando as metodologias otimizadas.

## Melhoramento genético do cafeeiro

O cultivo do cafeeiro no Brasil iniciou-se com introduções de algumas poucas plantas, em 1727. Sementes de uma planta da espécie *C. arabica* que estava no Jardim Botânico de Amsterdam foram enviadas ao Suriname e às Guianas, de onde o café foi trazido ao Brasil (Carvalho, 2007). Essas plantas foram denominadas de cultivar Typica, cafeeiro que apresentava baixa produtividade. Essa foi, praticamente, a única cultivar plantada no Brasil até meados do século XIX. Visando ampliar a base genética do cafeeiro produzido no Brasil, em 1859, foi introduzida a cultivar Bourbon Vermelho, proveniente da Ilha de Reunião. Novas introduções foram realizadas, como as cultivares Sumatra e Híbrido de Timor, que foram e ainda são largamente exploradas no melhoramento genético (Mendes et al., 2008; Pereira et al., 2010). Outras introduções foram realizadas, mas o material genético não despertou interesse agrônômico, como as cultivares Murta, Polysperma, Laurina, Mokka, entre outras.

Até início da década de 1930, as cultivares plantadas no Brasil foram resultantes de introduções, raras mutações ou recombinações resultantes de hibridações naturais entre as plantas existentes (Carvalho, 2007). A partir dessa data, com a participação do Instituto Agrônomo (IAC), o melhoramento genético do cafeeiro passou a ser conduzido com base em informações científicas e então abrangiam pesquisas básicas de taxonomia, evolução, citologia, biologia de reprodução, anatomia, análise genética e

morfologia. O programa de melhoramento foi conduzido por meio de seleção de progênies e hibridações (Carvalho, 1952). Nessa fase, o melhoramento do cafeeiro arábica era realizado utilizando-se basicamente dois métodos: (1) seleção de linhagens em população segregante originadas de cruzamentos naturais; e (2) hibridações seguidas de condução da população segregante pelo método genealógico. Os objetivos eram direcionados, principalmente, ao desenvolvimento de cultivares com alta produção, vigor, longevidade e adaptadas às diferentes regiões do país.

A partir de 1970, com o aparecimento da ferrugem, doença causada pelo fungo *Hemileia vastatrix*, grande impulso foi dado ao melhoramento genético do cafeeiro, sendo implantados programas de melhoramento em diferentes regiões brasileiras produtoras de café. A partir de então, resistência à ferrugem e a outras doenças passou a ser um dos principais objetivos dos programas de melhoramento dessa espécie. Ampliação dos métodos de melhoramento também ocorreu, sendo aplicados principalmente a introdução, a seleção de plantas individuais seguida de teste de progênie, o método genealógico, os retrocruzamentos, a seleção recorrente e a hibridação interespecífica. Além de resistência à ferrugem e a outras doenças, o objetivo dos métodos consistiu em desenvolver cultivares adaptadas às diferentes regiões cafeeiras e aos sistemas de cultivo, com elevada produtividade de grãos, épocas diferenciadas de colheita e tolerância a condições adversas de meio ambiente.

Outra característica que tem sido foco dos programas de melhoramento do cafeeiro arábica é a qualidade de bebida. A demanda por cafés com melhores qualidades, chamados cafés especiais, tem aumentado consideravelmente no mercado nacional e internacional. Dessa forma, na seleção dos cafeeiros, os melhoristas passaram também a buscar qualidade física e sensorial dos grãos (Barbosa et al., 2019). Nesse contexto, programas de melhoramento do café em todo o mundo buscam recursos genéticos, principalmente no que diz respeito aos atributos sensoriais do café para a produção de cafés diferenciados voltados ao mercado de cafés especiais (Tessema et al., 2011; Sobreira et al., 2016). Dessa forma, na seleção de genótipos superiores, melhoristas estão focando em uniformidade de maturação, ciclo de maturação, aumento do tamanho dos grãos, conversão do café beneficiado, redução de grãos mocas e outras características que afetam a qualidade do café (Barbosa et al., 2019).

Mais recentemente, o amplo desenvolvimento tecnológico da cafeicultura brasileira resultou na implementação da mecanização em diversas operações de campo, demandando cultivares de café arábica adaptadas a essa nova condição. Dessa forma, características que facilitam a mecanização, como estrutura, tamanho e forma de plantas, passaram a ser também buscadas no melhoramento. Além do rendimento, genótipos adaptados para mecanização necessitam apresentar facilidade de desprendimento do fruto, menor desfolha e vigor vegetativo para aumentar a eficiência na colheita mecanizada dos frutos (Dias et al., 2020).

Além de todos esses desafios enfrentados pelos melhoristas, nas últimas décadas, a obtenção de cultivares de café arábica melhoradas tem ganhado maiores interesses, que pode ser resultado da nova crise de ferrugem que ocorreu na América Central e Latina no início de 2012. Essa crise pode ter despertado para a vulnerabilidade da cultura, que tem sido formada por lavouras contendo cultivares suscetíveis à ferrugem e pouco produtivas. Além disso, a crescente evidência do impacto das mudanças climáticas no futuro da cafeicultura e o crescimento do mercado de cafés especiais têm motivado o uso de novas cultivares (Avelino et al., 2015; Pruvot-Woehl et al., 2020).

A outra espécie de *Coffea* de interesse agrônômico, *C. canephora*, também tem sido largamente trabalhada nos programas de melhoramento. O cultivo e o melhoramento de cafeeiros da espécie *C. canephora* foram provavelmente estimulados pela constatação de grande incidência da ferrugem nas lavouras de *C. arabica*, uma vez que o cafeeiro canéfora, em geral, apresentava resistência às doenças. O cultivo foi intensificado com o surgimento do café solúvel e seu uso em *blends* de café torrados e moídos (Ferrão et al., 2019). O melhoramento realizado nos cafeeiros canéfora são bem distintos do arábica, pois essas espécies diferem nos aspectos de sistemas de reprodução e propagação, número de cromossomos e base genética. *C. arabica* é autógama, alotetraploide ( $2n=4x=44$ ) e de base genética estreita, enquanto *C. canephora* é alógama, autocompatível, diploide ( $2n=2x=22$ ) e apresenta grande diversidade genética. Dessa forma, as principais estratégias que vêm sendo utilizadas nos programas de melhoramento para *C. canephora* são introdução de germoplasma, seleção clonal, hibridação intra e interespecífica e seleção recorrente (Ferrão et al., 2019).

Apesar da diferença na obtenção de cultivares, os objetivos gerais nos quais se baseiam os programas de melhoramento das duas espécies são semelhantes. As cultivares melhoradas apresentam aumento na capacidade produtiva das plantas, melhoria na arquitetura, resistência a estresse biótico e abiótico e bebida de qualidade superior. O objetivo é atender às exigências dos produtores e consumidores, bem como garantir a sustentabilidade da atividade e promover maior retorno socioeconômico para a cafeicultura e para a sociedade (Ferrão et al., 2019).

Atualmente, como resultados de extensos programas de melhoramento genético, já se dispõe de várias cultivares resistentes a doenças e pragas, com produção elevada e estável e com precocidade da primeira colheita. Essas cultivares são também portadoras de outras características de interesse, como alto vigor vegetativo, adaptação específica a diferentes regiões cafeeiras, adaptação ao cultivo sob irrigação, tolerância à seca, porte baixo e formato de copa adequado para a colheita manual e mecânica, maturação uniforme, boa qualidade de bebida e de grãos.

Apesar de extenso trabalho de melhoramento já ter sido realizado para o cafeeiro, o potencial do germoplasma disponível ainda não foi totalmente explorado. O recente progresso na área de biotecnologia, particularmente com os marcadores moleculares e novas plataformas de sequenciamento de DNA, tem permitido a descoberta de novos genes e acelerado os programas de melhoramento do cafeeiro (Mishra, 2019).

## Seleção assistida por marcadores moleculares

Marcadores moleculares permitem a detecção de variações nas sequências do DNA entre indivíduos de uma mesma espécie e de espécies distintas. Destacam-se pela capacidade de marcar um gene/região cromossômica desejada, por serem considerados estáveis e não serem afetados pelo ambiente e pelos efeitos pleiotrópicos e epistáticos (Adhikari et al., 2017). Dessa forma, marcadores moleculares têm sido eficientemente usados no melhoramento para a análise da variabilidade genética, uma vez que representam estratégia precisa para associar a variabilidade fenotípica à genotípica.

Uma das aplicações de maior impacto dos marcadores moleculares no melhoramento de plantas consiste em sua utilização como ferramenta auxiliar nos procedimentos de seleção, evidenciando que o uso das técnicas moleculares tem sido largamente empregado para fins de melhoramento assistido. Para aplicação das estratégias moleculares, mapas genéticos de ligação são construídos e são usados para identificação de marcadores associados às características de interesse agrônomo (Brito et al., 2010). Esses marcadores permitem a seleção eficiente e rápida de genótipos de interesse, o que possibilitou a aplicação da técnica de Seleção Assistida por Marcadores Moleculares (SAM) nos mais diversos programas de melhoramento genético. No entanto, para que seja eficiente, o marcador deve estar fisicamente próximo ou inserido no(s) gene(s) de interesse; ser altamente polimórfico, permitindo a discriminação dos diferentes genótipos; e consistir em metodologia rápida, simples e de custo baixo (Garrido-Cardenas et al., 2018).

A SAM, proposta por Lande e Thompson (1990), reúne dados fenotípicos e moleculares em ligação gênica próxima aos locos controladores de características qualitativas ou quantitativas (QTL – *Quantitative Trait Loci*). A SAM é baseada no conceito de que é possível inferir a presença de um gene pela presença de uma marca genética que está estreitamente ligada ao gene. O uso de marcadores moleculares e a incorporação da SAM no melhoramento de plantas trazem vantagens competitivas para os programas de melhoramento, sejam eles públicos ou privados. Algumas das vantagens e benefícios proporcionados pela técnica da SAM são: aceleração substancial do processo de seleção de genótipos de interesse e obtenção de uma nova cultivar, pois a seleção molecular pode ser realizada em qualquer fase de desenvolvimento da cultura; melhor entendimento nos estudos de herança; identificação e garantia da presença do alelo/QTL a ser introgridido e auxílio na seleção contra o genoma remanescente do genitor doador, visando acelerar as etapas de retrocruzamento; seleção de vários caracteres ao mesmo tempo e combinação de dois ou mais genes de interesse em um único genótipo (piramidação de genes); identificação de quebra de ligações gênicas indesejáveis; redução de recursos, pois o processo de seleção em cada geração é realizado com um número mínimo de plantas no campo; seleção indireta de características fenotípicas de difícil e/ou de alto custo de avaliação; e seleção visando ao melhoramento preventivo, como, por exemplo, na ausência do patógeno ou raça específica responsável pela ocorrência de determinada doença (Romero et al., 2014; Alkimim et al., 2017).

Para incorporar a SAM no melhoramento, inicialmente são identificados marcadores moleculares ligados ao(s) gene(s) de interesse e, por meio desses marcadores, a herança e a introgressão do(s) gene(s) pode(m) ser analisada(s) nos programas. Alguns marcadores moleculares ligados a genes conferindo resistência a doenças foram identificados em café, como resistência à ferrugem do cafeeiro (*Hemileia vastatrix* Berk. & Br.), ao nematoide das galhas (*Meloidogyne exigua*) e à *Coffee Berry Disease* – CBD (*Colletotrichum kahawae*) (Noir et al., 2003; Prakash et al., 2004; Gichuru et al., 2008; Mahé et al., 2008; Brito et al., 2010; Diola et al., 2011).

A resistência à ferrugem do cafeeiro é governada por pelo menos nove genes dominantes ( $S_H1$  a  $S_H9$ ), isoladamente ou em associação (Bettencourt; Rodrigues Júnior, 1988). Para essa doença, até o momento já foram identificados e disponibilizados marcadores ligados aos genes  $S_H3$ , derivado da introgressão de *C. liberica* em *C. arabica* (Prakash et al., 2004; Mahé et al., 2008), e ao gene  $S_H?$  (correspondente ao gene  $S_H7$ ,  $S_H8$  ou  $S_H9$  ou até mesmo outro gene desconhecido), provenientes de um acesso de Híbrido de Timor e que corresponde a um gene introgridido de *C. canephora* (Brito et al., 2010; Diola et al., 2011). Ligados ao gene  $S_H3$ , foram identificados 21 marcadores do tipo AFLP (*Amplified Fragment Length Polymorphism*) (Prakash et al., 2004). Quatro desses marcadores AFLP foram transformados com sucesso em marcadores do tipo SCAR (*Sequence Characterized Amplified Regions*) e juntamente com outros três SCAR provenientes de extremidade de uma biblioteca genômica de BAC (*Bacterial Artificial Chromosome*) e três marcadores do tipo SSR (*Simple Sequence Repeat*) formaram dois pequenos grupos de ligação (5,7 cM e 5,9 cM) contendo o gene de resistência (Mahé et al., 2008). Vale a pena ressaltar que a SAM é mais efetiva quanto mais próximo o marcador molecular estiver do loco de interesse e quando se utilizam marcadores flanqueando o gene (Singh et al., 2001; Balachiranjeevi et al., 2015). Dessa forma, esses 10 marcadores moleculares, por estarem proximamente ligados ao gene de resistência, podem ser utilizados para identificar o gene  $S_H3$ , por meio de SAM, em programas de melhoramento genético do cafeeiro.

Brito et al. (2010) e Diola et al. (2011) identificaram marcadores ligados aos genes  $S_H?$ , presente no acesso de Híbrido de Timor (HdT) UFV 427-15, que confere resistência a outras raças de *H. vastatrix*. Brito et al. (2010) iden-

tificaram três marcadores AFLP, por meio da técnica BSA (*Bulk Segregant Analysis*) (Michelmore et al., 1991) e, em continuação a esse trabalho, Diola et al. (2011) desenvolveram um mapa genético de alta densidade, com seis marcadores SCAR delimitando uma região cromossômica de 9,45 cM e flanqueando o gene  $S_H?$  em 0,7 e 0,9 cM. Esses marcadores também apresentam potencial para serem usados em SAM e, quando utilizados na etapa adequada do processo de melhoramento, permitem uma seleção precoce de indivíduos de interesse.

Quatro marcadores do total identificado por Mahé et al. (2008) e três obtidos no trabalho conduzido por Diola et al. (2011), ligados aos genes de efeito maior,  $S_H3$  e  $S_H?$ , foram validados por Alkimim et al. (2017). Esses marcadores foram utilizados para selecionar cafeeiros portadores dos genes  $S_H3$  e  $S_H?$ , em população resultante de cruzamentos entre seleções indianas com cultivares de *C. arabica*. Com base nos resultados desse trabalho, foi possível identificar cafeeiros arábica que são portadores de mais de um gene de resistência à ferrugem introgridido de outras espécies cafeeiras. Portanto, esses genótipos são importantes para o programa de melhoramento de *C. arabica* visando à resistência múltipla e durável (piramidação de genes).

Mais recentemente, outros dois marcadores moleculares associados a genes que conferem resistência a outras raças de *H. vastatrix* foram identificados e clonados. Barka et al. (2020) e Almeida et al. (2021), usando a técnica de biologia molecular, encontraram dois genes diferentes em HdT CIFC 832/1 e HdT CIFC 832/2. O gene identificado por Barka et al. (2020), aqui denominado de  $S_H10$ , corresponde a um gene de resistência análogo (RGA) e foi completamente sequenciado e caracterizado. Ao analisar esse gene nos clones diferenciadores de raças de *H. vastatrix*, observou-se que o loco do gene  $S_H$  clonado apresenta um polimorfismo característico que confere diferentes fenótipos de resistência à ferrugem do café (Barka et al., 2020). O gene identificado por Almeida et al. (2021) foi clonado com base em uma biblioteca de cromossomos artificiais bacterianos (BAC) do café proveniente do genótipo HdT CIFC 832/2 e selecionado usando um marcador funcional. Um análogo de gene de resistência a doenças foi clonado e caracterizado e mostrou um motivo típico RLK (*Receptor-Like Knase*) de planta. A análise da presença/ausência desse gene na série diferenciadora de raças sugeriu que o gene

candidato clonado não corresponde a nenhum dos nove genes  $S_H$  relatados anteriormente e também é diferente do  $S_H10$ , sendo, portanto, denominado de  $S_H11$ .

Além de incorporar genes de resistência à ferrugem, por meio da SAM também tem sido possível introduzir genes de resistência à CBD em cultivares de café. Essa doença é responsável por destruir cerca de 80% da lavoura de cultivares suscetíveis de *C. arabica*, caso nenhum controle seja aplicado (Gimase et al., 2020). CBD, que atualmente se encontra restrita ao continente africano, representa grande ameaça para as lavouras cafeeiras na América e Ásia, onde o patógeno ainda não foi detectado. Com as alterações climáticas que vêm ocorrendo ao longo dos anos, doenças e/ou pragas sem importância significativa podem passar a ser limitantes à cultura. Portanto, o estudo e a seleção de genótipos resistentes se fazem necessários, mesmo em locais onde o patógeno/praga ainda não foi detectado. Como essa doença não se encontra presente no Brasil, a seleção de materiais resistentes nesse caso de ausência do patógeno só é possível por meio da utilização da SAM (Alkimim et al., 2017). Essa estratégia permite um melhoramento preventivo, a fim de evitar prejuízos futuros.

A resistência à CBD em *C. arabica* é governada por três genes denominados T, R e k, presentes no Híbrido de Timor (gene T), Rume Sudan (genes R e k) e K7 (gene k). Os genes T e R são dominantes e o gene k recessivo (Vossen; Walyaro, 1980). Gichuru et al. (2008) identificaram o loco de resistência a *C. kahawae* e este foi designado *Ck-1*, provável sinônimo do gene T previamente caracterizado. Os autores obtiveram oito marcadores AFLP e dois SSR ligados ao gene *Ck-1*, que ficou localizado em um segmento de 11 cM. Gichimu et al. (2014) confirmaram a ligação de um dos marcadores SSR identificado, demonstrando o potencial do seu uso na SAM visando resistência à CBD. Alkimim et al. (2017), trabalhando com cafeeiros resultantes de cruzamentos entre Seleções Indianas (contendo genes introgrididos de *C. liberica*) com cultivares de *C. arabica*, validaram os dois marcadores SSR identificados por Gichuru et al. (2008). Com base nos resultados alcançados, foram identificados genótipos que contêm o gene de resistência à CBD, além de genes de resistência à ferrugem (Alkimim et al., 2017).

Embora o gene T ou *Ck-1* tenha sido mapeado com sucesso, é notória a necessidade de mapeamento genético dos outros genes que conferem resistência à CBD (genes R e k) para melhorar a eficiência da seleção (Gimase et al., 2019). Nesse sentido, Gimase et al. (2019) conduziram um estudo com objetivo de avaliar a adequação de uma população  $F_2$  oriunda do cruzamento entre os parentais, Rume Sudan e SL28 (cultivar altamente suscetível à CBD), para o mapeamento genético do gene R. Os resultados demonstraram diferenças significativas entre os genótipos para a resistência à CBD e foi observada herança monogênica dominante. Em continuação a esse trabalho, Gimase et al. (2020) conduziram um estudo com essa mesma população  $F_2$  (Rume Sudan x SL28) e, usando a técnica de genotipagem GBS (*Genotyping by Sequencing*) e estudo de associação genômica (GWAS-*Genome-Wide Association Study*), identificaram dois marcadores SNP (*Single Nucleotide Polymorphisms*) significativamente associados à resistência à CBD. Esses marcadores foram denominados *Ck-2* e *Ck-3*, complementar ao *Ck-1* previamente identificado (Gichuru et al., 2008), e foram recomendados para uso na SAM em programas de melhoramento de cafeeiros arábica para resistência à CBD e, possivelmente, em estudos de seleção genômica.

Para resistência ao nematoide *Meloidogyne exigua*, foram identificados 14 marcadores moleculares do tipo AFLP ligados ao gene dominante *Mex-1*, primeiro gene para resistência a nematoide identificado em café. Esses marcadores foram utilizados para construção de um mapa contendo o gene de resistência. Os resultados representam um importante ponto de partida para a seleção precoce de cafeeiros resistentes a *M. exigua* (Noir et al., 2003).

## Seleção genômica ampla

Como em outras áreas do conhecimento, os métodos seletivos utilizados nos programas de melhoramento sofreram profundas modificações e, nesse sentido, a SAM tem sido utilizada rotineiramente. No entanto, a SAM tem-se mostrado mais apropriada para características menos complexas, mono ou oligogênica e com alta herdabilidade. Por outro lado, a Seleção Genômica (GS) ou Seleção Genômica Ampla (GWS) tem evidenciado altas acurácias seletivas, inclusive para características com alto grau de complexidade. Essa estratégia de seleção foi proposta por Meuwissen et al. (2001) como uma

forma de aumentar a eficiência e acelerar o melhoramento genético. A GWS refere-se ao uso de informações moleculares para se estimar efeitos de marcadores genéticos com base em um modelo de predição de fenótipos ainda não expressos pelos indivíduos. Como a análise é genômica, a seleção precoce e acurada de genótipos mais promissores em uma população pode ser realizada (Begum et al., 2015). Com o número expressivo de marcadores moleculares distribuídos ao longo do genoma, há elevada probabilidade de se identificar um ou mais marcadores de DNA em desequilíbrio de ligação com os QTL (*Quantitative Trait Loci*) que explicam uma característica (Heffner et al., 2009), sem, contudo, exigir o prévio conhecimento do mapeamento genético da espécie. Isso permite explicar grande parte da variação genética do caráter quantitativo (Meuwissen et al., 2001), fundamentando o uso dessa técnica seletiva em características agrônômicas de alta complexidade.

O método envolve, basicamente, três etapas: (i) estimação; (ii) validação do modelo de predição dos efeitos dos locos marcados em indivíduos das populações de estimação e de validação; e (iii) predição do valor genético genômico (VGG) dos indivíduos em uma população de seleção e, a partir destes, a seleção daqueles indivíduos mais promissores para uma dada característica (Resende et al., 2008). O VGG representa o quanto da composição genética de um indivíduo contribui para o valor fenotípico da próxima geração.

Os efeitos dos locos marcados são estimados simultaneamente e sem nenhuma comparação de diferenças significativas entre eles (Heslot et al., 2012), pois se utiliza modelo linear misto, no qual os efeitos dos marcadores são considerados como aleatórios (Resende et al., 2008). Assim, após a estimação e validação de modelos preditivos adequados, a GWS apresenta alta acurácia seletiva, pois a seleção é baseada exclusivamente em marcadores de DNA (Goddard; Hayes 2007). Por não envolver comparação estatística dos efeitos dos locos genéticos, não está sujeita aos erros tipo II, associados à seleção de marcadores ligados a QTL.

Na população de validação, normalmente, utiliza-se conjunto de dados menor do que aquele que compõe a população de estimação. Os indivíduos também são genotipados e fenotipados para as características de interesse. Esta amostra independente é utilizada para testar e verificar as acurácias das equações de predição de VGGs. Para computar essa acurácia, os VGGs são

preditos, usando os efeitos dos locos estimados na população de estimação, e submetidos à análise de correlação com os valores fenotípicos corrigidos para cada característica agrônômica (Resende et al., 2008). Como a amostra de validação não foi envolvida na predição dos efeitos dos marcadores, os erros nas estimativas dos valores genéticos genômicos e dos valores fenotípicos são independentes e toda covariância entre esses valores é de natureza genética (Resende et al., 2008).

Embora as populações de estimação e validação devam ser genotipadas e fenotipadas, as demais populações, que serão submetidas à seleção, não requerem fenotipagem. Essa é uma das grandes vantagens desse método seletivo, pois, além de facilitar, possibilita a seleção precoce dos indivíduos desde os seus estádios iniciais (Resende et al., 2012). Na seleção genética, que é praticada, normalmente, pelo procedimento BLUP (*Best Linear Unbiased Prediction*), avalia-se o fenótipo com intuito de estimar o genótipo. Já na GWS, avalia-se o genótipo, por meio de milhares de marcadores de DNA, e, por meio destes, o VGG pode ser estimado. A GWS é mais acurada porque trabalha com matriz de parentesco estimada pelos marcadores de DNA, enquanto, no método BLUP, trabalha-se com a matriz de parentesco estimada pelo parentesco entre os indivíduos (Grattapaglia; Resende, 2011).

A GWS propicia uma forma de seleção precoce direta (SPD), pois captura, precocemente, os efeitos de genes que serão expressos na idade adulta. Ao contrário, a seleção precoce tradicional é indireta, pois atua (via avaliação fenotípica) sobre genes ativados na idade precoce, esperando que esses informem parcialmente sobre genes expressos na idade adulta. Assim, a SPD propiciada pela GWS é especialmente importante para o melhoramento de organismos perenes como o cafeeiro.

Em resumo, a superioridade da GWS sobre a seleção baseada em fenótipos pode ser atribuída a quatro fatores: (i) uso da matriz de parentesco real e própria de cada caráter, fato que aumenta a acurácia seletiva; (ii) viabilização da SPD, que aumenta o ganho genético por unidade de tempo; (iii) permissão da avaliação repetida de cada alelo (propicia repetição experimental) sem o uso de testes clonais e de progênies, fato que aumenta a acurácia seletiva; e (iv) uso de maior número de informações, combinando três tipos de informações (fenotípica, genotípica e genealógica) para corrigir e desregressar os dados e fazer a análise genômica, fato que aumenta a acurácia (Resende et al., 2008).

Após a proposição da GWS em 2001, o procedimento permaneceu discreto até 2007, quando a evolução e aperfeiçoamento das plataformas de sequenciamento viabilizaram a obtenção de grandes conjuntos de dados de DNA. A partir do desenvolvimento de genotipagem em larga escala, vários trabalhos abordaram diferentes metodologias estatísticas de GWS e suas acurácias no melhoramento animal e vegetal (Meuwissen et al., 2001; Bernardo; Yu 2007; Fernando et al., 2007; Goddard; Hayes 2007). Outros trabalhos relataram que a GWS é o novo paradigma em genética quantitativa (Resende et al., 2008; Campos et al., 2009) no melhoramento animal (Hayes et al., 2009), no melhoramento de plantas anuais (Heffner et al., 2009) e no de espécies perenes (Peixoto et al., 2017; Sousa et al., 2019; Alkimim et al., 2020).

Em plantas perenes, a aplicabilidade da GWS assume importância acentuada por possibilitar a seleção precoce dos indivíduos (Sousa et al., 2019). Embora possua elevada importância econômica e social, ainda são escassos os trabalhos com seleção genômica publicados para as espécies de cafeeiros. Nessa cultura, basicamente, a seleção dos indivíduos tem sido praticada por meio de análises biométricas que utilizam, principalmente, dados fenotípicos de produção e resistência às doenças. Com o aperfeiçoamento das plataformas de sequenciamento, de forma análoga ao que aconteceu com os cereais, os genomas dos cafés arábica e canéfora foram disponibilizados (Denoëud et al., 2014; Tran et al., 2018). Isso tem possibilitado a identificação de milhares de marcadores de DNA, sendo disponibilizados vários conjuntos de marcadores SNP (*Single Nucleotide Polymorphisms*) (Moncada et al., 2016; Sant'Ana et al., 2018; Tran et al., 2018; Merot-L'anthoene et al., 2019; Sousa et al., 2019; Alkimim et al., 2020).

De forma pioneira para o café arábica, Sousa et al. (2019) utilizaram a seleção genômica em cafeeiros arábica, pertencentes a diferentes gerações em melhoramento. As análises de seleção genômica para 18 características morfoagronômicas permitiram a redução do tempo de ciclo seletivo em 50%, o que possibilita desenvolver uma cultivar de café em 12 anos, maximizando os ganhos de seleção por unidade de tempo. Carvalho et al. (2020) estimaram o efeito da bienalidade na precisão dos modelos de seleção genômica ao se avaliar a característica produtividade. Os cafeeiros foram genotipados com marcadores moleculares SNP e tiveram

a característica produção dos indivíduos fenotipada, sendo realizada nos anos de baixa (2005 e 2007) e alta (2006 e 2008) produção. As análises consideraram os efeitos dos genótipos e, também, a interação entre genótipo x ano. O cenário de predição incluiu grupos de treinamento de um único ano e grupos bienais. Eles observaram que treinar os modelos genômicos em biênios de anos sucessivos e prever o biênio seguinte foi a estratégia mais eficiente. No geral, assim como Sousa et al. (2019), verificaram um ganho com a seleção genômica em relação à fenotípica, principalmente devido ao reduzido tempo por ciclo seletivo.

De forma semelhante, outros trabalhos, realizados com a outra espécie de café de interesse agrônomo, *C. canephora*, demonstraram a potencialidade da seleção genômica na redução do tempo necessário para o desenvolvimento de cultivares de café (Ferrão et al., 2019; Alkimim et al., 2020). Ferrão et al. (2019) avaliaram o desempenho de 13 metodologias estatísticas utilizadas na seleção genômica implementadas em dados de duas populações (precoce e intermediária) de seleção recorrente de *C. canephora*, avaliadas em dois locais (Marilândia e Sooretama, ES). A genotipagem se deu por marcadores moleculares SNP e a fenotipagem realizada em três características importantes no café, produção de grãos de café, incidência de ferrugem e produtividade de grãos limpos e secos. Os autores observaram resultados similares entre as diferentes metodologias de seleção genômica avaliada, havendo um destaque para o método RR-BLUP (*Ridge Regression* – BLUP) por apresentar boa capacidade preditiva e ser relativamente mais simples, o que reduz o custo computacional. No método RR-BLUP é assumido que todas as marcas possuem efeitos aleatórios com variância comum (Meuwissen, et al., 2001). Assumir variância comum entre os efeitos é diferente de assumir efeitos iguais para os locos marcados (Bernardo; Yu, 2007), mas significa dizer que todas as marcas possuem o mesmo *shrinkage*. Alkimim et al. (2020) utilizaram a seleção genômica em populações de *C. canephora*, incluindo os dois grupos varietais, Conillon e Robusta, e híbridos entre eles. Os indivíduos foram genotipados com marcadores moleculares SNP e tiveram oito características morfoagronômicas fenotipadas. Os resultados evidenciaram a importância da seleção genômica como uma ferramenta promissora para o melhoramento de *C. canephora*.

Conforme demonstrado, os resultados a partir das metodologias de seleção genômica nas duas principais espécies de café, arábica e canéfora, evidenciaram que a utilização dessa metodologia seletiva é altamente promissora para os cafeeiros. Resultados similares foram observados para outras espécies de vegetais e animais. Diferentes metodologias genético-estatísticas têm sido propostas. Isso é de extrema importância, pois os modelos preditivos são baseados em suposições biológicas e estatísticas, então se espera que seu desempenho varie dependendo de quão bem essas suposições se alinham com a verdadeira arquitetura genética do fenótipo das populações analisadas. Com a viabilidade da genotipagem em larga escala para a maioria das espécies, as metodologias de seleção genômica são alternativas reais para remodelar os esquemas tradicionais de seleção nos programas de melhoramento genético, resultando na possibilidade de ganho genético por unidade de tempo, reduzindo a duração do ciclo de seleção e, conseqüentemente, o desenvolvimento de cultivares de café.

## Métodos de seleção genômica

Um método ideal para seleção genômica ampla (GWS) deve contemplar três atributos: (i) **acomodar a arquitetura genética** do caráter em termos de genes de pequenos e grandes efeitos e suas distribuições; (ii) realizar a **regularização** do processo de estimação em presença de multicolinearidade e grande número de marcadores, usando para isso estimadores do tipo *shrinkage*; (iii) realizar a **seleção de covariáveis** (marcadores) que afetam a característica em análise. O problema principal da GWS é a estimação de um grande número de efeitos a partir de um limitado número de observações e também as colinearidades advindas do desequilíbrio de ligação entre os marcadores. Os estimadores do tipo *shrinkage* lidam adequadamente com isso, tratando os efeitos de marcadores como variáveis aleatórias e estimando-os simultaneamente (Resende, 2007). Os principais métodos para a GWS são baseados em regressão aleatória e podem ser divididos em três grandes classes: regressão explícita, implícita e com redução dimensional. Na primeira classe, destacam-se os métodos RR-BLUP, Lasso (*Least Absolute Shrinkage and Selection Operator*), BayesA e BayesB, dentre outros. Na classe de regressão implícita, citam-se os métodos de redes neurais e RKHS (*Reproducing Kernel Hilbert Spaces*), que é um

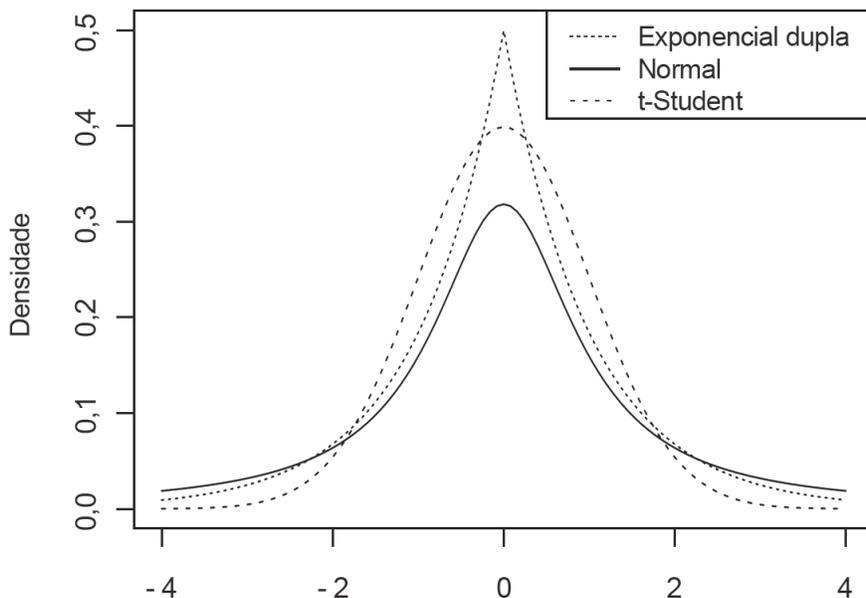
método semiparamétrico. Dentre os métodos de regressão com redução dimensional, destacam-se os de componentes independentes, quadrados mínimos parciais e de componentes principais.

Os métodos de regressão explícita são divididos em dois grupos: (i) métodos de estimação penalizada (RR-BLUP, Lasso); e (ii) métodos de estimação bayesiana (BayesA, BayesB, Fast BayesB, BayesC $\pi$ , BayesD $\pi$ , Regressão Bayesiana, BayesR, BayesRS, Blasso, llasso e outros). Cada método sem seleção de covariáveis tem o seu similar com seleção de covariáveis. Assim, têm-se os pares sem e com seleção: BayesA – BayesB; BRR – BayesC $\pi$ ; Blasso – llasso.

Se os efeitos de marcadores são tomados como fixos, não é possível considerar as covariâncias entre os efeitos de marcadores. Com alta densidade de marcadores, mais de um marcador estará em desequilíbrio de ligação com um QTL segregante, o que resultará em covariância entre efeitos de marcadores. A maioria dos marcadores não terá efeito algum sobre um caráter. Assim, os efeitos estimados desses marcadores vazios serão falsos. Este problema é maior no caso em que os marcadores são considerados de efeitos fixos, pois esses pseudoefeitos não serão regressados em direção a zero.

As distribuições assumidas para os efeitos genéticos de marcadores nos diferentes métodos de GWS são: RR-BLUP - Normal de Gauss com variância comum; Métodos Bayesianos - t de Student dado priori qui-quadrado para as variâncias; Lasso - Exponencial Dupla de Laplace. A Figura 1 ilustra as formas das distribuições normal (RR-BLUP), t (BayesA) e exponencial (Lasso).

Observa-se que, em relação ao RR-BLUP, a densidade a priori utilizada no Lasso Bayesiano apresenta maior massa de densidade no valor zero e caudas mais robustas, exercendo maior encurtamento sobre coeficientes de regressão próximos de zero e menor encurtamento sobre coeficientes de regressão distantes de zero. A densidade a priori utilizada no BayesA também apresenta maior massa de densidade no valor zero e caudas mais robustas do que a normal usada no RR-BLUP. O Lasso Bayesiano também exerce maior encurtamento sobre coeficientes de regressão próximos de zero do que o BayesA. No entanto, as caudas das distribuições são similares pelos dois métodos (Figura 1).



**Figura 1.** Funções densidade de probabilidade das distribuições exponencial dupla, normal e t de Student, todas com médias iguais a zero e variâncias iguais à unidade.

O método BayesA implica grande número de marcadores com efeitos pequenos e poucos marcadores com efeitos moderados a grandes. O Blasso implica grande número de marcadores com efeitos próximos a zero e poucos marcadores com efeitos moderados a grandes. O RR-BLUP implica grande número de marcadores com efeitos pequenos.

Um refinamento da seleção genômica pode ser conseguido pelo uso de QTN (*Quantitative Trait Nucleotides*) em lugar de marcadores SNP. A evolução da tecnologia genômica é previsível e a mutação causal de uma variação genética em nível de nucleotídeo (QTN) poderá ser acessada em um futuro próximo. Assim, a seleção genômica poderá ser aperfeiçoada pelo uso direto dos QTN em lugar dos SNP. O uso dos QTN trará as seguintes vantagens: (i) a GWS não dependerá do desequilíbrio de ligação, pois o QTN será acessado diretamente e não via marcadores, o que aumentará a durabilidade da predição genômica e será útil também a longo prazo; (ii) a predição genômica poderá ter validade (transferibilidade) através de diferentes populações e espécies em um mesmo gênero; (iii) a predição genômica usará QTNs específicos para cada caráter, ao

contrário do G-BLUP via SNP, o qual usa a mesma matriz de parentesco G para todas as características; (iv) os índices de seleção multicaracterísticos ponderarão diretamente os QTNs e não as características fenotípicas; (v) a seleção genômica poderá usar um menor número de gerações (apenas as últimas) para a composição da matriz G, isto trará maior ganho genético e menor massa de dados a serem processados; (vi) as frequências alélicas dos QTNs serão acessadas diretamente e não via desequilíbrio de ligação com marcadores SNP.

## Capacidade preditiva, viés da predição e acurácia da GWS

A capacidade preditiva e o viés da predição são medidas práticas da capacidade de o método realizar predições de forma acurada e não viesada. A capacidade preditiva ( $r_{gy}$ ) é obtida por meio da correlação entre os valores genômicos preditos (g) e os valores fenotípicos observados (y), sendo equivalente à capacidade da GWS em estimar os fenótipos. O viés da predição (b) é obtido por meio do coeficiente de regressão dos valores genômicos preditos sobre os valores fenotípicos (Resende et al., 2012). A acurácia é obtida pelo estimador:  $r_{gg} = r_{gy}/\sqrt{h^2}$ , em que  $r_{gy}$  equivale à capacidade preditiva da GWS e  $h^2$  à herdabilidade individual (Resende et al., 2008, 2012).

A herdabilidade molecular ( $h^2m$ ) deve ser menor ou igual à herdabilidade do caráter ( $h^2$ ), pois  $h^2m$  é uma fração de  $h^2$  que é capturada. Além disso, o limite máximo do quadrado da acurácia da GWS é a  $h^2m$  (Campos et al., 2013).

Os coeficientes de correlação e regressão envolvendo valores observados e preditos são medidas práticas da capacidade de os métodos predizerem de forma acurada e não viesada, respectivamente. A correlação fornece a capacidade preditiva. O coeficiente de regressão equivale algebricamente a 1. Coeficientes de regressão abaixo de 1 indicam que os valores genéticos são superestimados e apresentam variabilidade além da esperada, e os acima de 1 indicam que os valores genéticos estimados apresentam variabilidade aquém da esperada. Não vício é importante quando a seleção envolve indivíduos de muitas gerações usando efeitos dos marcadores estimados em uma só geração. Coeficientes de regressão próximos de 1 indicam que as avaliações são não viesadas e são efetivas em predizer as reais magnitudes das diferenças entre os indivíduos em avaliação.

O coeficiente de regressão tem valor esperado igual a 1 e, nessa situação, indica que a predição foi não viesada. Assim sendo, pode-se também usar o coeficiente de regressão para estimar a herdabilidade molecular ( $h^2_M$ ). Vários valores de herdabilidade são avaliados, e aquele que fornece uma regressão igual a 1 deve ser escolhido como melhor estimativa. Se a regressão der resultado menor que 1, o valor de herdabilidade avaliado foi de alta magnitude e deve ser diminuído até a convergência para 1. Se a regressão der resultado maior que 1, o valor de herdabilidade avaliado foi de pequena magnitude e deve ser aumentado até a convergência para 1. Esse é o método R (Reverter et al., 2000; Resende, 2002) de estimação de componentes de variância e está implementado no *software* Selegen Genômica.

## Aumento da eficiência seletiva do melhoramento de plantas

O aumento da eficiência seletiva com o uso da GWS pode ocorrer pela alteração dos quatro componentes da expressão do progresso genético, dada por  $G_s = (k r_{gg} \sigma_g) / L$ , em que  $k$  é o diferencial de seleção padronizado (dependente da intensidade de seleção),  $r_{gg}$  é a acurácia seletiva,  $\sigma_g$  é o desvio-padrão genético (variabilidade genética) do caráter na população e  $L$  é o tempo necessário para completar um ciclo seletivo.

Nas espécies perenes, como o cafeeiro, o benefício da GWS se dá devido ao aumento de  $r_{gg}$  e redução em  $L$ . O aumento em  $r_{gg}$  é resultante do uso da matriz de parentesco real (Resende, 2007) e esse aumento depende do tamanho da população de estimação e da densidade de marcadores. O fator  $L$  é enormemente reduzido com a GWS, pois a predição genômica e a seleção podem ser feitas no estágio de plântulas. Assim, mesmo que  $r_{gg}$  seja menor ou de mesma magnitude que aquela obtida com a seleção fenotípica, a GWS será ainda superior à seleção baseada em fenótipos, devido à redução em  $L$ . A GWS explorando essas vantagens foi implementada por Resende et al. (2008), Grattapaglia e Resende (2011), Resende et al. (2012) e Resende Junior et al. (2012a, 2012b) em espécies florestais; por Cavalcanti et al. (2012) em cajueiros; por Oliveira et al. (2012) em mandioca; por Ferrão et al. (2017, 2019) e Alkimin et al. (2020) em café canéfora e por Sousa et al. (2019) e Carvalho et al. (2020) em café arábica.

## GWS em *Coffea arabica*

Até o momento, poucos trabalhos com GWS foram realizados em cafeeiro, podendo ser citado três em *C. canephora* (Ferrão et al., 2017, 2019; Alkimin et al., 2020) e dois em *C. arabica* (Sousa et al., 2019; Carvalho et al., 2020). Neste tópico, são apresentados os principais resultados obtidos no programa de Seleção Genômica da Embrapa Café/Epamig/ Universidade Federal de Viçosa (UFV).

O objetivo dos trabalhos realizados no programa foi avaliar a acurácia seletiva e a eficiência da GWS utilizada para prever o valor genético genômico (GEBV) dos indivíduos que compõem a população de cafeeiro arábica obtida por cruzamentos entre genitores do grupo de cultivares Catuaí e Híbrido de Timor. Um total de 195 indivíduos de café arábica, pertencentes a 13 famílias em geração  $F_2$ , retrocruzamentos suscetíveis e retrocruzamentos resistentes à ferrugem do cafeeiro foram fenotipados para 18 características agrônômicas e genotipados com 21.211 marcadores moleculares SNP.

## Material genético

A partir de cruzamentos entre três genitores do grupo de cultivares Catuaí e três genitores Híbrido de Timor (HdT), contrastantes em relação à ferrugem do cafeeiro, 13 progênies foram obtidas no programa de melhoramento genético de *C. arabica* da Epamig/UFV/Embrapa. Estas progênies são gerações de retrocruzamentos resistentes (RCr), retrocruzamentos suscetíveis (RCs) e  $F_2$ . Em cada progênie, foram selecionados 15 genótipos, totalizando 195 indivíduos.

## Avaliações fenotípicas

As avaliações fenotípicas de 18 características agrônômicas foram realizadas (11 contínuas e sete categóricas) nos 195 genótipos de *C. arabica*. As características contínuas foram avaliadas a partir de mensurações, conforme descrito na Tabela 1. Já as características categóricas, conforme detalhado a seguir, foram avaliadas por notas. O tamanho dos frutos maduros foi avaliado com notas 1, 2 e 3 para frutos pequenos, médios e grandes, respectivamente. A uniformidade de maturação dos frutos foi avaliada pela atribuição das notas 1, 2, 3 e 4 para maturação uniforme, medianamente

uniforme, medianamente desuniforme e desuniforme, respectivamente. O ciclo de maturação dos frutos foi avaliado com notas de 1 a 5, em que as notas 1, 2, 3, 4 e 5 se referiram às maturações precoce, precoce a média, média, média para tardia e tardia, respectivamente. As incidências de ferrugem do cafeeiro, cercosporiose e bicho-mineiro foram avaliadas com notas de 1 a 5, atribuídas para genótipos sem nenhum sintoma das enfermidades e para genótipos altamente suscetíveis às mesmas, respectivamente. A característica vigor vegetativo foi avaliada por notas de 1 a 10, sendo a nota 1 atribuída a plantas totalmente depauperadas e a nota 10 para plantas consideradas de vigor vegetativo máximo.

**Tabela 1.** Características fenotípicas avaliadas no estudo de seleção genômica ampla em *Coffea arabica*.

Característica	Forma de avaliação
Produção	Litros de café recém-colhido por planta
Comprimento de uma folha (cm)	Medidos em folha do 3º ou 4º par de um ramo plagiotrópico do terço médio da planta
Largura de uma folha (cm)	
Comprimento de um ramo (cm)	Medidos no ramo plagiotrópico localizado no terço médio da planta
Número de nós reprodutivos	
Número de nós vegetativos	
Número total de frutos	
Volume de frutos	
Altura de plantas (cm)	Medida do ramo ortotrópico (superfície do solo até o ponto final de crescimento do ramo)
Diâmetro da copa (cm)	Medido no sentido transversal à linha de plantio, medindo-se a maior projeção da “saia” do cafeeiro
Diâmetro do caule (cm)	Medido na região do coleto da planta (cerca de 5 cm da superfície do solo)
Tamanho dos frutos maduros	Avaliado por notas de 1 a 3
Uniformidade de maturação dos frutos	Avaliada por notas de 1 a 4
Ciclo de maturação dos frutos	Avaliados por notas de 1 a 5
Incidência de ferrugem	
Incidência de cercosporiose	
Infestação de bicho-mineiro	
Vigor vegetativo	Avaliado por notas de 1 a 10

## Análise de qualidade de marcadores moleculares

O controle de qualidade dos marcadores moleculares baseou-se em uma CR (*Call Rate*) superior a 90% e MAF (*Minor Allele Frequency*) de 5%. O ponto de corte para inclusão de uma marca na análise foi dado por  $MAF = 1 / (2N)^{1/2}$ ; isto advém do desvio-padrão de uma proporção, dado por  $(pq)^{1/2} / (2N)^{1/2}$ , em que N é o número de indivíduos genotipados, significando que quanto menor N, maior deve ser a MAF adotada para que o efeito da marca seja estimado com precisão (Resende, 2015). Além disso, para evitar a ocorrência de SNP falsos (Vidal et al., 2010), os marcadores moleculares que não apresentaram variância entre os indivíduos que compõem a população em estudo foram eliminados. Os 195 genótipos foram genotipados para 21.211 e, após análises de qualidade, 11.187 SNP foram selecionados e validados em análises de diversidade e estrutura genética de populações.

Os dados fenotípicos foram analisados por meio dos modelos mistos (REML/BLUP), implementados no *software* Selegen-REML/BLUP (Resende, 2016). As análises de GWS foram realizadas utilizando o método G-BLUP via procedimento RKHS (*Reproducing Kernel Hilbert Spaces*) com algoritmo Bayesiano, por meio do pacote BGLR (Pérez; Campos, 2014), implementado no *software* R.

### Eficiência da GWS

A eficiência seletiva da GWS, em comparação com a seleção baseada apenas em fenótipos, foi calculada usando a expressão:  $Ef = (r_{gg}L_f)/(r_{yy}L_{GWS})$ , em que  $r_{gg}$  é a acurácia seletiva da GWS;  $L_f$  é o tempo médio requerido para o ciclo de seleção com base em fenótipos;  $r_{yy}$  é a acurácia da seleção fenotípica;  $L_{GWS}$  é o tempo médio requerido para o ciclo de seleção com base na GWS (Resende et al., 2012). As análises de eficiências foram estimadas considerando a necessidade de seis anos para obtenção das acurácias fenotípicas, conforme tempo de obtenção dos dados fenotípicos. Para a obtenção das acurácias da GWS, foram considerados os tempos de um a seis anos, conforme possibilidade da técnica.

### Resultados e discussão

Os valores estimados de herdabilidade genômica  $h^2$  variaram de 0,16 para a característica diâmetro do caule (DCa) a 0,46 para as características número de nós vegetativos (NNV) e altura de plantas (API) (Tabela 2). As estimativas

de h2a para todas as características avaliadas apresentaram erro padrão igual ou inferior a 0,05. Para as características diâmetro do caule (DCa), número de nós reprodutivos (NNR), uniformidade de maturação dos frutos (UMat), comprimento de folha (CF) e largura de folha (LF) foram observadas estimativas de h2a de magnitude moderada. No entanto, para tais características, as estimativas de rgy foram mais baixas. Para características que apresentam baixas estimativas de herdabilidades, são esperadas menores estimativas de capacidade preditivas (Legarra et al., 2008).

**Tabela 2.** Estimativas de parâmetros genéticos obtidos por meio de análises de modelos mistos (REML/BLUP), resultados da análise da Seleção Genômica Ampla (GWS) para dezoito características morfoagronômicas em população de melhoramento de *Coffea arabica*.<sup>(1)</sup>

Característica	REML/BLUP			GWS					
	h2fen	ryy	h2a	sdh	rgy	sdr	b	sdb	rgg
Prod	0,55	0,74	0,26	0,03	0,13	0,27	1,63	3,01	0,25
Fer	0,61	0,78	0,31	0,04	0,26	0,22	1,5	1,34	0,46
Cer	0,38	0,62	0,44	0,05	0,31	0,3	1,45	1,55	0,47
BM	0,3	0,55	0,3	0,04	0,18	0,24	1,34	1,71	0,33
Vig	0,7	0,84	0,34	0,04	0,21	0,19	1,38	1,39	0,36
CMat	0,72	0,85	0,31	0,05	0,12	0,19	1,31	2,15	0,21
UMat	0,3	0,55	0,28	0,03	0,03	0,27	0,62	3,13	0,06
TFr	0,5	0,71	0,36	0,04	0,23	0,24	1,52	2,01	0,39
CF	0,42	0,65	0,29	0,02	0,06	0,23	0,87	2,21	0,12
LF	0,44	0,66	0,32	0,05	0,03	0,25	0,44	2,62	0,06
CRP	0,78	0,88	0,41	0,04	0,32	0,34	1,2	1,3	0,5
NNR	0,49	0,7	0,23	0,02	-0,01	0,21	0,25	3,25	-
NNV	0,44	0,66	0,46	0,04	0,38	0,21	1,92	1,63	0,56
NFRP	0,49	0,7	0,34	0,05	0,14	0,2	1,33	2,1	0,23
VFRP	0,57	0,76	0,25	0,03	0,11	0,17	1,23	1,9	0,21
API	0,9	0,95	0,46	0,04	0,38	0,18	1,18	0,77	0,56
DCo	0,9	0,95	0,45	0,03	0,4	0,22	1,46	0,9	0,61
DCa	0,01	0,1	0,16	0,01	0,06	0,26	1,14	2,9	0,14

<sup>(1)</sup>h2fen = herdabilidade fenotípica; ryy = acurácia da seleção obtida pelo método REML/BLUP utilizando dados fenótipos; h2a = herdabilidade genômica; sdh = erro padrão das estimativas de h2a; rgy = capacidade preditiva da GWS; sdr = erro padrão das estimativas de rgy; b = viés da predição; sdb = erro padrão das estimativas de b; rgg = acurácia seletiva da GWS; Prod = produção de frutos de café; Fer = incidência de ferrugem; Cer = incidência de Cercospora; BM = infestação de bicho-mineiro; Vig = vigor vegetativo; CMat = ciclo de maturação dos frutos; UMat = uniformidade de maturação dos frutos; TFr = tamanho dos frutos maduros; CF = comprimento de uma folha; LF = largura de uma folha; CRP = comprimento de um ramo plagiotrópico; NNR = número de nós reprodutivos; NNV = número de nós vegetativos; NFRP = número de frutos por ramo plagiotrópico; VFRP = volume de frutos por ramo plagiotrópico; API = altura de planta; DCo = diâmetro da copa; DCa = diâmetro do caule.

O coeficiente de correlação ou capacidade preditiva ( $rgy$ ) e o coeficiente de regressão ou viés da predição ( $b$ ), envolvendo os valores fenotípicos observados e os valores genéticos preditos, são medidas práticas da capacidade dos métodos fazerem predições acuradas e não viesadas, respectivamente (Resende et al., 2014). As estimativas da capacidade preditiva ( $rgy$ ) das 18 características em análises variam de -0,01 a 0,40 para as características número de nós reprodutivos (NNR) e diâmetro da copa (DCo), respectivamente (Tabela 2). Tais estimativas apresentaram erro padrão variando de 0,17 a 0,34. Além do DCo, as maiores estimativas da capacidade preditiva foram observadas nas características número de nós vegetativos (NNV) e altura de plantas (API). Os resultados de  $h^2a$  e  $rgy$  demonstram alta associação positiva, apresentando coeficiente de correlação de 88%. Assim como observado nesse trabalho, associação entre a capacidade preditiva e a herdabilidade tem sido relatada (Cavalcanti et al., 2012; Gois et al., 2016). As estimativas do viés ( $b$ ) da predição variaram de 0,25 a 1,92 para as características número de nós reprodutivos (NNR) e número de nós vegetativos (NNV). A maioria das características avaliadas apresentou estimativa de  $b$  próxima à unidade indicando predição não viesada para as características. Portanto, são efetivas em prever as reais magnitudes das diferenças entre os indivíduos (Resende et al., 2012). Tais estimativas apresentaram erro padrão variando de 0,77 a 3,25.

A acurácia seletiva ( $rgg$ ) refere-se à correlação entre o valor genotípico verdadeiro do tratamento genético e aquele estimado ou predito a partir das informações fenotípicas (Gois et al., 2016). Valores adequados de  $rgg$  são aqueles próximos à unidade. A acurácia é tanto mais alta quanto menores forem os desvios absolutos entre os valores genéticos paramétricos e os valores genéticos estimados ou preditos (Resende; Duarte, 2007). O valor desta medida indica o quão preciso é o modelo em estimar o GEBV. As estimativas da acurácia seletiva obtidas com a utilização da GWS ( $rgg$ ) estão apresentadas na Tabela 2. Para a característica número de nós reprodutivos (NNR), a  $rgg$  não foi estimada, uma vez que, para essa característica, a capacidade preditiva foi próxima de zero. Para as demais características, os valores estimados de  $rgg$  variaram de 0,06 para uniformidade de maturação dos frutos (UMat) e largura de folha (LF) a 0,61 para o diâmetro da copa (DCo), sendo consideradas de magnitudes baixa à moderada (Resende; Duarte, 2007). Foi verificada alta correlação

entre as estimativas de rgg com h<sup>2</sup>a (82%) e com rgy (99%). Correlação positiva entre acurácia seletiva e herdabilidade também foi relatada para ferrugem amarela e ferrugem do colmo, em trigo (Ornella et al., 2012).

As baixas magnitudes de rgg observadas em algumas características avaliadas podem ser explicadas pelo reduzido tamanho populacional e, sobretudo, pelo tamanho efetivo da população. Em estudos com populações de trigo, o aumento do tamanho populacional possibilitou incrementos nas estimativas de acurácias seletivas (Heffner et al., 2011a, 2011b). O sucesso da seleção genômica é influenciado por diversos fatores, os quais interferem na acurácia seletiva de um modelo de GWS. Grattapaglia e Resende (2011) e Desta e Ortiz (2014) destacaram o tamanho da população de treinamento, tamanho efetivo populacional, densidade de marcadores, herdabilidade da característica e número de QTL que governam as características. Destes, a herdabilidade e o número de QTL envolvidos no caráter são características inerentes à arquitetura genética da característica (Resende et al., 2014). Além disso, a estrutura genética da população pode influenciar as predições genômicas (Zhang et al., 2010). Nesse sentido, as diferentes frequências alélicas entre subpopulações podem produzir associações falsas entre os dados moleculares e fenotípicos (Price et al., 2010) e, dessa forma, superestimar as herdabilidades e reduzir a acurácia seletiva (Riedelsheimer et al., 2012; Wray et al., 2013).

A eficiência das análises de GWS em relação à seleção fenotípica está apresentada na Tabela 3. Essa análise não foi realizada para a característica número de nós reprodutivos (NNR) porque a estimativa da capacidade preditiva dessa característica foi próxima de zero (Tabela 2). Os resultados demonstraram que a possibilidade de redução do tempo do ciclo seletivo permite aos melhoristas obtenção de ganhos genéticos precocemente, sendo esta uma das grandes vantagens da GWS. Ao se reduzir o ciclo seletivo de 24 para 12 anos (50%), observa-se que a GWS foi mais eficiente que a seleção com base nos dados fenotípicos para nove características avaliadas, as quais incluem incidências de ferrugem (Fer), cercosporiose (Cer) e bicho-mineiro (BM). Assim, ao se selecionar precocemente os indivíduos superiores por meio das análises de GWS, o melhorista concentrará esforços em genótipos potenciais, sendo os indesejáveis eliminados. Por essa razão, os custos de manutenção de populações de melhoramento em campo podem ser reduzidos. Além disso, a

seleção genômica precoce possibilita condução de populações de melhoramento com maior desempenho agrônomo, o que resultará na maximização dos ganhos genéticos. O fato de a seleção com base nos dados fenotípicos ser mais eficiente que a seleção genômica para algumas características pode ser explicado pelo número de genótipos avaliados. Outros genótipos de cafeeiro podem ser analisados com os SNPs selecionados neste estudo, de forma a aumentar as acurácias seletivas.

**Tabela 3.** Acurácia seletiva estimada a partir de diferentes densidades de marcadores SNP e eficiência da seleção genômica ampla (GWS) em relação à seleção fenotípica em população de melhoramento de *Coffea arabica*.

Característica <sup>(1)</sup>	Acurácia de acordo com o número de SNP <sup>(2)</sup>							Eficiência da GWS em função dos anos <sup>(3)</sup>	
	1000	4000	8000	12000	16000	20000	20477	24	12
Prod	-0,08	0,1	0,21	0,25	0,27	0,26	0,25	0,34	0,68
Fer	0,19	0,42	0,4	0,24	0,24	0,38	0,46	0,59	1,18
Cer	0,39	0,37	0,35	0,36	0,41	0,44	0,47	0,76	1,52
BM	0,23	0,28	0,23	0,23	0,27	0,26	0,33	0,6	1,2
Vig	0,27	0,46	0,43	0,37	0,49	0,37	0,36	0,43	0,86
CMat	-0,07	0,11	0,01	0,27	0,1	0,18	0,21	0,25	0,49
UMat	0,07	0,15	0,17	0,09	0,01	0,25	0,06	0,11	0,22
TFr	0,33	0,47	0,3	0,4	0,31	0,36	0,39	0,54	1,09
CF	0,14	0,18	0,22	0,07	0,12	0,16	0,12	0,18	0,37
LF	0	0,26	0,11	0,27	0,22	0,19	0,06	0,09	0,18
CRP	0,42	0,44	0,5	0,51	0,51	0,56	0,5	0,56	1,13
NNR	0,03	0,04	-0,07	0,12	0,09	0,06	-0,01	-	-
NNV	0,44	0,33	0,43	0,44	0,46	0,5	0,56	0,84	1,68
NFRP	0,08	0,18	0,29	0,33	0,27	0,25	0,23	0,33	0,67
VFRP	0,23	0,27	0,3	0,24	0,21	0,19	0,21	0,28	0,55
API	0,56	0,48	0,58	0,57	0,48	0,54	0,56	0,59	1,17
DCo	0,37	0,48	0,52	0,58	0,59	0,57	0,61	0,64	1,28
DCa	-0,03	0,14	0,16	-0,04	-0,02	-0,05	0,14	1,38	2,77

<sup>(1)</sup>Prod = produção de frutos de café; Fer = incidência de ferrugem; Cer = incidência de Cercospora; BM = infestação de bicho-mineiro; Vig = vigor vegetativo; CMat = ciclo de maturação dos frutos; UMat = uniformidade de maturação dos frutos; TFr = tamanho dos frutos maduros; CF = comprimento de uma folha; LF = largura de uma folha; CRP = comprimento de um ramo plagiotrópico; NNR = número de nós reprodutivos; NNV = número de nós vegetativos; NFRP = número de frutos por ramo plagiotrópico; VFRP = volume de frutos por ramo plagiotrópico; API = altura de planta; DCo = diâmetro da copa; DCa = diâmetro do caule. <sup>(2)</sup>Acurácia seletiva estimada a partir de diferentes densidades de marcadores SNP. <sup>(3)</sup>Eficiência da seleção genômica ampla (GWS) em relação à seleção fenotípica.

## GWS em *Coffea canephora*

### Material genético

A população de estudo foi composta por clones dos grupos varietais Conilon e Robusta e por híbridos interpopulacionais originados de cruzamentos entre estes grupos. O material genético de Conilon foi obtido do Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural (Incaper) e o material de Robusta foi obtido do Centro Agronómico Tropical de Investigación y Enseñanza (Catie). Essas populações compõem o programa de melhoramento da Empresa de Pesquisa Agropecuária de Minas Gerais (Epamig) em parceria com a Universidade Federal de Viçosa (UFV) e a Empresa Brasileira de Pesquisa Agropecuária - Café (Embrapa Café), localizado no município de Oratórios, MG.

Os grupos varietais Conilon e Robusta foram constituídos por 51 e 32 genótipos, respectivamente. Além desses genótipos, 82 híbridos interpopulacionais foram obtidos por meio de cruzamentos artificiais entre cinco genótipos do grupo Conilon (genitores masculinos) e cinco do grupo Robusta (genitores femininos), avaliados em dialelo parcial interpopulacional.

A população em estudo, inicialmente, foi composta por 72 genótipos nos quais os marcadores SNP foram identificados e validados. A seleção fenotípica, por meio do procedimento REML/BLUP, foi realizada em 192 genótipos de cafeeiros. Por fim, a população avaliada por meio da GWS foi composta por 165 genótipos.

Foram avaliados 165 genótipos, sendo 51 genótipos Conilon, 32 Robusta e 82 híbridos interpopulacionais. Utilizando a plataforma de sequenciamento da empresa RAPID Genomics, foram identificados 18.111 marcadores SNP, restando após análises de qualidade 14.429 marcadores.

### Avaliações fenotípicas

As avaliações fenotípicas foram realizadas para oito características durante três anos consecutivos (2014 a 2016). Foram avaliadas cinco características categóricas e três características contínuas. Estas foram realizadas na época de maturidade fisiológica dos frutos dos cafeeiros.

As características categóricas avaliadas foram: vigor vegetativo (Vig), avaliação das incidências, em campo, de ferrugem (Fer) e de cercosporiose (Cer), época de maturação dos frutos (Mat) e tamanho de fruto (TF). O vigor vegetativo foi avaliado pelo aspecto geral da planta, observando-se o enfolhamento, a coloração das folhas, o estado nutricional e a sanidade dos cafeeiros. A escala de notas de 1 a 10 foi utilizada, sendo a nota 1 atribuída às plantas totalmente depauperadas e a 10 às plantas altamente vigorosas; a incidência à ferrugem do cafeeiro, causada pelo fungo *Hemileia vastatrix* Berk. & Br., foi avaliada com notas de 1 a 5. A nota 1 foi atribuída para genótipos sem nenhum sintoma do patógeno e a nota 5 para genótipos altamente suscetíveis; a incidência à cercosporiose, causada pelo fungo *Cercospora coffeicola*, foi avaliada por notas de 1 a 5. A nota 1 foi atribuída para genótipos sem nenhum sintoma do patógeno e a nota 5 para genótipos altamente suscetíveis; a época de maturação dos frutos do cafeeiro foi classificada em precoce, média e tardia, recebendo notas de 1 a 3, respectivamente; e o tamanho dos frutos foi classificado como pequeno, médio e graúdo, notas 1, 2 e 3, respectivamente.

As características contínuas avaliadas foram: altura da planta (AP), diâmetro da projeção da copa (DC) e produção em litros por planta (Prod). A altura da planta, em centímetros (cm), foi determinada pela medida da ramificação ortotrópica mais desenvolvida, do nível do solo até o último ponto apical do cafeeiro com o auxílio de uma trena métrica afixada em uma haste de madeira; o diâmetro da projeção da copa da planta foi determinado em centímetros (cm) por meio de régua no sentido perpendicular à linha de plantio; e a produção por planta de cafeeiros foi avaliada colhendo todos os frutos presentes em um genótipo e mensurado o volume total em litros de café recém-colhido.

## Resultados e discussão

Com o sequenciamento dos 165 genótipos utilizando 10.000 sondas, distribuídas ao longo de todo o genoma, foram identificados 18.111 SNP. Após as análises de qualidade, foram obtidos 14.429 marcadores SNP. As avaliações de qualidade dos SNP permitem a identificação de marcadores com os critérios de qualidade estabelecidos como ideais. As análises de qualidade são vantajosas por removerem os marcadores de baixa qualidade antes das análises estatísticas e, dessa forma, reduzirem a ocorrência de falso-positivos (erro tipo I) e falso-negativos (erro tipo II) (Anderson et al., 2010). Alta densidade de marcadores é

importante para capturar genes de efeito menor e maior e, conseqüentemente, aumentar a probabilidade de explicar a maior parte da variação genética do caráter em estudo (Resende et al., 2008; Resende, 2016). Em trabalho realizado com dados simulados, foi verificado que o uso de maiores densidades de marcadores é necessário, para que se obtenha acurácia de predição de alta magnitude (> 70%) (Valente et al., 2016).

As estimativas dos valores de herdabilidade genômica, capacidade preditiva da GWS, viés da predição, acurácia com base nos dados fenótipos, acurácia da GWS e número de QTLs que controlam a característica estão apresentadas na Tabela 4.

**Tabela 4.** Resultados da seleção genômica ampla (GWS) para oito características morfoagronômicas em população de melhoramento de *Coffea canephora*.<sup>(1)</sup>

Característica	$h^2_a \pm sd$	$rgy \pm sd$	$b \pm sd$	$ryy$	$Rgg$
Vig	0,43±0,04	0,44±0,15	1,18±0,50	0,6	0,67
Fer	0,37±0,04	0,48±0,15	1,40±0,67	0,39	0,79
Cer	0,43±0,04	0,54±0,20	1,31±0,57	-	0,82
AP	0,36±0,04	0,41±0,22	1,35±0,86	0,67	0,68
DC	0,53±0,03	0,58±0,13	1,23±0,41	0,57	0,8
Mat	0,21±0,19	-0,03±0,19	1,08±2,99	-	-
TF	0,21±0,19	0,00±0,29	0,60±1,42	-	-
Prod	0,15±0,02	-0,02±0,30	-0,11±2,12	-	-

<sup>(1)</sup> $h^2_a$  = herdabilidade genômica; sd = desvio padrão; rgy = capacidade preditiva; b = viés da predição; ryy = acurácia da seleção com base nos dados fenótipos obtidos pelo método REML/BLUP; rgg = acurácia da GWS; Vig = vigor vegetativo; Fer = incidência de ferrugem; Cer = incidência de cercosporiose; AP = altura da planta; DC = diâmetro da projeção da copa; Mat = época de maturação dos frutos; TF = tamanho de fruto; Prod = produção por planta.

Os valores estimados de herdabilidade genômica ( $h^2_a$ ) variaram de 0,15 para a característica produção por planta (Prod) a 0,53 para a característica diâmetro da projeção da copa (DC). Esses resultados indicam que a seleção genômica é herdável. Embora valores consideráveis de  $h^2_a$  tenham sido obtidos para as características época de maturação dos frutos (0,21), tamanho de fruto (0,21) e produção por planta (0,15), não foi possível obter boa capacidade preditiva para essas características. Espera-se que as estimativas de capacidade preditiva sejam menores para as características que apresentam baixa herdabilidade (Legarra et al., 2008). Nesse sentido, como os menores valores de  $h^2_a$  foram

obtidos para essas três características, pode-se justificar o fato de não terem sido obtidos bons valores de capacidade preditiva para tais características.

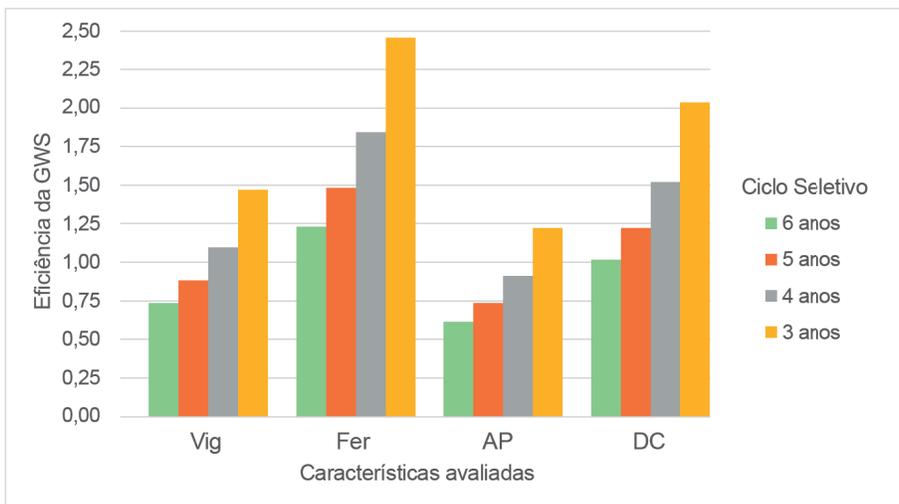
A capacidade preditiva ( $rgy$ ) equivale à correlação entre os valores genéticos genômicos e os valores fenotípicos observados e reflete a consistência das informações moleculares em informar sobre o fenótipo (Cavalcanti et al., 2012). Foi observada boa capacidade preditiva para as características Vig (0,44), Fer (0,48), Cer (0,54), AP (0,41) e DC (0,58), indicando a capacidade de se antever fenótipos para essas características. Esses dados mostram que, de modo geral, os valores de  $rgy$  foram maiores para as características que apresentaram os maiores valores de  $h^2_a$ .

Assim como observado nesse trabalho, foi verificada resposta da capacidade preditiva em função da herdabilidade em estudo de GWS com cajueiro (*Anacardium occidentale*) (Cavalcanti et al., 2012). O viés da predição ( $b$ ), obtido por meio do coeficiente de regressão dos valores fenotípicos sobre os valores genômicos preditos, apresentou valores próximos de 1,0 para as características: vigor vegetativo, incidência de ferrugem, incidência de cercosporiose, altura da planta, diâmetro da projeção da copa, época de maturação dos frutos e tamanho de fruto. Valores de viés próximos de 1,0 indicam que a predição foi não viesada e, portanto, são efetivas em prever as reais magnitudes das diferenças entre os indivíduos avaliados (Resende et al., 2012). A característica produção por planta não apresentou viés da predição próximo da unidade, indicando, para essa característica, que a predição foi viesada. Essa característica apresentou o menor valor de estimativa de herdabilidade genômica, sendo uma possível justificativa para o viés observado. Além disso, características governadas por maior quantidade de genes exigem populações com maior tamanho amostral.

A acurácia da GWS ( $rgg$ ) reflete a quantidade e qualidade das informações e procedimentos utilizados na predição dos valores genéticos genômicos dos indivíduos. Ela se baseia na correlação entre o valor genético verdadeiro e aquele estimado ou predito (Rabier et al., 2016). A acurácia pode ser classificada em muito alta, quando maior que 90%; alta, quando entre 70% e 90%; moderada, entre 50% e 70%, e baixa, quando menor que 50% (Resende; Duarte, 2007). Para as características época de maturação dos frutos, tamanho de fruto e produção por planta, não foi verificada capacidade preditiva satisfatória e, portanto, para tais características, não foram estimadas

as acurácias. Para as demais características, as estimativas dos valores de acurácia foram obtidas. Os valores de acurácia observados nesse trabalho foram considerados de magnitude moderada a alta, variando entre 67% (Vig) e 82% (Fer). Valores moderados a altos de r<sub>gg</sub>, 68% e 79%, foram obtidos até mesmo para as características altura da planta e incidência de ferrugem, que apresentaram menores valores de ha, 0,36 e 0,37, respectivamente. Esses resultados confirmam a eficiência da GWS na seleção de características que apresentam baixa herdabilidade, estando de acordo com outros trabalhos (Legarra et al., 2008; Zhang et al., 2010).

Na Figura 2 é apresentada a eficiência da GWS com a redução do ciclo seletivo em relação à seleção baseada somente em dados fenotípicos, de seis anos, para todas as características que apresentaram boa capacidade preditiva, exceto incidência de cercosporiose (Cer). Para essa variável, a herdabilidade no sentido amplo obtida com base nos dados fenotípicos no *software* Selegen REML/BLUP foi igual a zero e, portanto, não foi incluída na análise de eficiência da GWS. Dessa forma, foi possível estimar a eficiência da GWS somente para as características vigor vegetativo, incidência de ferrugem, altura da planta e diâmetro da projeção da copa. Em espécies perenes, como *C. canephora*, uma das vantagens da seleção genômica é a redução do ciclo de seleção ao praticar a seleção precoce.



**Figura 2.** Eficiência da GWS em relação à seleção baseada somente em dados fenotípicos, de seis anos, em população de melhoramento de *Coffea canephora*, para as variáveis vigor vegetativo (Vig), incidência de ferrugem (Fer), altura da planta (AP) e diâmetro da projeção da copa (DC).

Com base na Figura 2 é possível verificar o aumento da eficiência seletiva com o uso da GWS, mesmo para características como vigor vegetativo e altura da planta, as quais apresentaram altos valores de estimativa de acurácia a partir da seleção baseada em dados fenotípicos, 60% e 67%, respectivamente (Tabela 3). Esse aumento da eficiência se deve à redução do tempo necessário para completar um ciclo seletivo com o uso da GWS. Dessa forma, com a redução do ciclo, variando de seis a três anos, tem-se um aumento na eficiência seletiva da GWS para todas as características ao longo da redução do ciclo. Portanto, mesmo quando a acurácia da seleção genômica seja de mesma magnitude que a obtida com a seleção baseada em dados fenotípicos, a GWS proporcionará ganhos genéticos superiores, devido à redução do ciclo de seleção (Gois et al., 2016).

Com a redução do ciclo seletivo de seis para três anos, a GWS foi superior (variando entre 22% a 146%) para todas as características. Dessa forma, observa-se que a redução no tempo necessário para completar um ciclo seletivo é expressiva com a GWS. Isso se deve ao fato de que a predição genômica e a seleção podem ser feitas no estádio de plântulas, e, portanto, a GWS é mais eficiente por unidade de tempo (Resende et al., 2012; Gois et al., 2016). Resultados semelhantes foram observados em outros trabalhos. Em trabalho avaliando a eficiência seletiva da GWS com redução de 50% no ciclo seletivo do Citrus, a GWS foi superior para todas as características avaliadas (variando entre 31% a 160%) (Gois et al., 2016). Também foi verificado, em trabalho conduzido com dendê (*Elaeis guineensis* Jacq.), a redução do ciclo de seleção de 19 para seis anos com o uso da GWS (Wong; Bernardo, 2008). Em trabalho conduzido com milho (*Zea mays* L.) foi observado que, com o uso da GWS, houve aumento significativo na acurácia seletiva e nos ganhos genéticos por unidade de tempo (Fritsche-Neto et al., 2012). Em outros trabalhos, a seleção genômica também revelou o grande potencial do seu uso em aumentar a eficiência do melhoramento, via resultados de dados simulados (Resende et al., 2008; Valente et al., 2016).

Mesmo para as características que apresentaram menores valores de ha, altura da planta (0,36) e incidência de ferrugem (0,37), foram obtidos ganhos em eficiência com a GWS. Esses dados confirmam a importância da GWS na seleção de características que apresentam

baixa herdabilidade, fato este também observado em outros trabalhos (Resende et al., 2012; Gois et al., 2016). Os dados obtidos nesse trabalho demonstram que a seleção genômica ampla é ferramenta útil para o melhoramento de *C. canephora*, por predizer com acurácia os fenótipos dos indivíduos e possibilitar a redução no tempo necessário para completar o ciclo de seleção, proporcionando ganhos em eficiência seletiva por unidade de tempo.

## Conclusões

Na seleção genômica, os efeitos genéticos dos marcadores sobre fenótipos de caracteres quantitativos são somados e usados na predição de valores genéticos de indivíduos apenas genotipados, candidatos à seleção em programas de melhoramento genético. A predição e a seleção podem ser realizadas em fases muito juvenis das plantas, acelerando assim o processo de melhoramento genético. Adicionalmente, a própria predição tende a ser mais acurada por considerar o real parentesco genético dos indivíduos em avaliação, em detrimento do parentesco médio esperado matematicamente. A seleção genômica propicia uma forma de seleção precoce direta (SPD), pois atua precocemente sobre genes expressos na idade adulta. Assim, a SPD propiciada pela GWS é especialmente importante para o melhoramento de organismos perenes como animais, espécies florestais, fruteiras (e outras frutíferas), forrageiras, cana-de-açúcar, café, dentre outras.

Em relação ao café arábica, a seleção genômica ampla mostrou-se potencial para seu melhoramento, por possibilitar a obtenção acurada dos GEBVs e pela maior eficiência em relação à seleção fenotípica, possibilitando a redução no tempo necessário para completar o ciclo de seleção.

As herdabilidades fenotípicas e as acurácias seletivas, com base nos dados fenotípicos, foram estimadas e apresentaram magnitudes moderadas a altas, para a maioria das características avaliadas. Os resultados das análises de GWS foram promissores, sendo a GWS mais eficiente que a seleção utilizando exclusivamente os dados fenotípicos e indicaram a possibilidade de se maximizar os ganhos com seleção por unidade de tempo. O efeito da densidade de marcadores nas análises de GWS foi avaliado. Foi possível concluir que a seleção genômica ampla mostrou-se potencial para o

melhoramento de *C. arabica* por possibilitar a obtenção acurada dos GEBVs e pela maior eficiência em relação à seleção tradicional, possibilitando a redução no tempo necessário para completar o ciclo de seleção.

No que se refere ao café canéfora, os resultados obtidos nesse trabalho demonstram que a seleção genômica ampla é ferramenta útil para seu melhoramento, por prever com acurácia os fenótipos dos indivíduos e possibilitar a redução no tempo necessário para completar o ciclo de seleção, proporcionando ganhos em eficiência seletiva por unidade de tempo.

Foi observada boa capacidade preditiva para as características, exceto para época de maturação dos frutos, tamanho de fruto e produção por planta. Os menores valores de herdabilidade genômica encontrados para essas características podem justificar os baixos valores de capacidade preditiva obtidos. Os valores de acurácia estimados foram considerados de magnitude moderada a alta, variando entre 67% a 82%. Com a redução do tempo necessário para completar o ciclo seletivo de seis para três anos, a GWS proporcionou eficiência seletiva variando entre 22% a 146%. Esses resultados indicam que é possível alcançar maiores ganhos por unidade de tempo por meio da GWS. Portanto, a GWS mostrou ser ferramenta útil para o melhoramento genético de *C. canephora*, por prever com acurácia os fenótipos dos indivíduos e possibilitar a redução no tempo necessário para completar o ciclo de seleção, proporcionando ganhos em eficiência seletiva por unidade de tempo.

## Referências

- ADHIKARI, S.; SAHA, S.; BISWAS, A.; RANA, T. S.; BANDYOPADHYAY, T. K.; GHOSH, P. Application of molecular markers in plant genome analysis: a review. **The Nucleus**, v. 60, n. 3, p. 283-297, Sept. 2017. DOI: 10.1007/s13237-017-0214-7.
- ALKIMIM, E. R.; CAIXETA, E. T.; SOUSA, T. V.; PEREIRA, A. A.; OLIVEIRA, A. C. B. de; ZAMBOLIM, L.; SAKIYAMA, N. S. Marker-assisted selection provides arabica coffee with genes from other *Coffea* species targeting on multiple resistance to rust and coffee berry disease. **Molecular Breeding**, v. 37, article number 6, Jan. 2017. DOI: 10.1007/s11032-016-0609-1.
- ALKIMIM, E. R.; CAIXETA, E. T.; SOUSA, T. V.; RESENDE, M. D. V.; SILVA, F. L. da; SAKIYAMA, N. S.; ZAMBOLIM, L. Selective efficiency of genome-wide selection in *Coffea canephora* breeding. **Tree Genetics & Genomes**, v. 16, article number 41, Mar. 2020. DOI: 10.1007/s11295-020-01433-3.

ALMEIDA, D. P. de; CASTRO, I. S. L.; MENDES, T. A. de O.; ALVES, D. R.; BARKA, G. D.; BARREIROS, P. R. R. M.; ZAMBOLIM, L.; SAKIYAMA, N. S.; CAIXETA, E. T. Receptor-Like Kinase (RLK) as a candidate gene conferring resistance to *Hemileia vastatrix* in coffee. **Scientia Agricola**, v. 78, n. 6, e20200023, 2021. DOI: 10.1590/1678-992x-2020-0023.

ANDERSON, C. A.; PETTERSSON, F. H.; CLARKE, G. M.; CARDON, L. R.; MORRIS, A. P.; ZONDERVAN, K. T. Data quality control in genetic case-control association studies. **Nature Protocols**, v. 5, n. 9, p. 1564-1573, 2010. DOI: 10.1038/nprot.2010.116.

AVELINO, J.; CRISTANCHO, M.; GEORGIU, S.; IMBACH, P.; AGUILAR, L.; BORNEMANN, G.; LÄDERACH, P.; ANZUETO, F.; HRUSKA, A. J.; MORALES, C. The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. **Food Security**, v. 7, n. 2, p. 303-321, 2015. DOI: 10.1007/s12571-015-0446-9.

BABOVA, O.; OCCHIPINTI, A.; MAFFEI, M. E. Chemical partitioning and antioxidant capacity of green coffee (*Coffea arabica* and *Coffea canephora*) of different geographical origin. **Phytochemistry**, v. 123, p. 33-39, Mar. 2016. DOI: 10.1016/j.phytochem.2016.01.016.

BALACHIRANJEEVI, C. H.; BHASKAR, N. S.; ABHILASH, V.; AKANKSHA, S.; VIRAKTAMATH, B. C.; MADHAV, M. S.; HARIPRASAD, A. S.; LAHA, G. S.; PRASAD, M. S.; BALACHANDRAN, S. M.; NEERAJA, C. N.; SATENDRA KUMAR, M.; SENGUTTUVEL, P.; KEMPARAJU, K. B.; BHADANA, V. P.; RAM, T.; HARIKA, G.; MAHADEVASWANY, H. K.; HAJIRA, S. K.; YUGANDER, A.; PRANATHI, k.; ANILA, M.; REKHA, G.; KOUSIK, M. B. V. N.; DILIP KUMAR, T.; SWAPNIL, R. K.; ARCHANA GIRI; SUNDARAM, R. M. Marker-assisted introgression of bacterial blight and blast resistance into DRR17B, an elite, fine-grain type maintainer line of rice. **Molecular Breeding**, v. 35, n. 7, article number 151, 2015. DOI: 10.1007/s11032-015-0348-8.

BARBOSA, I. de P.; COSTA, W. G. DA; NASCIMENTO, M.; CRUZ, C. D.; OLIVEIRA, A. C. B. de Recommendation of *Coffea arabica* genotypes by factor analysis. **Euphytica**, v. 215, article number 178, 2019. DOI: 10.1007/s10681-019-2499-x.

BARKA, G. D.; CAIXETA, E. T.; FERREIRA, S. S.; ZAMBOLIM, L. In silico guided structural and functional analysis of genes with potential involvement in resistance to coffee leaf rust: a functional marker based approach. **PLoS ONE**, v. 15, e0222747, 2020. DOI: 10.1371/journal.pone.0222747.

BEGUM, H.; SPINDEL, J. E.; LALUSIN, A.; BORROMEO, T.; GREGORIO, G.; HERNANDEZ, J.; VIRK, P.; COLLARD, B.; McCOUCH, S. R. Genome-wide association mapping for yield and other agronomic traits in an elite breeding population of tropical rice (*Oryza sativa*). **PLoS ONE**, v. 10, e0119873, Mar. 2015. DOI: 10.1371/journal.pone.0119873.

BERNARDO, R.; YU, J. Prospects for genomewide selection for quantitative traits in maize. **Crop Science**, v. 47, n. 3, p. 1082-1090, May-June 2007. DOI: 10.2135/cropsci2006.11.0690.

BETTENCOURT, A.; RODRIGUES JUNIOR, C. J. Principles and practice of coffee breeding for resistance to rust and other diseases. In: CLARKE, R. J.; MACRAE, R. (ed.). **Coffee: agronomy**. London: Elsevier Applied Science Publishers, 1988. v. 4. p. 199-234.

BRITO, G. G. de; CAIXETA, E. T.; GALLINA, A. P.; ZAMBOLIM, E. M.; ZAMBOLIM, L.; DIOLA, V.; LOUREIRO, M. E. Inheritance of coffee leaf rust resistance and identification of AFLP markers linked to the resistance gene. **Euphytica**, v. 173, n. 2, p. 255264, May 2010. DOI: 10.1007/s10681-010-0119-x.

CAMPOS, G. de los; HICKEY, J. M.; PONG-WONG, R.; DAETWYLER, H. D.; CALUS, M. P. L. Whole-genome regression and prediction methods applied to plant and animal breeding. **Genetics**, v. 193, n. 2, p. 327-345, Feb. 2013. DOI: 10.1534/genetics.112.143313.

CAMPOS, G. de los; NAYA, H.; GIANOLA, D.; CROSSA, J.; LEGARRA, A.; MANFREDI, E.; WEIGEL, K.; COTES, J. M. Predicting quantitative traits with regression models for dense molecular markers and pedigree. **Genetics**, v. 182, n. 1, p. 375-385, May 2009. DOI: 10.1534/genetics.109.101501.

CARVALHO, A. **Histórico do desenvolvimento do cultivo do café no Brasil**. Campinas: Instituto Agrônomo, 2007. 8 p. (IAC. Documentos, 34). Disponível em: <http://www.iac.sp.gov.br/publicacoes/arquivos/iacdoc34.pdf>. Acesso em: 3 fev. 2021.

CARVALHO, A. Melhoramento do cafeeiro: VI - estudo e interpretação, para fins de seleção, de produções individuais na variedade bourbon. **Bragantia**, v. 12, n. 4-6, p. 179-200, June 1952. DOI: 10.1590/S0006-87051952000200007.

CARVALHO, H. F.; GALLI, G.; FERRÃO, L. F. V.; NONATO, J. V. A.; PADILHA, L.; MALUF, M. P.; RESENDE JÚNIOR, M. F. R. de; GUERREIRO FILHO, O.; FRITSCHÉ-NETO, R. The effect of bienniality on genomic prediction of yield in arabica coffee. **Euphytica**, v. 216, article number 101, June 2020. DOI: 10.1007/s10681-020-02641-7.

CAVALCANTI, J. J. V.; RESENDE, M. D. V. de; SANTOS, F. H. C. dos; PINHEIRO, C. R. Predição simultânea dos efeitos de marcadores moleculares e seleção genômica ampla em cajueiro. **Revista Brasileira de Fruticultura**, v. 34, n. 3, p. 840-846, set. 2012. DOI: 10.1590/S0100-29452012000300025.

DENOEU, F.; CARRETERO-PAULET, L.; DEREPPER, A.; DROC, G.; GUYOT, R.; PIETRELLA, M.; ZHENG, C.; ALBERTI, A.; ANTHONY, F.; APREA, G.; AURY, J.-M.; BENTO, P.; BERNARD, M.; BOCS, S.; CAMPA, C.; CENCI, A.; COMBES, M.-C.; CROUZILLAT, D.; SILVA, C. da; DADDIEGO, L.; BELLIS, F. de; DUSSERT, S.; GARSMEUR, O.; GAYRAUD, T.; GUIGNON, V.; JAHN, K.; JAMILLOUX, V.; JOET, T.; LABADIE, K.; LAN, T.; LECLERCQ, J.; LEPELLEY, M.; LEROY, T.; LI, L.-T.; LIBRADO, P.; LOPEZ, L.; MUÑOZ, A.; NOEL, B.; PALLAVICINI, A.; PERROTTA, G.; PONCET, V.; POT, D.; PRIYONO; RIGOREAU, M.; ROUARD, M.; ROZAS, J.; TRANCHANT-DUBREUIL, C.; VANBUREN, R.; ZHANG, Q.; ANDRADE, A. C.; ARGOUT, X.; BERTRAND, B.; KOCHKO, A. de; GRAZIOSI, G.; HENRY, R. J.; JAYARAMA; MING, R.; NAGAI, C.; ROUNSLEY, S.; SANKOFF, D.; GIULIANO, G.; ALBERT, V. A.; WINCKER, P.; LASHERMES, P. The coffee genome provides insight into the convergent evolution of caffeine biosynthesis. **Science**, v. 345, n. 6201, p. 1181-1184, Sept. 2014. DOI: 10.1126/science.1255274.

DESTA, Z. A.; ORTIZ, R. Genomic selection: genome-wide prediction in plant improvement. **Trends in Plant Science**, v. 19, n. 9, p. 592-601, Sept. 2014. DOI: 10.1016/j.tplants.2014.05.006.

DIAS, R. E. B. A.; DIAS, R. A. A.; BOTELHO, C. E.; ABRAHÃO, J. C. de R.; REZENDE, T. T.; CARVALHO, G. R. Genetic determination of characteristics related to semi-mechanized coffee harvests. **Crop Breeding and Applied Biotechnology**, v. 20, n. 1, e218820116, 2020. DOI: 10.1590/1984-70332020v20n1n16.

DIOLA, V.; BRITO, G. G. de; CAIXETA, E. T.; MACIEL-ZAMBOLIM, E.; SAKIYAMA, N. S.; LOUREIRO, M. E. High-density genetic mapping for coffee leaf rust resistance. **Tree Genetics & Genomes**, v. 7, n. 6, p. 1199-1208, July 2011. DOI: 10.1007/s11295-011-0406-2.

FERNANDO, R. L.; HABIER, D.; STRICKER, C.; DEKKERS, J. C. M.; TOTIR, L. R. Genomic selection. **Acta Agriculturae Scandinavica, Section A - Animal Science**, v. 57, n. 4, p. 192-195, Dec. 2007. DOI: 10.1080/09064700801959395.

FERRÃO, L. F. V.; FERRÃO, R. G.; FERRÃO, M. A. G.; FONSECA, A. F. A. da; CARBONETTO, P.; STEPHENS, M.; GARCIA, A. A. F. Accurate genomic prediction of *Coffea canephora* in multiple environments using whole-genome statistical models. **Heredity**, v. 122, n. 3, p. 261-275, 2019. DOI: 10.1038/s41437-018-0105-y.

FERRÃO, L. F. V.; FERRÃO, R. G.; FERRÃO, M. A. G.; FRANCISCO, A.; GARCIA, A. A. F.; VENTORIM FERRÃO, L. F.; GAVA FERRÃO, R.; FERRÃO, M. A. G.; FONSECA, A. F. A. da; GARCIA, A. A. F. A mixed model to multiple harvest-location trials applied to genomic prediction in *Coffea canephora*. **Tree Genetics & Genomes**, v. 13, article number 95, 2017. DOI: 10.1007/s11295-017-1171-7.

FRITSCHÉ-NETO, R.; RESENDE, M. D. V.; MIRANDA, G. V.; DOVALE, J. C. Seleção genômica ampla e novos métodos de melhoramento do milho. **Revista Ceres**, v. 59, n. 6, p. 794-802, nov./dez. 2012. DOI: 10.1590/S0034-737X2012000600009.

GARRIDO-CARDENAS, J. A.; MESA-VALLE, C.; MANZANO-AGUGLIARO, F. Trends in plant research using molecular markers. **Planta**, v. 247, n. 3, p. 543-557, mar. 2018. DOI: 10.1007/s00425-017-2829-y.

GICHIMU, B. M.; GICHURU, E. K.; MAMATI, G. E.; NYENDE, A. B. Occurrence of Ck-1 gene conferring resistance to coffee berry disease in *Coffea arabica* cv. Ruiru 11 and its parental genotypes. **Journal of Agricultural and Crop Research**, v. 2, n. 3, p. 51-61, Mar. 2014. Disponível em: [http://www.sciencewebpublishing.net/jacr/archive/2014/March/pdf/Gichimu et al.pdf](http://www.sciencewebpublishing.net/jacr/archive/2014/March/pdf/Gichimu%20et%20al.pdf). Acesso em: 3 fev. 2021.

GICHURU, E. K.; AGWANDA, C. O.; COMBES, M. C.; MUTITU, E. W.; NGUGI, E. C. K.; BERTRAND, B.; LASHERMES, P. Identification of molecular markers linked to a gene conferring resistance to coffee berry disease (*Colletotrichum kahawae*) in *Coffea arabica*. **Plant Pathology**, v. 57, n. 6, p. 1117-1124, Dec. 2008. DOI: 10.1111/j.1365-3059.2008.01846.x.

GIMASE, J. M.; THAGANA, W. M.; OMONDI, C. O.; CHESEREK, J. J.; GICHIMU, B. M.; GICHURU, E. K.; ZIYOMO, C.; SNELLER, C. H. Genome-Wide Association Study identify the genetic loci conferring resistance to coffee berry disease (*Colletotrichum kahawae*) in *Coffea arabica* var. Rume Sudan. **Euphytica**, v. 216, article number 86, May 2020. DOI: 10.1007/s10681-020-02621-x.

GIMASE, J. M.; THAGANA, W. M.; OMONDI, C. O.; ITHIRU, J. M. Evaluation of coffee berry disease resistance (*Colletotrichum kahawae*) in F2 populations derived from Arabica coffee varieties Rume Sudan and SL 28. **Journal of Plant Breeding and Crop Science**, v. 11, n. 9, p. 225-233, article number 0375FAF62286, Oct./Dec. 2019. DOI: 10.5897/JPBCS2019.0829.

GODDARD, M. E.; HAYES, B. J. Genomic selection. **Journal of Animal Breeding and Genetics**, v. 124, n. 6, p. 323-330, Dec. 2007. DOI: 10.1111/j.1439-0388.2007.00702.x.

GOIS, I. B.; BORÉM, A.; CRISTOFANI-YALY, M.; RESENDE, M. D. V. de; AZEVEDO, C. F.; BASTIANEL, M.; NOVELLI, V. M.; MACHADO, M. A. Genome wide selection in citrus breeding. **Genetics and Molecular Research**, v. 15, n. 4, gmr15048863, Oct. 2016. DOI: 10.4238/gmr15048863.

GRATTAPAGLIA, D.; RESENDE, M. D. V. de. Genomic selection in forest tree breeding. **Tree Genetics & Genomes**, v. 7, n. 2, p. 241-255, Apr. 2011. DOI: 10.1007/s11295-010-0328-4.

HAYES, B. J.; BOWMAN, P. J.; CHAMBERLAIN, A. J.; GODDARD, M. E. Invited review: genomic selection in dairy cattle: progress and challenges. **Journal of Dairy Science**, v. 92, n. 2, p. 433-443, Feb. 2009. DOI: 10.3168/jds.2008-1646.

HEFFNER, E. L.; JANNINK, J.-L.; IWATA, H.; SOUZA, E.; SORRELLS, M. E. Genomic selection accuracy for grain quality traits in biparental wheat populations. **Crop Science**, v. 51, n. 6, p. 2597-2606, Nov. 2011b. DOI: 10.2135/cropsci2011.05.0253.

HEFFNER, E. L.; JANNINK, J.-L.; SORRELLS, M. E. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. **The Plant Genome**, v. 4, n. 1, p. 65-75, Mar. 2011a. DOI: 10.3835/plantgenome2010.12.0029.

HEFFNER, E. L.; SORRELLS, M. E.; JANNINK, J.-L. Genomic selection for crop improvement. **Crop Science**, v. 49, n. 1, p. 1-12, Jan. 2009. DOI: 10.2135/cropsci2008.08.0512.

HESLOT, N.; YANG, H.-P.; SORRELLS, M. E.; JANNINK, J.-L. Genomic selection in plant breeding: a comparison of models. **Crop Science**, v. 52, n. 1, p. 146-160, Jan. 2012. DOI: 10.2135/cropsci2011.06.0297.

LANDE, R.; THOMPSON, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. **Genetics**, v. 124, n. 3, p. 743-756, Mar. 1990. Disponível em: <https://www.genetics.org/content/genetics/124/3/743.full.pdf>. Acesso em: 3 fev. 2021.

LEGARRA, A.; ROBERT-GRANIÉ, C.; MANFREDI, E.; ELSÉN, J.-M. Performance of genomic selection in mice. **Genetics**, v. 180, n. 1, p. 61-618, Sept. 2008. DOI: 10.1534/genetics.108.088575.

MAHÉ, L.; COMBES, M.-C.; VÁRZEA, V. M. P.; GUILHAUMON, C.; LASHERMES, P. Development of sequence characterized DNA markers linked to leaf rust (*Hemileia vastatrix*) resistance in coffee (*Coffea arabica* L.). **Molecular Breeding**, v. 21, p. 105-113, 2008. DOI: 10.1007/s11032-007-9112-z.

MEDINA FILHO, H. P.; BORDIGNON, R.; CARVALHO, C. H. S. de. Desenvolvimento de novas cultivares de café arábica. In: CARVALHO, C. H. S. de (ed.). **Cultivares de café: origem, características e recomendações**. Brasília, DF: Embrapa Café, 2008. cap. 5, p. 79-102.

MENDES, A. N. G.; CARVALHO, G. R.; BOTELHO, C. E.; FAZUOLI, L. C.; SILVAROLLA, M. B. História das primeiras cultivares de café plantadas no Brasil. In: CARVALHO, C. H. S. de (ed.). **Cultivares de café: origem, características e recomendações**. Brasília, DF: Embrapa Café, 2008. cap. 4, p. 69-78.

MEROT-L'ANTHOENE, V.; TOURNEBIZE, R.; DARRACQ, O.; RATTINA, V.; LEPELLEY, M.; BELLANGER, L.; TRANCHANT-DUBREUIL, C.; COULÉE, M.; PÉGARD, M.; METAIRON, S.; FOURNIER, C.; STOFFELEN, P.; JANSSENS, S. B.; KIWUKA, C.; MUSOLI, P.; SUMIRAT, U.; LEGNATÉ, H.; KAMBALE, J.-L.; COSTA NETO, J. F. da; REVEL, C.; KOCHKO, A. de; DESCOMBES, P.; CROUZILLAT, D.; PONCET, V. Development and evaluation of a genome-wide Coffee 8.5K <sc>SNP</sc> array and its application for high-density genetic mapping and for investigating the origin of *Coffea arabica* L. **Plant Biotechnology Journal**, v. 17, n. 7, p. 1418-1430, July 2019. DOI: 10.1111/pbi.13066.

MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, n. 4, p. 1819-1829, Apr. 2001. DOI: 10.1093/genetics/157.4.1819.

MICHELMORE, R. W.; PARAN, I.; KESSELI, R. V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. **Proceedings of the National Academy of Sciences**, v. 88, n. 21, p. 9828-9832, Nov. 1991. DOI: 10.1073/pnas.88.21.9828.

MISHRA, M. K. Genetic Resources and Breeding of Coffee (*Coffea* spp.). In: AL-KHAYRI, J. M.; JAIN, S. M.; JOHNSON, D. V. (ed.). **Advances in plant breeding strategies: nut and beverage crops**. Cham: Springer International Publishing, 2019. v. 4. p. 475-515. DOI:10.1007/978-3-030-23112-5\_12.

MONCADA, M. D. P.; TOVAR, E.; MONTOYA, J. C.; GONZÁLEZ, A.; SPINDEL, J.; McCOUCH, S. A genetic linkage map of coffee (*Coffea arabica* L.) and QTL for yield, plant height, and bean size. **Tree Genetics & Genomes**, v. 12, article number 5, 2016. DOI: 10.1007/s11295-015-0927-1.

NOIR, S.; ANTHONY, F.; BERTRAND, B.; COMBES, M.-C.; LASHERMES, P. Identification of a major gene (Mex-1) from *Coffea canephora* conferring resistance to Meloidogyne exigua in *Coffea arabica*. **Plant Pathology**, v. 52, n. 1, p. 97-103, Feb. 2003. DOI: 10.1046/j.1365-3059.2003.00795.x.

OLIVEIRA, E. J. de; RESENDE, M. D. V. de; SANTOS, V. da S.; FERREIRA, C. F.; OLIVEIRA, G. A. F.; SILVA, M. S. da; OLIVEIRA, L. A. de; AGUILAR-VILDOSO, C. I. Genome-wide selection in cassava. **Euphytica**, v. 187, p. 263-276, Sept. 2012. DOI: 10.1007/s10681-012-0722-0.

ORNELLA, L.; SINGH, S.; PEREZ, P.; BURGUEÑO, J.; SINGH, R.; TAPIA, E.; BHAVANI, S.; DREISIGACKER, S.; BRAUN, H.-J.; MATHEWS, K.; CROSSA, J. Genomic prediction of genetic values for resistance to wheat rusts. **The Plant Genome**, v. 5, n. 3, p. 136-148, Nov. 2012. DOI: 10.3835/plantgenome2012.07.0017.

PEIXOTO, L. de A.; LAVIOLA, B. G.; ALVES, A. A.; ROSADO, T. B.; BHERING, L. L. Breeding *Jatropha curcas* by genomic selection: A pilot assessment of the accuracy of predictive models. **PLoS ONE**, v. 12, n. 3, e0173368, 2017. DOI: 10.1371/journal.pone.0173368.

PEREIRA, A. A.; CARVALHO, G. R.; MOURA, W. de M.; BOTELHO, E. E.; REZENDE, J. C. de; OLIVEIRA, A. C. B. de; SILVA, F. L. da. Cultivares: origem e suas características. In: REIS, P. R.; CUNHA, R. L. (ed.). **Café arábica: do plantio à colheita**. Lavras: Epamig, 2010. v. 1, cap. 3, p. 163-222.

PÉREZ, P.; CAMPOS, G. de los. Genome-wide regression and prediction with the BGLR statistical package. **Genetics**, v. 198, n. 2, p. 483-495, Oct. 2014. DOI: 10.1534/genetics.114.164442.

PRAKASH, N. S.; MARQUES, D. V.; VARZEA, V. M. P.; SILVA, M. C.; COMBES, M. C.; LASHERMES, P. Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into *C. arabica* L. **Theoretical and Applied Genetics**, v. 109, n. 6, p. 1311-1317, Oct. 2004. DOI: 10.1007/s00122-004-1748-z.

PRICE, A. L.; ZAITLEN, N. A.; REICH, D.; PATTERSON, N. New approaches to population stratification in genome-wide association studies. **Nature Reviews Genetics**, v. 11, n. 7, p. 459-463, July 2010. DOI: 10.1038/nrg2813.

- PRUVOT-WOEHL, S.; KRISHNAN, S.; SOLANO, W.; SCHILLING, T.; TONIUTTI, L.; BERTRAND, B.; MONTAGNON, C. Authentication of *Coffea arabica* varieties through DNA fingerprinting and its significance for the coffee sector. **Journal of AOAC International**, v. 103, n. 2, p. 325–334, Mar./Apr. 2020. DOI: 10.1093/jaoacint/qs2003.
- RABIER, C.-E.; BARRE, P.; ASP, T.; CHARMET, G.; MANGIN, B. On the accuracy of genomic selection. **PLoS ONE**, v. 11, n. 6, p. e0156086, June 2016. DOI: 10.1371/journal.pone.0156086.
- RESENDE JÚNIOR, M. F. R.; MUÑOZ, P.; ACOSTA, J. J.; PETER, G. F.; DAVIS, J. M.; GRATTAPAGLIA, D.; RESENDE, M. D. V.; KIRST, M. Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. **New Phytologist**, v. 193, n. 3, p. 617–624, Feb. 2012a. DOI: 10.1111/j.1469-8137.2011.03895.x.
- RESENDE JÚNIOR, M. F. R.; MUÑOZ, P.; RESENDE, M. D. V.; GARRICK, D. J.; FERNANDO, R. L.; DAVIS, J. M.; JOKELA, E. J.; MARTIN, T. A.; PETER, G. F.; KIRST, M. Accuracy of genomic selection methods in a standard data set of loblolly pine (*Pinus taeda* L.). **Genetics**, v. 190, n. 4, p. 1503–1510, Apr. 2012b. DOI: 10.1534/genetics.111.137026.
- RESENDE, M. D. V. de. Desequilíbrio de ligação e seleção genômica. In: RESENDE, M. D. V. de. **Genética quantitativa e de populações**. Viçosa, MG: UFV, 2015. p. 289–311.
- RESENDE, M. D. V. de. Estimação de componentes de variância via REML. In: RESENDE, M. D. V. de. **Genética biométrica e estatística no melhoramento de plantas perenes**. Brasília, DF: Embrapa Informação Tecnológica; Colombo: Embrapa Florestas, 2002. p. 384–476.
- RESENDE, M. D. V. de. Seleção genômica ampla (GWS) e modelos linearers mistos. In: RESENDE, M. D. V. de. **Matemática e estatística na análise de experimentos e no melhoramento genético**. Colombo: Embrapa Florestas, 2007. p. 517–533.
- RESENDE, M. D. V. de. *Software* Selegen-REML/BLUP: a useful tool for plant breeding. **Crop Breeding and Applied Biotechnology**, v. 16, n. 4, p. 330–39, Oct./Dec. 2016. DOI: 10.1590/1984-70332016v16n4a49.
- RESENDE, M. D. V. de; AZEVEDO, C. F.; SILVA, F. F. e. Seleção genômica. In: RESENDE, M. D. V. de; SILVA, F. F. e; AZEVEDO, C. F. **Estatística matemática, biométrica e computacional: modelos mistos, multivariados, categóricos e generalizados (REML/BLUP), inferência bayesiana, regressão aleatória, seleção genômica, QTI-GWAS, estatística espacial e temporal, competição, sobrevivência**. Viçosa, MG: UFV, 2014. p. 627–768.
- RESENDE, M. D. V. de; DUARTE, J. B. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, v. 37, n. 3, p. 182–194, 2007. Disponível em: <https://www.alice.cnptia.embrapa.br/alice/bitstream/doc/313450/1/Pecisaoecontroledequidadeemexperimentos.pdf>. Acesso em: 3 fev. 2021.
- RESENDE, M. D. V. de; LOPES, P. S.; SILVA, R. L. da; PIRES, I. E. Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. **Pesquisa Florestal Brasileira**, v. 56, p. 63–77, jan./jun. 2008. Disponível em: <https://pfb.cnpf.embrapa.br/ojs-3.2.1-3/index.php/pfb/article/view/63/65>. Acesso em: 3 fev. 2021
- RESENDE, M. D. V. de; SILVA, F. F. e; LOPES, P. S.; AZEVEDO, C. F. Fundamentos da genome Wide selection. In: RESENDE, M. D. V. de; SILVA, F. F. e; LOPES, P. S.; AZEVEDO, C. F. **Seleção genômica ampla (GWS) via modelos mistos (REML/BLUP), inferência bayesiana (MCMC), regressão aleatória multivariada e estatística espacial**. Viçosa, MG: UFV, 2012. p. 136–159. Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/92345/1/Deon-livro-selecao-genomica-ampla.pdf>. Acesso em: 3 fev. 2021.

REVERTER, A.; JOHNSTON, D. J.; GRASER, H. U.; WOLCOTT, M. L.; UPTON, W. H. Genetic analyses of live-animal ultrasound and abattoir carcass traits in Australian Angus and Hereford cattle. **Journal of Animal Science**, v. 78, n. 7, p. 1786-1795, July 2000. DOI: 10.2527/2000.7871786x.

RIEDELSCHEIMER, C.; CZEDIK-EYSENBERG, A.; GRIEDER, C.; LISEC, J.; TECHNOW, F.; SULPICE, R.; ALTMANN, T.; STITT, M.; WILLMITZER, L.; MELCHINGER, A. E. Genomic and metabolic prediction of complex heterotic traits in hybrid maize. **Nature Genetics**, v. 44, n. 2, p. 217-220, Feb. 2012. DOI: 10.1038/ng.1033.

ROMERO, G.; VÁSQUEZ, L. M.; LASHERMES, P.; HERRERA, J. C. Identification of a major QTL for adult plant resistance to coffee leaf rust (*Hemileia vastatrix*) in the natural Timor hybrid (*Coffea arabica* x *C. canephora*). **Plant Breeding**, v. 133, n. 1, p. 121-129, Feb. 2014. DOI: 10.1111/pbr.12127.

SANT'ANA, G. C.; PEREIRA, L. F. P.; POT, D.; IVAMOTO, S. T.; DOMINGUES, D. S.; FERREIRA, R. V.; PAGIATTO, N. F.; SILVA, B. S. R. da; NOGUEIRA, L. M.; KITZBERGER, C. S. G.; SCHOLZ, M. B. S.; OLIVEIRA, F. F. de; SERA, G. H.; PADILHA, L.; LABOUISSSE, J.-P.; GUYOT, R.; CHARMETANT, P.; LEROY, T. Genome-wide association study reveals candidate genes influencing lipids and diterpenes contents in *Coffea arabica* L. **Scientific Reports**, v. 8, article number 465, 2018. DOI: 10.1038/s41598-017-18800-1.

SAVADI, S.; PRASAD, P.; KASHYAP, P. L.; BHARDWAJ, S. C. Molecular breeding technologies and strategies for rust resistance in wheat (*Triticum aestivum*) for sustained food security. **Plant Pathology**, v. 67, n. 4, p. 771-791, May 2018. DOI: 10.1111/ppa.12802.

SINGH, S.; SIDHU, J. S.; HUANG, N.; VIKAL, Y.; LI, Z.; BRAR, D. S.; DHALIWAL, H. S.; KHUSH, G. S. Pyramiding three bacterial blight resistance genes (*xa5*, *xa13* and *Xa21*) using marker-assisted selection into indica rice cultivar PR106. **Theoretical and Applied Genetics**, v. 102, p. 1011-1015, May 2001. DOI: 10.1007/s001220000495.

SOBREIRA, F. M.; OLIVEIRA, A. C. B. de; PEREIRA, A. A.; MARTINS, A. G.; SAKIYAMA, N. S. Divergence among arabica coffee genotypes for sensory quality. **Australian Journal of Crop Science**, v. 10, n. 10, p. 1442-1448, Oct. 2016. DOI: 10.21475/ajcs.2016.10.10.p7430.

SOUSA, T. V.; CAIXETA, E. T. E.; ALKIMIM, E. R.; OLIVEIRA, A. C. B. de; PEREIRA, A. A.; SAKIYAMA, N. S.; ZAMBOLIM, L.; RESENDE, M. D. V. de. Early selection enabled by the implementation of genomic selection in *coffea arabica* breeding. **Frontiers in Plant Science**, v. 9, article 1934, 2019. DOI: 10.3389/fpls.2018.01934.

SOUSA, T. V.; CAIXETA, E. T.; ALKIMIM, E. R.; OLIVEIRA, A. C. B. de; PEREIRA, A. A.; ZAMBOLIM, L.; SAKIYAMA, N. S. Molecular markers useful to discriminate *Coffea arabica* cultivars with high genetic similarity. **Euphytica**, v. 213, article number 75, 2017. DOI: 10.1007/s10681-017-1865-9.

TESSEMA, A.; ALAMEREW, S.; KUFA, T.; GAREDEW, W. Genetic diversity analysis for quality attributes of some promising *coffea arabica* germplasm collections in southwestern Ethiopia. **Journal of Biological Sciences**, v. 11, n. 3, p. 236-244, 2011. DOI: 10.3923/jbs.2011.236.244.

TRAN, H. T. M.; FURTADO, A.; VARGAS, C. A. C.; SMYTH, H.; SLADE LEE, L.; HENRY, R. SNP in the *coffea arabica* genome associated with coffee quality. **Tree Genetics & Genomes**, v. 14, article number 72, 2018. DOI: 10.1007/s11295-018-1282-9.

- VALENTE, M. S. F.; VIANA, J. M. S.; RESENDE, M. D. V. de; SILVA, F. F. e; LOPES, M. T. G. Seleção genômica para melhoramento vegetal com diferentes estruturas populacionais. **Pesquisa Agropecuária Brasileira**, v. 51, n. 11, p. 1857-1867, nov. 2016. DOI: 10.1590/s0100-204x2016001100008.
- VIDAL, R. O.; MONDEGO, J. M. C.; POT, D.; AMBRÓSIO, A. B.; ANDRADE, A. C.; PEREIRA, L. F. P.; COLOMBO, C. A.; VIEIRA, L. G. E.; CARAZZOLLE, M. F.; PEREIRA, G. A. G. A High-throughput data mining of single nucleotide polymorphisms in coffee species expressed sequence tags suggests differential homeologous gene expression in the allotetraploid *Coffea arabica*. **Plant Physiology**, v. 154, n. 3, p. 1053-1066, Nov. 2010. DOI: 10.1104/pp.110.162438.
- VOSSSEN, H. A. M. van der; WALYARO, D. J. Breeding for resistance to coffee berry disease in *Coffea arabica* L. II. Inheritance of the resistance. **Euphytica**, v. 29, n. 3, p. 777-791, Nov. 1980. DOI: 10.1007/BF00023225.
- WONG, C. K.; BERNARDO, R. Genomewide selection in oil palm: increasing selection gain per unit time and cost with small populations. **Theoretical and Applied Genetics**, v. 116, n. 6, p. 815-824, Apr. 2008. DOI: 10.1007/s00122-008-0715-5.
- WRAY, N. R.; YANG, J.; HAYES, B. J.; PRICE, A. L.; GODDARD, M. E.; VISSCHER, P. M. Pitfalls of predicting complex traits from SNPs. **Nature Reviews Genetics**, v. 14, n. 7, p. 507-515, July 2013. DOI: 10.1038/nrg3457.
- ZHANG, Z.; LIU, J.; DING, X.; BIJMA, P.; KONING, D.-J. de; ZHANG, Q. Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix. **PLoS ONE**, v. 5, n. 9, e12648, Sept. 2010. DOI: 10.1371/journal.pone.0012648.

**Embrapa**

---

**Café**



MINISTÉRIO DA  
AGRICULTURA, PECUÁRIA  
E ABASTECIMENTO

