

**UNIVERSIDADE ESTADUAL PAULISTA - UNESP
CÂMPUS DE JABOTICABAL**

**INFLUÊNCIA DA DISTÂNCIA GENÉTICA ENTRE
GENITORES NA RECUPERAÇÃO DO PARENTAL
RECORRENTE NA CONVERSÃO DE LINHAGENS DE
MILHO**

Elton Otacílio de Souza

Biólogo

2021

UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CÂMPUS DE JABOTICABAL

**INFLUÊNCIA DA DISTÂNCIA GENÉTICA ENTRE
GENITORES NA RECUPERAÇÃO DO PARENTAL
RECORRENTE NA CONVERSÃO DE LINHAGENS DE
MILHO**

Elton Otacílio de Souza

Orientador: Prof. Dr. Gustavo Vitti Môro

Coorientador: Dr. Ivan Schuster

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – UNESP, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Agronomia (Genética e Melhoramento de Plantas)

2021

S729i Souza, Elton Otacílio de
Influência da distância genética entre genitores na recuperação do parental
recorrente na conversão de linhagens de milho / Elton Otacílio de Souza. --
Jaboticabal, 2021
31 p.

Dissertação (mestrado profissional) - Universidade Estadual Paulista
(Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal
Orientadora: Gustavo Vitti Môro
Coorientadora: Ivan Schuster

1. Plantas Melhoramento genético. 2. Diversidade genética. 3.
Marcadores genéticos. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal



CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: INFLUÊNCIA DA DISTÂNCIA GENÉTICA ENTRE GENITORES NA RECUPERAÇÃO DO PARENTAL RECORRENTE NA CONVERSÃO DE LINHAGENS DE MILHO

AUTOR: ELTON OTACÍLIO DE SOUZA

ORIENTADOR: GUSTAVO VITTI MÔRO

COORIENTADOR: IVAN SCHUSTER

Aprovado como parte das exigências para obtenção do Título de Mestre em AGRONOMIA (GENÉTICA E MELHORAMENTO DE PLANTAS), pela Comissão Examinadora:

Prof. Dr. GUSTAVO VITTI MÔRO (Participação Virtual)
Departamento de Ciências da Produção Agrícola (Produção Vegetal) / FCAV / UNESP - Jaboticabal

Pesquisador Dr. RICARDO RODRIGUES (Participação Virtual)
Departamento de Pesquisa e Desenvolvimento - LONGPING HIGH-TECH / Cravinhos/SP

PI

Prof. Dr. RINALDO CESAR DE PAULA (Participação Virtual)
Departamento de Ciências da Produção Agrícola (Produção Vegetal) / FCAV / UNESP - Jaboticabal

PI

Jaboticabal, 29 de junho de 2021

DADOS CURRICULARES DO AUTOR

Elton Otacílio de Souza – filho de Juvenal Otacílio de Souza e Vera Lúcia da Silva e Silva, nasceu em 26 de setembro de 1976 na cidade de Guaíra, São Paulo.

Ingressou em 1992 no curso de Técnico em Agropecuária no Colégio José Bonifácio - Unesp de Jaboticabal, obtendo em dezembro de 1994 o título de Técnico em Agropecuária; em 2002 ingressou no curso de Biologia na UNIFRAN, obtendo em dezembro 2005 o título de Bacharel em Biologia e em 2006 ingressou no curso de Pós-Graduação em Especialização em Biologia Molecular e Biotecnologia na UNIFRAN, obtendo em dezembro de 2006 o título de Especialista em Biologia Molecular e Biotecnologia. Durante o curso de biologia realizou estágio no laboratório de Genética Molecular do Prof. Dr. Luís Eduardo Aranha Camargo na USP de Piracicaba.

Em fevereiro 1995, iniciou atividade de trainee na empresa Santa Eulália Comercio de Embriões no município de Pirassununga, estado de São Paulo, como técnico em pecuária; em agosto de 1995, iniciou atividade de bolsista do CNPq no projeto NAP na empresa Colorado Sementes em Guaíra, estado de São Paulo, como técnico de campo; em julho de 1997, iniciou seu trabalho na empresa Colorado Sementes, no município de Guaíra, estado de São Paulo, como técnico de pesquisa; em julho de 1998, iniciou seu trabalho na empresa Dow AgroSciences Ltda nos municípios de Guaíra e Jardinópolis, estado de São Paulo, como técnico de pesquisa em diversos projetos e funções; em janeiro de 2012 se tornou coordenador de atividades do laboratório de Biologia Molecular na Dow AgroSciences na cidade de Cravinhos, estado de São Paulo; em dezembro de 2018 iniciou seu trabalho na empresa LongPing High Tech Ltda como coordenador de atividades no Laboratório de Biologia Molecular, técnico responsável do projetos de campo de organismo geneticamente modificados regulados e/ou desregulados e criação e avaliação de bioensaios de insetos para as culturas de milho e soja da LPHT Brasil.

Em março de 2019, iniciou o curso de Mestrado em Agronomia (Genética e Melhoramento de Plantas) na Universidade Estadual Paulista “Júlio de Mesquita Filho” sob orientação do Prof. Dr. Gustavo Vitti Moro.

“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes”

Marthin Luther King

Dedico

A minha mãe, Vera Lúcia, por todo carinho, empenho e dedicação e aos meus amigos e orientadores profissionais: João Rosolem, Paulo Dion, Alexandre Takahashi e Ivan Schuster que sempre me incentivaram para o meu crescimento profissional e pessoal.

Ofereço

A Deus e a minha esposa, Edna e aos meus filhos, Luiz Olavo e Pedro Henrique

AGRADECIMENTOS

Ao Professor Dr. Gustavo Vitti M^oro, pela orienta^ço, pelos conselhos e por me ajudar nesse caminho cheio de obst^áculos chamado mestrado.

Ao Pesquisador Dr. Ivan Schuster, pela coorienta^ço, pelos incentivos, ensinamentos, conselhos e apoio para conquista do meu t^ítulo de mestrado um sonho que se torna realidade.

A Professora Dra. Luciana Rossini Pinto e ao pesquisador Dr. Gabriel Belfort Rodrigues o segundo autor desse trabalho e companheiro de todas as horas, que participaram da banca de qualifica^ço e que ajudaram a enriquecer o artigo de mestrado.

Ao Professor Dr. Rinaldo C^ésar de Paula, pelos conselhos e por toda ajuda e suporte durante o mestrado e ao Pesquisador Dr. Ricardo Rodrigues pelos ensinamentos em biotecnologia e por participarem da banca geral de defesa do curso de mestrado.

Aos Pesquisadores Jo^ão Batista Rosolem, MS. Paulo Eduardo Dion e Ms. Alexandre Takahashi, pelos conselhos e por toda ajuda e suporte durante toda minha trajet^ória profissional, gradua^ço, p^ós gradua^ço e curso de mestrado que me qualifica^ço para chegar a este momento t^ão honrado.

A minha fam^ília, esposa Edna, filhos Luiz Olavo e Pedro Henrique e m^ãe Vera Lucia que compartilharei das minhas dificuldades, anseios, e felicidades durante essa trajet^ória, sem esse apoio e essa base de compreens^ão, amor, carinho e cumplicidade n^ão seria poss^ível ter chegado at^é aqui.

INFLUÊNCIA DA DISTÂNCIA GENÉTICA ENTRE GENITORES NA RECUPERAÇÃO DO PARENTAL RECORRENTE NA CONVERSÃO DE LINHAGENS DE MILHO

RESUMO - Nas últimas duas décadas o uso de híbridos transgênicos agregou um valor considerável para o aumento da produção de milho. O método dos retrocruzamentos (RC) é o mais utilizado para a obtenção de linhagens e híbridos transgênicos em milho. Utilizando-se marcadores moleculares para seleção do genoma recorrente nas gerações de retrocruzamento é possível reduzir o tempo para a obtenção de linhagens convertidas. É possível que a utilização de parentais doadores mais similares aos parentais recorrentes possa reduzir ainda mais o tempo de conversão dessas linhagens. Este trabalho teve por objetivo avaliar a recuperação do genoma recorrente e a similaridade das plantas de RC com os parentais recorrentes em função da distância genética entre esses genitores em um programa de conversão de linhagens de milho. Foram avaliadas nove populações de retrocruzamento, com distâncias genéticas entre os parentais doadores e recorrentes variando de 0,238 a 0,499. As populações foram divididas em três grupos, um com distâncias genéticas de 0,238 a 0,286, outro de 0,307 a 0,335 e outro de 0,408 a 0,499. Nas gerações RC1, RC2 e RC3 foram usados marcadores moleculares para identificação das plantas com o maior percentual de recuperação do genoma recorrente e com maior similaridade com o parental recorrente. Na primeira geração de RC as diferenças na similaridade entre as plantas de RC e os parentais recorrentes foram proporcionais as distâncias genéticas iniciais, mas na terceira geração de RC todas as populações obtiveram plantas com similaridade de pelo menos 98% com os parentais recorrentes. Para as populações com menores distâncias genéticas entre os parentais, as linhagens elites já haviam sido recuperadas na segunda geração de RC. Quando o parâmetro usado foi recuperação do genoma recorrente, não foram observadas diferenças entre os grupos em RC3. Esse parâmetro não permitiria observar que nas populações cuja distância genética entre os parentais era menor as linhagens elites poderiam ser finalizadas na segunda geração de RC.

Palavras-chave: milho, retrocruzamento assistido por marcadores moleculares, marcadores SNP, diversidade genética

INFLUENCE OF GENETIC DISTANCE BETWEEN PARENTS IN THE RECOVERY OF THE RECURRENT PARENTAL IN THE CONVERSION OF MAIZE INBRED LINES

ABSTRACT - In the last two decades, the use of transgenic hybrids has added considerable value to the increase in maize production. The backcrossing method (BC) is the most used to obtain transgenic hybrid and inbred lines in maize. Using molecular markers for recurrent genome selection in backcross generations, it is possible to reduce the time to obtain converted inbred lines. It is possible that the use of donor parents more like those of parents may further reduce the conversion time of these inbred line. The objective of this work was to evaluate the recovery of the recurrent genome and the similarity of BC plants with recurrent parents as a function of the genetic distance between donors and recurrent parents in a maize inbred line conversion program. And compare the parameters of genetic similarity and recurrent genome recovery for decision management in BC programs assisted by molecular markers. Nine backcrossing populations were evaluated, with genetic distances between donors and recurrent parents ranging from 0.238 to 0.499. As the populations were divided into three groups, a genetic distance of 0.238 to 0.286, another from 0.307 to 0.335 and another from 0.408 to 0.499. In the BC1, BC2 and BC3 generations, molecular markers were used to identify the plants with the highest percentage of recurrent genome recovery and with greater similarity to the parental recurrent. In the first generation of BC as differences in similarity between BC plants and recurrent parent plants, the initial genetic distances were proportional, but in the third generation of BC all populations obtained plants with a similarity of 98% less with the recurrent of the parents. For populations with shorter genetic distances between parents, such as elite inbred line has already been recovered in the second generation of BC. When the parameter used was recurrent genome recovery, no differences were observed between the groups. This parameter does not allow us to observe that in populations whose genetic distances between parents were shorter, because elite lineages could be completed in the second generation of BC.

Keywords: maize, molecular marker-assisted backcrossing, SNP markers, genetic diversity

Lista de Abreviaturas

BC – Backcrossing (Retrocruzamento)

cM – Centimorgan

PD – Parental doador

PR – Parental recorrente

RC – Retrocruzamento

RC1 – Primeira geração de retrocruzamento

RC2 – Segunda geração de retrocruzamento

RC3 – Terceira geração de retrocruzamento

DNA – Ácido desoxirribonucleico

PCR – Reação em cadeia polimerase

A – Adenina

T – Timina

G – Guanina

C – Citosina

Lista de Tabelas

	Página
Tabela 1. Distância Genética e Similaridade entre os parentais doadores e recorrentes.....	8
Tabela 1. Número de marcadores e distribuição dos marcadores SNPs em cM em cada cromossomo.....	10
Tabela 3. Similaridade genética entre parentais doadores e recorrente Similaridades genéticas entre populações da primeira geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento.	21
Tabela 4. Similaridade genética entre as plantas selecionadas da primeira geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento.....	22
Tabela 5. Similaridade genética entre parentais doadores e recorrente similaridades genéticas entre populações da segunda geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento.....	23
Tabela 6. Similaridade genética entre as plantas selecionadas da segunda geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento.....	24
Tabela 7. Similaridade genética entre parentais doadores e recorrente similaridades genéticas entre populações da terceira geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento.....	25
Tabela 8. Similaridade genética entre as plantas selecionadas da terceira geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento.....	26

Lista de Figuras

	Página
Figura 1. Esquema do método de retrocruzamento. A barra vermelha representa o gene que se deseja transferir (7-Caixeta et al., 2015)	4

1. INTRODUÇÃO

No Brasil, dos 17,5 milhões de ha plantados com milho em 2019 (Conab, 2021), 16,3 milhões foram plantados com milho transgênico (ISAAA, 2019), o que significa 93% da área de milho sendo cultivada com milho geneticamente modificado.

O principal método utilizado para a conversão de linhagens e híbridos de milho para os eventos transgênicos comerciais é o método de Retrocruzamentos (RC). Este método tem por objetivo a introgressão de um ou poucos genes em uma variedade ou linhagem elite, e a subsequente recuperação do genoma do genitor recorrente. O processo geralmente é utilizado para corrigir genótipos-elites nas características em que são deficientes, por meio do cruzamento com genótipos portadores das características que se deseja introduzir, denominados genótipos doadores. Uma das limitações do método é o longo tempo necessário para se completar o processo, fazendo com que o genótipo-elite utilizado como genitor recorrente, muitas vezes, se torne obsoleto ao final do programa de retrocruzamento (Mesquita et al., 2005).

Uma das alternativas para reduzir o tempo de recuperação do genoma recorrente em uma linhagem ou variedade elite é a utilização de marcadores moleculares para a seleção das plantas com maior proporção de genoma recorrente em cada geração de retrocruzamento. A proporção de genoma recorrente recuperado em cada planta é variável em uma mesma geração de RC. O uso de marcadores moleculares permite identificar as plantas com maior proporção de genoma recorrente recuperado. Selecionando-se em cada geração de RC as plantas com maior proporção de genoma recorrente recuperado, pode-se reduzir o número de gerações de RC necessárias para a recuperação da linhagem elite contendo o gene transferido do genitor doador (Openshaw et al., 1994; Hospital, 2001; Bouchez et al., 2002).

A probabilidade de sucesso de um programa de retrocruzamento depende, em grande parte, da quantidade de germoplasma do genitor doador que precisa ser eliminada no processo de retrocruzamento. Assim, a utilização de genitores doadores mais similares aos genitores recorrentes poderia, em

teoria, facilitar a recuperação do genoma recorrente nesses programas de retrocruzamento. A similaridade genética entre os genitores doador e recorrente pode significar que alguns segmentos cromossômicos no doador em potencial podem ser idênticos por descendência ou pelo menos em estado com o do parental recorrente. Esses segmentos cromossômicos semelhantes, em essência, já estão convertidos para o genótipo do parental recorrente, podendo acelerar a recuperação do germoplasma do parental recorrente (Peng et al.; 2013).

Embora a vantagem da utilização de genitores doadores mais similares aos genitores recorrentes possa ser demonstrada teoricamente ou por simulação, não existem trabalhos usando dados reais que ilustrem essa vantagem. Por este motivo, o presente trabalho teve como objetivo avaliar a influência da distância genética entre os parentais doador e recorrente na recuperação do genótipo recorrente e na avaliação da similaridade genética das plantas de RC e o parental recorrente, em um programa de retrocruzamento assistido por marcadores moleculares.

2. REVISÃO DE LITERATURA

2.1 Aspecto econômico da cultura do milho

A cultura do milho (*Zea mays* L.) é uma das mais importantes do Brasil, sendo produzida em quase todo o território nacional (Engelsing et al., 2012). A Companhia Nacional de Abastecimento (Conab) estima que o país produzirá 268,3 milhões de toneladas de grãos na safra 2020/2021. O número apresenta um crescimento de 4,4%% (ou 11,4 milhões de toneladas), se comparado ao da safra anterior (Conab, 2021).

Os plantios de culturas biotecnológicas aumentaram aproximadamente 113 vezes desde 1996, com uma área acumulada de 2,5 bilhões de hectares, mostrando que a biotecnologia é a tecnologia agrícola de mais rápida adoção no mundo (ISAA, 2019). Nos países que mais utilizam cultivos transgênicos, particularmente EUA, Brasil, Argentina, Canadá e Índia, as taxas de adoção de grandes culturas estão em níveis próximos a 100%, indicando que os

agricultores favorecem essa tecnologia de cultura em relação às variedades convencionais (ISAA, 2019).

Nas últimas décadas, plantas transgênicas que expressam toxinas de *Bacillus thuringiensis* (*Bt*) vêm sendo amplamente utilizados para o controle genético de pragas no Brasil e no mundo. Após o ano de 2007, com a liberação comercial do primeiro evento *Bt* com resistência a lagarta no Brasil, a adoção desta tecnologia por parte dos agricultores foi muito rápida, neste período houve diversos lançamentos de híbridos transgênicos *Bt*, contendo desde um até seis genes *Bt* no mesmo híbrido. Este processo dinâmico de lançamentos de novos eventos transgênicos que expressam proteínas *Bt* é necessário pelo dinamismo da eficácia destas tecnologias (Ramos, 2018).

2.2 Método de retrocruzamento.

O método de retrocruzamento tem por objetivo a introgressão de um ou poucos genes em uma variedade ou linhagem elite, e a subsequente recuperação do genoma do genitor recorrente (Borém e Miranda, 2005). O processo geralmente é utilizado para corrigir genótipos-elites nas características em que são deficientes, por meio do cruzamento com genótipos portadores das características que se deseja introduzir, denominados genótipos doadores. O método de RC consiste no cruzamento da linhagem elite cujo defeito se deseja corrigir, denominada de genitor recorrente, com uma linhagem que contenha o gene que corrige este defeito, denominada de genitor doador. As plantas resultantes desse cruzamento e que contenham o gene que se deseja transferir são cruzadas novamente com o genitor recorrente e o processo de seleção das plantas que contém o gene de interesse e cruzamento com o genitor recorrente é repetido várias vezes. Em média, a cada geração de retrocruzamento recupera-se metade do genoma do genitor recorrente que ainda falta recuperar (Figura 1), de forma que a cada geração as progênies se tornam mais similares ao genitor recorrente (Guimarães et al., 2016).

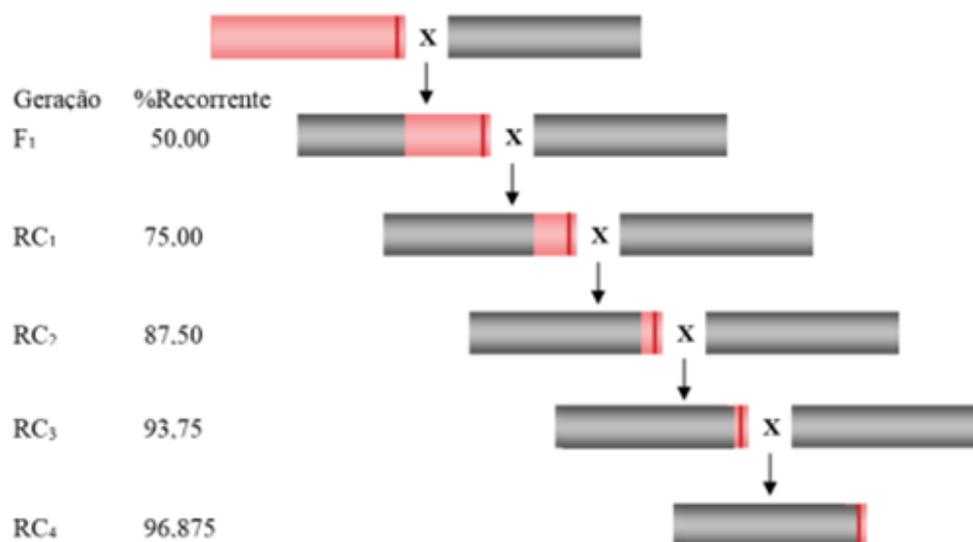


Figura 1. Esquema do método de retrocruzamento. A barra vermelha representa o gene que se deseja transferir (Caixeta et al., 2015)

Uma das limitações do método é o longo tempo necessário para se completar o processo, fazendo com que o genótipo-elite utilizado como genitor recorrente muitas vezes se torne obsoleto ao final do programa de retrocruzamento (Mesquita et al., 2005). Apesar disso este método é de extrema importância no processo de conversão de genótipos convencionais de milho no desenvolvimento de cultivares transgênicos (Mesquita et al., 2005).

Uma das alternativas para reduzir o tempo de recuperação do genoma recorrente em uma linhagem ou variedade elite é a utilização de marcadores moleculares para a seleção das plantas com maior proporção de genoma recorrente em cada geração de retrocruzamento (Openshaw et al., 1994; Hospital, 2001; Bouchez et al., 2002).

A probabilidade de sucesso de um programa de retrocruzamento depende, em grande parte, da quantidade de germoplasma do genitor doador que precisa ser eliminada no processo de retrocruzamento. Assim, a utilização de genitores doadores mais similares aos genitores recorrentes poderia facilitar a recuperação do genoma recorrente nesses programas de retrocruzamento. A similaridade genética entre os genitores doador e recorrente poderia fazer com que alguns segmentos cromossômicos no doador em potencial fossem idênticos por descendência ou pelo menos em estado com o do parental recorrente. Esses

segmentos cromossômicos semelhantes em essência já estão convertidos para o genótipo do parental recorrente, podendo acelerar a recuperação do germoplasma do parental recorrente (Peng et al., 2013).

2.3 Macho esterilidade.

O método tradicional para a produção de sementes híbridas de milho envolve o despendoamento manual das plantas femininas, deixando apenas pólen das plantas masculinas para fertilizar as plantas fêmeas. O despendoamento manual pode não ser efetivo, levando à disponibilidade de pólen de algumas plantas femininas e, assim, autopolinização e finalmente, uma redução no vigor híbrido. O despendoamento mecânico é mais rápido e econômico, mas a maioria das máquinas de despendoamento danificam as plantas durante o processo, causando uma redução linear no rendimento dos grãos para cada folha adicional removida com o pendão, (Wilhelm et al., 1995).

A macho esterilidade citoplasmática foi descrita em muitas espécies de plantas e está ligada a genes mitocondriais que codificam proteínas tóxicas. As proteínas citoplasmáticas masculinas estéreis (CMS) são tipicamente codificadas por um gene mitocondrial quimérico, (Chen e Liu, 2014).

No milho, o citoplasma masculino estéril pode ser dividido em três categorias de acordo com os padrões de restauração da fertilidade exibidos nos híbridos: Texas (T), USDA(S) e Charrua (C) (Beckett, 1971).

A esterilidade masculina citoplasmática tipo C (CMS-C) apresenta uma esterilidade masculina estável e afeta positivamente a produção de grãos, contribuindo para seu grande valor de aplicação na produção de sementes híbridas de milho (Stevanovic et al., 2016).

No geral, o complexo mecanismo de restauração da fertilidade CMS-C traz sérios desafios à aplicação do CMS-C na produção de sementes híbridas. O gene *Rf4* pode restaurar completamente a fertilidade da maioria das linhas de milho CMS-C e tem grande valor potencial na popularização e utilização híbrida do milho CMS-C. Muitos trabalhos têm sido realizados para *Rf4*, que foi previamente mapeado para o braço curto do cromossomo 8 (Tang et al., 2001).

2.4 Marcadores SNPs

Polimorfismo de nucleotídeo único (SNP) é uma variação de sequência de DNA que ocorre quando um único nucleotídeo (A, T, G ou C) difere entre os membros de uma espécie. O SNP é o sistema de marcadores mais abundante tanto em genomas animais como vegetais e emergiu como a nova geração de marcadores moleculares para várias aplicações (Arif et al., 2010).

Os marcadores SNP possuem status binário ou codominante, e são capazes de discriminar eficientemente entre alelos homocigotos e heterocigotos. Além disso, ao contrário dos microssatélites, seu poder não vem do número de alelos, mas do grande número de loci que podem ser avaliados (Turcheto-Zolet et al., 2017). Além disso, estes marcadores são favoráveis à automação de alto rendimento, permitindo uma genotipagem rápida e eficiente de um grande número de amostras (Tsuchihashi e Dracopoli, 2002).

O estudo de marcadores SNPs basicamente envolve duas etapas principais: a identificação (descoberta) dos SNPs no genoma da espécie de interesse e a genotipagem destes marcadores na população desta espécie ou no indivíduo de interesse para posterior análise. A diferença entre essas duas etapas é que na descoberta dos SNPs pode ser utilizado um número pequeno (representativo) de indivíduos da espécie estudada enquanto na genotipagem é utilizado um número maior de indivíduos, os quais representem uma ou mais populações, dependendo do objetivo do estudo. Tanto a identificação quanto a genotipagem de marcadores SNPs pode ser realizada utilizando métodos de pequena ou larga escala (Turcheto-Zolet et al., 2017).

A genotipagem pode envolver diferentes categorias de métodos e técnicas, destacando-se os métodos baseados em hibridização, como a hibridização alelo específico (Prince et al., 2001), hibridização de sondas (ex Sistema TaqMan por PCR em Tempo Real – McGuigan e Ralston 2002), Allele Specific PCR (Kasp – Cheng et al., 2017) e hibridização em arranjos multiplex (SNP array) (Hehir-Kwa et al., 2007); métodos de ligação de oligonucleotídeo baseado em PCR (Podder et al., 2008) e métodos baseados em NGS (Next Generation Sequencing, Torkamaneh et al., 2016).

A Genotipagem por Sequenciamento (Genotyping-by-Sequencing - GBS) é um método de genotipagem de baixo custo e alto rendimento que depende de enzimas de restrição para reduzir a complexidade do genoma. A GBS está sendo amplamente utilizada para várias aplicações genéticas e de melhoramento. Avanços significativos na tecnologia de sequenciamento têm ocorridos nos últimos anos e levaram a uma melhor compreensão do genoma do milho. GBS é um método para genotipar muitas amostras e fornece mais SNPs do que a genotipagem utilizando de SNPs arrays. (Wang et al., 2020).

Uma aplicação mais prática do uso da genotipagem por sequenciamento na seleção assistida por marcadores moleculares é a genotipagem por sequenciamento de regiões alvo (Target Genotyping by Sequence – tGBS). Diferentemente da GBS, que se utiliza enzimas de restrição para a redução da complexidade do genoma antes da etapa de ligação dos adaptadores, a tGBS utiliza PCR multiplex para enriquecer as regiões alvo do sequenciamento. Nesse caso os SNPs são pré-definidos através do desenho de oligonucleotídeos para amplificação de regiões de interesse e o sequenciamento é alvo-direcionado. Essa estratégia oferece ótimo custo benefício por possibilitar a genotipagem de grande quantidade de amostras e marcadores SNPs simultaneamente, além de ser relativamente simples sua customização e automatização (He et al., 2014), possibilitando o seu uso para aumentar a densidade da genotipagem para seleção do genoma recorrente em programas de retrocruzamento.

3. MATERIAL E MÉTODOS

Para a realização do estudo foram avaliadas nove populações de RC conduzidas na unidade de pesquisa da LongPing High-Tech em Cravinhos-SP. Para evitar conflitos de interesse, o gene alvo foi um gene nativo do milho (*Rf4*), mas os resultados são diretamente aplicáveis a programas de retrocruzamentos para eventos transgênicos.

Todas as populações foram obtidas com o mesmo parental doador (PD1) para transferência do gene de restauração de fertilidade masculina *Rf4*. Os nove parentais recorrentes foram denominados PR1, PR2, PR3, PR4, PR5, PR6, PR7, PR8 e PR9. Os parentais recorrentes e as respectivas populações de RC foram

classificadas em três grupos: Grupo1: PR similares ao PD; Grupo2: PR com similaridade/distância intermediária ao PD; e Grupo3: PR geneticamente mais distantes do PD.

As distâncias genéticas entre os parentais doador e recorrentes variaram de 0,238 a 0,499, formando 3 grupos distintos com três linhagens em cada grupo de avaliação sendo o grupo 1 com a menor distância entre os parentais, o grupo 2 com distâncias intermediárias e o grupo 3 com maior distância entre os parentais doador e recorrentes. (Tabela 1).

Tabela 1. Distância Genética e Similaridade entre os parentais doadores e recorrentes

Recorrente	Doador	Distância Genética	Similaridade entre PD-PR	Classe de Distâncias	Grupo
PR1	PD1	0,238	76,2%	Similares	1
PR2	PD1	0,256	74,4%	Similares	1
PR3	PD1	0,286	71,4%	Similares	1
PR4	PD1	0,335	66,5%	Intermediário	2
PR5	PD1	0,307	69,3%	Intermediário	2
PR6	PD1	0,309	69,1%	Intermediário	2
PR7	PD1	0,452	54,8%	Distante	3
PR8	PD1	0,408	59,2%	Distante	3
PR9	PD1	0,499	50,1%	Distante	3

Inicialmente foram realizados os cruzamentos entre o parental doador com todos os parentais recorrentes. As plantas da geração F1 foram cruzadas novamente com as respectivas linhagens recorrentes. Nas gerações RC1 e RC2 as plantas de cada população foram selecionadas com base na análise presença do gene *Rf4* e na recuperação do genoma recorrente. As plantas que continham o gene *Rf4* e que possuíam a maior similaridade genética com o genitor recorrente foram novamente cruzadas com as respectivas linhagens recorrentes.

Em todas as gerações de retrocruzamento foram selecionadas as plantas que possuíam o gene nativo *Rf4* do programa de retrocruzamento.

Para extração de DNA foi coletado um disco foliar de 5 mm de diâmetro de cada planta, O DNA dos discos foliares foi extraído utilizando-se o kit Fast ID

Genomic (Genetic ID – Fairfield, Iowa, USA), seguindo as instruções do fabricante. Inicialmente as amostras de DNA foram utilizadas para a seleção do gene *Rf4*, utilizando-se primers e sondas específicos e o método TaqMan® (Holland et al., 1991). As plantas contendo o gene *Rf4* foram genotipadas com um painel de 1065 marcadores SNPs, utilizando-se o método de genotipagem por sequenciamento de regiões alvo (tGBS – target Genotyping by Sequence). Para esta genotipagem, foi utilizada a plataforma Ion GeneStudio S5 System (Thermo Fisher – Waltham, Massachusetts, USA), seguindo as instruções do fabricante. Também foram genotipadas amostras de DNA dos parentais doadores e recorrentes de cada população.

Uma vez que no painel de genotipagem também havia marcadores monomórficos entre os parentais doadores e recorrentes, foram avaliados dois parâmetros de comparação entre as plantas das gerações de RC e os genitores recorrentes: a Similaridade Genética e a Recuperação do Genoma Recorrente.

A similaridade genética (S) foi estimada usando-se a expressão:

$$S = \frac{\text{Número de SNP homozigotos iguais ao genitor recorrente}}{\text{Número total de SNP}}$$

As distâncias genéticas (D) foram estimadas por $D = 1 - S$. A recuperação do genoma recorrente de cada planta de RC foi estimada da mesma forma que a Similaridade Genética, porém usando apenas os marcadores SNP polimórficos entre os genitores doador e recorrente. Neste caso, considera-se que os genitores recorrente e doador tem similaridade zero. A similaridade genética entre as plantas RC1, RC2 e RC3 e o genitor recorrente, considerando-se apenas marcadores polimórficos, representa a quantidade de genoma recorrente que era divergente entre os parentais doador e recorrente, que já foi recuperado nestas plantas.

Os valores de distância genética foram expressos em decimais, que podem variar de zero a 1. A similaridade e a recuperação do genoma recorrente foram expressas em porcentagem, multiplicando-se o valor estimado por 100.

As cinco plantas com maior similaridade com o seu genitor recorrente foram selecionadas para produzirem a geração seguinte de retrocruzamento.

Para avaliar a similaridade genética e a recuperação do genoma recorrente entre os três grupos, em cada geração de retrocruzamento foram estimadas as médias, desvio padrão e amplitude das similaridades genéticas das plantas RC1, RC2 e RC3 e os genitores recorrentes em cada população.

Em todas as gerações de RC foram utilizados 1065 marcadores moleculares SNPs para selecionar as plantas com maior recuperação do genoma recorrente para compor as próximas gerações de RC (Tabela 2).

Tabela 2. Número de marcadores e distribuição dos marcadores SNPs em cM em cada cromossomo.

Cromossomo	Número de Marcadores	cM
1	262	230,56
2	134	182,10
3	130	188,34
4	104	166,04
5	97	217,86
6	69	160,42
7	75	143,55
8	59	155,46
9	72	143,71
10	63	133,70
Total	1065	1727,25

As análises de similaridade genética foram realizadas com auxílio do programa Tassel 5 (Bradbury et al., 2007). Para comparar os três grupos de populações, as médias de similaridade genética e de recuperação do genoma recorrente foram utilizadas para a análise de variância, em que os grupos foram os tratamentos. Utilizou-se o modelo inteiramente casualizados, com três repetições, que foram as três populações de cada grupo. As médias foram comparadas pelo teste de Tukey com significância de 5%. Também foram avaliadas as correlações de Pearson entre a similaridade genética dos parentais recorrente e doador com as similaridades genéticas e recuperação do genoma recorrente nas diferentes populações de RC. As significâncias das correlações

foram avaliadas pelo teste de Mantel, com 5000 simulações e teste de t. As análises de variância, teste de média e correlações foram realizadas com auxílio do programa Genes (Cruz, 2014).

4. RESULTADOS E DISCUSSÃO

4.1 Recuperação do genoma recorrente

A recuperação do genoma recorrente significa a quantidade do genoma que era divergente entre os genitores doador e recorrente que se tornou similar ao genoma do genitor recorrente em cada geração de RC (Peng et. al., 2013), expressa em porcentagem. Neste caso, apenas a parte divergente do genoma é considerada na avaliação de similaridade, o que significaria dizer que a similaridade original entre os genitores doador e recorrente era zero.

Após a primeira geração de retrocruzamento, a média da recuperação do genoma recorrente das plantas de todas as populações RC1 foi de 78,29% (Tabela 3). No grupo 1 a recuperação do genoma recorrente foi em média de 78,49%, no grupo 2 de 77,36% e no grupo 3 de 79,03%. As médias obtidas foram ligeiramente superiores à média esperada de recuperação do genoma recorrente na geração RC1, que é de 75% (Borém e Miranda, 2005), sendo que a diferença na recuperação do genoma recorrente entre os grupos 2 e 3 foi significativa, enquanto que em relação ao grupo 1, nenhum dos outros dois grupos apresentou recuperação do genoma recorrente significativamente diferente (Tabela 3).

A correlação entre a média da recuperação do genoma recorrente das plantas da geração RC1 com as distâncias genéticas entre os genitores em cada população foi de 0,50. Embora este valor de correlação indique uma tendência de maior recuperação do genoma recorrente nas populações cujos genitores doador e recorrente eram mais divergentes, essa correlação não foi significativa (Tabela 3).

Em todas as populações de RC é possível observar uma grande variabilidade na recuperação do genoma recorrente entre as plantas em acordo com base nos marcadores moleculares em acordo com Wang et al. (2007). A amplitude na variação da recuperação do genoma recorrente nas plantas RC1

variou de 16,6 pontos percentuais do genoma na população PR7xPD1 até 26,58 pontos percentuais do genoma na população PR2xPD1 (Tabela 3). Uma vez que há esta variabilidade na proporção do genoma recorrente dentro de cada população, é possível selecionar as plantas com a maior proporção de recuperação do genoma recorrente para a próxima geração de RC (Peng et al., 2013). Usando-se plantas com maior proporção do genoma recorrente recuperado para gerar a próxima geração de RC, pode-se acelerar a recuperação desejada do genoma recorrente nos programas de RC, e finalizar os projetos com menor número de gerações de RC de acordo com Pinto et al., 2018.

Cinco plantas de cada população RC1 com maior recuperação do genoma do parental recorrente foram selecionadas para a próxima geração de retrocruzamento. Considerando apenas as plantas selecionadas, a média de recuperação do genoma recorrente entre todas as populações foi de 86,73% (Tabela 4). No grupo 1 a média foi de 87,29%, no grupo 2 de 85,38% e no grupo 3 a média foi de 87,51% (Tabela 4). As médias de recuperação do genoma recorrente nas plantas selecionadas do grupo 1 e do grupo 3 diferiram significativamente da média da recuperação do genoma recorrente das plantas selecionadas do grupo 2. Mas essa diferença não deve ser devida às distâncias genéticas entre os genitores, pois a correlação entre a distância genética e a recuperação do genoma recorrente não foi significativa (Tabela 4). O número de plantas avaliadas por marcadores do grupo 2 foi 84, que é 5,6% menor do que o grupo 1 e 8,6% menor do que o grupo 3. Como o número de plantas selecionadas foi fixo entre os grupos, houve diferentes intensidades de seleção. Sob intensidade de seleção fixa, observa-se que nos doadores mais similares com os parentais recorrentes há maior probabilidade de obtenção de maior recuperação (Peng et al., 2013).

Considerando-se a recuperação média do genoma recorrente das plantas selecionadas, a correlação entre a recuperação do genoma recorrente e as distâncias genéticas entre os genitores foi de 0,20, e não foi significativa (Tabela 4). Na geração RC1, a recuperação do genoma recorrente não tem relação com as distâncias genéticas entre os genitores, tanto na média geral das populações quanto na média das plantas selecionadas.

Em um programa de retrocruzamento sem auxílio de marcadores moleculares, é esperado que na geração RC2 a média de recuperação do genoma recorrente seja de 87,5% (Borém e Miranda, 2005). Selecionando-se as cinco plantas com a maior proporção de genoma recorrente em cada população, com uso de marcadores moleculares, a proporção de recuperação do genoma recorrente esperada para geração RC2 foi obtida na geração RC1. O monitoramento da recuperação do genoma do recorrente via marcador viabiliza a redução do número de retrocruzamentos necessários (Mesquita, 2002).

Na geração RC2 a média da recuperação do genoma recorrente entre todas as populações foi de 92,57% (Tabela 5). No grupo 1 a média foi de 93,57%, no grupo 2 foi de 91,86% e no grupo 3 de 92,27% (Tabela 5). As diferenças entre as médias da recuperação do genoma recorrente entre os três grupos na geração RC2 não foram significativas. A correlação entre a recuperação média do genoma recorrente na geração RC2 e as distâncias genéticas entre os parentais foi -0,34507 e não foi significativa (Tabela 5).

Uma vez que na geração RC1 foi realizada seleção das plantas com maior proporção do genoma recorrente recuperado, a proporção média do genoma recorrente recuperado na geração RC2 é equivalente ao esperado para a geração RC3 em um programa sem a seleção por marcadores moleculares, que é de 93,5%. Isso era esperado, uma vez que as plantas usadas para o cruzamento na geração RC1, após a seleção por marcadores moleculares, já tinham proporção de genoma recorrente equivalente ao de uma geração RC2 sem o uso de marcadores moleculares.

A média de recuperação do genoma recorrente nas cinco plantas selecionadas de todas as populações foi de 96,18% (Tabela 6). No grupo 1 a média de recuperação nas plantas selecionadas foi de 96,56%, no grupo 2 foi de 96,00% e no grupo 3 foi de 95,99%. As diferenças entre estas médias também não foi significativa, e a correlação destas médias com as distâncias genéticas entre os genitores foi de -0,31262, também não significativa (Tabela 6).

Após a seleção por marcadores moleculares, as plantas a serem usadas para a próxima geração de RC apresentaram média de recuperação do genoma recorrente equivalente a uma população RC4 sem o uso de marcadores moleculares, que é de 96,875%.

Na geração RC3 a média da recuperação do genoma recorrentes nas plantas de todas as populações foi de 96,97% (Tabela 7). No grupo 1 a média foi de 96,90%, no grupo 2 foi de 96,89% e no grupo 3 de 97,14% (Tabela 7). As diferenças entre estas médias não foram significativas, e a correlação entre estas médias e as distâncias genéticas entre os genitores (0,24366) também não foi significativa (Tabela 7).

Entre as cinco plantas selecionadas e no conjunto de todas as populações a média de recuperação do genoma recorrente foi de 98,92% (Tabela 8). No grupo 1 a média de 98,69%, no grupo 2 foi de 98,77% e no grupo 3 foi de 99,29%. Estas médias também não diferiram entre si.

Na geração RC3, apenas duas populações não tiveram plantas com mais de 99% de recuperação do genoma recorrente. A população PR1xPD1 teve recuperação máxima de 98,28%, e a população PD6xPR1 teve recuperação máxima de 98,25%. Todas as demais populações tiveram plantas com mais de 99% de recuperação do genoma recorrente na geração RC3. Em programas de RC que não utilizem marcadores moleculares, esta proporção de recuperação do genoma recorrente só é esperada na geração RC6 (Borém e Miranda, 2005). Assim, demonstra-se com esses resultados a importância do uso de marcadores moleculares para estimar a recuperação do genoma no programa de retrocruzamentos.

4.2 Similaridade genética nas gerações RC1, RC2 e RC3

Após a primeira geração de retrocruzamento, a média da similaridade genética das plantas de todas as populações RC1 com os genitores recorrentes foi de 92,15% (Tabela 3). A média da similaridade genética das plantas RC1 do grupo 1 com os genitores recorrentes foi de 94,01%. No grupo 2 a média da similaridade entre as plantas RC1 e os parentais recorrentes foi de 92,34%, e no grupo 3 a média da similaridade com os parentais recorrentes foi de 90,10% (Tabela 3). As diferenças entre as médias de similaridade genética em cada grupo foram estatisticamente significativas, demonstrando o efeito da distância genética entre os genitores e a similaridade genética das plantas RC1 com o genitor recorrente (Tabela 3).

As similaridades genéticas das plantas da geração RC1 com os genitores recorrentes foram maiores nas populações cujas distâncias genéticas entre os parentais doadores e recorrentes eram menores. A correlação entre a similaridade genética média de cada população de RC e a distância genética entre os parentais foi -0,97, significativa a 5% de probabilidade, na geração RC1 (Tabela 3). A maior média de similaridade com o parental recorrente foi observada no cruzamento PR1xPD1 (94,74%) e a menor média de similaridade foi no cruzamento PR9xPD1 (89,86%) (Tabela 3).

Também foi observada maior variabilidade entre as plantas das populações RC1 com maior distância genética entre os parentais. Populações de retrocruzamento derivadas de parentais mais divergentes apresentaram maior desvio padrão das similaridades e maior amplitude das similaridades genéticas das plantas de cada população com o genitor recorrente (Tabela 3).

Considerando apenas as cinco plantas selecionadas em todas as populações, a média da similaridade genética com os genitores recorrentes foi de 95,11% (Tabela 4). Analisando cada grupo separadamente as médias de similaridade com o genitor recorrente foram 96,42% no grupo 1, 94,98% no grupo 2 e 93,92% no grupo 3. As médias de similaridade do grupo 1, considerando apenas as plantas selecionadas, é estatisticamente diferente das médias dos grupos 2 e 3, demonstrando a eficácia na recuperação do genoma do parental recorrente com o emprego de doadores mais próximos.

Assim como na média de todas as plantas das populações RC1, a maior e menor similaridade das plantas selecionadas com os genitores recorrentes também foi observada nas populações dos cruzamentos PR1xPD1 (97,24%) e PR9xPD1 (92,75%) (Tabela 4).

A correlação entre as similaridades médias das plantas selecionadas em cada população com as distâncias genéticas entre os genitores de cada população de retrocruzamento foi de -0,94003, significativa a 5% de probabilidade. Essa correlação elevada indica que, embora haja maior amplitude de valores de similaridade genética entre as plantas derivadas de cruzamentos mais divergentes, essa amplitude não é suficiente para compensar a quantidade de genoma que precisa ser recuperado nestas populações mais divergentes, na geração RC1 (Tabela 4).

Na geração RC2 a média da similaridade genética das plantas de todas as populações com os genitores recorrentes foi de 97,15% (Tabela 5). No grupo 1 a similaridade média das plantas RC2 com os genitores recorrentes foi de 98,22%. No grupo 2 foi de 97,08% e no grupo 3 foi de 96,14% (Tabela 5). Na geração RC2, as médias de similaridade são estatisticamente diferentes apenas entre o grupo 1 e o grupo 3. Após uma geração de retrocruzamento assistido por marcadores moleculares, as similaridades do grupo 3 já se igualaram as similaridades do grupo 2, e as similaridades do grupo 2 também se igualaram as similaridades do grupo 1. Isso demonstra a eficácia do uso de marcadores moleculares já na geração RC1 para seleção das plantas com maior recuperação do genoma do parental recorrente.

Também na geração RC2 as similaridades genéticas com os genitores recorrentes foram maiores nas populações cujas distâncias genéticas entre os parentais doadores e recorrentes eram menores, com correlação de -0,900358 (Tabela 5). A maior média de similaridade genética na geração RC2 também foi observada na população do cruzamento PR1xPD1, com média de 98,64%. Também na geração RC2, a menor média de similaridade foi observada no cruzamento PR9xPD1, com média de 95,25% (Tabela 5).

Entre as cinco plantas de cada população RC2 mais similares ao parental recorrente, a média de similaridade foi 98,40%, considerando-se a média de todas as populações (Tabela 6). No grupo 1 a média de similaridade das cinco plantas selecionadas foi 98,95%, no grupo 2 foi 98,48% e no grupo 3 foi 97,77%. Na geração RC1, entre as plantas selecionadas, as similaridades nos grupos 2 e 3 já haviam sido estatisticamente iguais. Na geração RC2, além dos grupos 2 e 3 não diferirem entre si, as similaridades do grupo 2 também não diferiram das similaridades do grupo 1.

Entre as plantas selecionadas, a maior e menor similaridade continuou sendo nas populações dos cruzamentos PR1xPD1 (99,18%) e PR9xPD1 (97,37%) (Tabela 6). Mais uma vez, mesmo considerando apenas as cinco plantas mais similares aos parentais recorrentes em cada população, a correlação entre a média de similaridade e as distâncias genéticas originais entre os parentais de cada população (-0,91702) foi elevada (Tabela 6).

Considerando a similaridade com o parental recorrente parâmetro de seleção quando avaliamos todas as plantas de todas as populações o grupo 1 com média acima de 98,0% poderia ser finalizado, porém quando utilizamos a média de similaridade para as cinco plantas selecionadas para similaridade com o parental recorrente tanto o grupo 1 quanto o grupo 2 poderiam ser finalizados ambos com mais de 98,0% de similaridade. Isso demonstra que a influência da distância genética inicial tem efeito no número de gerações de retrocruzamentos necessários para a finalização no processo de introgressão.

Na geração RC3 a média da similaridade genética das plantas de todas as populações RC3 com os genitores recorrentes foi de 98,61% (Tabela 7). Entre as populações do grupo 1 a média de similaridade foi de 98,84%, no grupo 2 foi de 98,60% e no grupo 3 foi de 98,39% (Tabela 7), sendo que as diferenças entre estas médias não foram significativas.

Na geração RC3 a correlação entre as similaridades médias em cada população com o seu genitor recorrente e as distâncias genéticas originais entre os genitores foi de menor magnitude (-0,45451) e não significativa (Tabela 7). A maior média de similaridade foi no cruzamento PR3xPD1 (99,22%) e a menor média foi nos cruzamentos PR7xPD1 e PR8xPD1 (98,35%) (Tabela 7) e a diferença entre as médias de similaridade é de apenas 0,87 pontos percentuais. Após duas gerações de retrocruzamento assistido por marcadores moleculares, já não há mais influência significativa das distâncias genéticas entre os genitores das populações de retrocruzamento e as similaridades das plantas de cada população com o seu genitor recorrente.

Considerando-se as cinco plantas selecionadas em todas as populações, a média de similaridade com os genitores recorrentes foi 99,44% (Tabela 8). No grupo 1 a média de similaridade entre as cinco plantas mais similares foi 99,40%, no grupo 2 foi 99,37% e no grupo 3 a média foi 99,55%. A diferença entre a maior e a menor média de similaridade entre as plantas selecionadas de cada grupo foi de 0,18%, e não há diferença estatística entre estas médias.

A população do cruzamento PR9xPD1, que tinha as plantas com menor similaridade com os genitores recorrentes nas duas primeiras gerações de RC, na geração RC3 foi a que apresentou a maior similaridade (99,78%). A

população com menor média de similaridade após três gerações de RC foi PR6xPD1 (98,99%), do grupo 2 (Tabela 8).

Considerando-se a similaridade genética entre os genitores doador e recorrentes, os três grupos de populações de retrocruzamento apresentavam diferenças entre si. Após uma geração de retrocruzamento, sem seleção, as médias de similaridade genética continuaram diferentes entre os grupos. Já na primeira geração de retrocruzamento, após a seleção das plantas mais similares aos genitores recorrentes em cada população, os grupos 2 e 3 já não apresentaram diferenças significativas nas médias de similaridades com os genitores recorrentes. Após três gerações de retrocruzamento e seleção com marcadores moleculares para a recuperação do genoma recorrente, todas as populações apresentaram média de recuperação do genoma recorrente próximo de 99%, não havendo mais nenhuma influência das distâncias genéticas originais na recuperação do genoma.

Na geração RC2, a média de similaridade das plantas selecionadas do grupo 1 com os genitores recorrentes foi próxima de 99%, sendo que em todas as populações foram obtidas plantas com mais de 99,22% de recuperação do genoma recorrente. Uma vez que já há uma grande similaridade entre as plantas selecionadas das populações de RC com os genitores recorrentes, não há mais muito ganho em similaridade genética da geração RC2 para a geração RC3. Nas três populações desse grupo todas poderiam ser encerradas a introgressão na geração RC2. Nos outros dois grupos, a terceira geração de RC aumentou a similaridade com os genitores recorrentes, e era necessário ter sido realizada

Na avaliação da recuperação do genoma recorrente, foi observado que após três gerações de RC todas as populações apresentaram recuperação do genoma recorrente próximo de 99%, independentemente da distância genética entre os genitores doador e recorrente. Até mesmo na geração RC2 na maior parte das populações houve plantas com recuperação do genoma recorrente acima de 97% (Tabela 6).

Na avaliação da recuperação do genoma recorrente considera-se apenas as regiões do genoma que contrastam entre os genomas dos genitores usados no cruzamento inicial. No entanto, a avaliação da recuperação do genoma recorrente sem considerar a similaridade genética com base na avaliação de

todo o genoma, e não apenas nas regiões contrastantes pode levar a erros de interpretação, e o resultado do programa de RC pode não ser o esperado.

Considerando o exemplo de duas populações, em que a primeira a distância genética entre os parentais seja de 0,2 e na segunda a distância genética seja de 0,5, e considerando-se o genoma do milho, que tem em torno de 1.700 cM, os genitores da primeira população divergem em 340 cM, enquanto que os genitores da segunda população divergem em 850 cM. Quando a primeira população tiver recuperado 98% do genoma recorrente, terá recuperado 332,2 cM, restando ainda 6,8 cM do genitor doador nas plantas convertidas. Quando a segunda população tiver recuperado 98% do genoma recorrente, terá recuperado 833 cM, restando ainda 17 cM do genitor doador nestas plantas, ou 2,5 vezes mais germoplasma do genitor doador do que na primeira população.

Pelos resultados obtidos, a avaliação da similaridade genética entre as plantas das populações de RC com o genitor recorrente oferece informação mais precisa para o gerenciamento dos programas de RC, e a definição do momento de encerrar os ciclos de RC e finalizar as linhagens. Para isso, é necessário utilizar um número de marcadores que amostram bem todo o genoma da espécie, para que as similaridades genéticas sejam estimadas com elevada acurácia. Diante disso, os marcadores representam uma ferramenta importante para detectar a recuperação do genoma do parental recorrente nas gerações de retrocruzamento (Bellon et.al., 2014).

A utilização de genitores doadores geneticamente mais próximos dos genitores recorrentes pode reduzir o número de gerações de RC necessárias para recuperar a quantidade de genoma recorrente desejada, como é o caso das populações do grupo 1 deste estudo. No entanto, com três gerações de RC, e uso de marcadores moleculares em cada geração para selecionar as plantas com maior proporção de genoma recorrente recuperado, o efeito da distância genética entre os parentais doadores e recorrente é anulada, e todas as populações conseguem chegar à similaridade genéticas próximas ou acima de 99%. Importante comentar que os marcadores representam ferramenta importante para estimar a recuperação do genoma do parental recorrente e que ainda pode haver diferença entre a linhagem convertida e sua versão original. Diante disso, é importante manter procedimento de finalização via técnicas

convencionais de melhoramento antes de escolher a melhor versão para cada linhagem (Mumm e Walters, 2001; Mumm, 2013).

Tabela 3. Similaridade genética entre parentais doadores e recorrente similaridades genéticas entre populações da primeira geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento e Correlação de Pearson / Teste de Mantel, médias das gerações RC1: Dist= Distância Genética, Sim= Similaridade, Recup= Recuperação do genoma recorrente de todas as plantas.

População	RCs	Distância Genética	No. Plants	Similaridade com Recorrente			Recuperação do Genoma Recorrente		
				Média	Desvio Padrão	Amplitude	Média	Desvio Padrão	Amplitude
PR1XPD1	RC1	0,238	88	94,74%	1,04%	92,90% - 97,24%	78,93%	4,20%	71,30% - 89,77%
PR2XPD1	RC1	0,256	91	93,74%	1,22%	89,96% - 97,23%	77,57%	4,49%	63,92% - 90,50%
PR3XPD1	RC1	0,286	89	93,55%	1,17%	89,49% - 96,41%	78,96%	3,89%	64,92% - 88,08%
Média Grupo1		0,260	89	94,01% a			78,49% ab		
PR4XPD1	RC1	0,335	84	91,94%	1,74%	84,16% - 95,70%	77,12%	4,45%	66,67% - 87,41%
PR5XPD1	RC1	0,307	89	92,60%	1,44%	89,11% - 96,11%	77,55%	4,45%	67,16% - 88,33%
PR6XPD1	RC1	0,309	80	92,49%	1,27%	89,16% - 95,59%	77,41%	3,91%	67,94% - 87,16%
Média Grupo2		0,317	84	92,34% b			77,36% b		
PR7XPD1	RC1	0,452	86	90,00%	2,08%	86,71% - 94,10%	79,07%	4,42%	70,77% - 87,37%
PR8XPD1	RC1	0,408	95	90,44%	2,03%	86,20% - 95,32%	78,28%	4,81%	66,35% - 89,30%
PR9XPD1	RC1	0,499	95	89,86%	1,84%	84,78% - 94,99%	79,74%	3,71%	69,33% - 90,17%
Média Grupo3		0,453	92	90,10% c			79,03% a		
Média Geral		0,343	89	92,15%			78,29%		
Correlação Dist x Sim				-0,97038**++					
Correlação Dist x Recup				0,50354					

Médias seguidas por letras iguais nas colunas não diferem pelo teste de Tukey a 5% de probabilidade

**, * - Significativo a 1 e 5% de probabilidade pelo teste t

++, + - Significativo a 1 e 5% de probabilidade pelo teste de Mantel baseado em 5000 simulações

Tabela 4. Similaridade genética entre as plantas selecionadas da primeira geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento e Correlação de Pearson / Teste de Mantel, médias das gerações RC1: Dist= Distância Genética, MedSel= Média das similaridades das cinco plantas selecionadas e RecSel= Média da recuperação do genomarecorrente das cinco plantas selecionas.

População	RCs	Distância Genética	No. Plants	Similaridade com Recorrente			Recuperação do Genoma Recorrente		
				Média	Desvio Padrão	Amplitude	Média	Desvio Padrão	Amplitude
PR1XPD1	RC1	0,238	5	97,00%	0,19%	96,73% - 97,24%	88,09%	1,00%	87,09% - 89,77%
PR2XPD1	RC1	0,256	5	96,34%	0,56%	95,89% - 97,23%	87,05%	2,04%	85,62% - 90,50%
PR3XPD1	RC1	0,286	5	95,91%	0,46%	95,26% - 96,41%	86,73%	1,38%	84,78% - 88,08%
Média Grupo1		0,260		96,42% a			87,29% a		
PR4XPD1	RC1	0,335	5	94,64%	0,64%	94,06% - 95,70%	84,92%	1,42%	83,91% - 87,41%
PR5XPD1	RC1	0,307	5	95,26%	0,54%	94,71% - 96,11%	85,70%	1,65%	84,31% - 88,33%
PR6XPD1	RC1	0,309	5	95,05%	0,40%	94,57% - 95,59%	85,51%	1,15%	84,04% - 87,16%
Média Grupo2		0,317		94,98% b			85,38% b		
PR7XPD1	RC1	0,452	5	93,59%	0,34%	93,23% - 94,10%	86,65%	0,50%	86,01% - 87,37%
PR8XPD1	RC1	0,408	5	94,67%	0,47%	94,26% - 95,32%	88,04%	0,88%	87,08% - 89,30%
PR9XPD1	RC1	0,499	5	93,50%	0,94%	92,75% - 94,99%	87,85%	1,93%	85,56% - 90,17%
Média Grupo3		0,453		93,92% b			87,51% a		
Média Geral		0,343		95,11%			86,73%		
Correlação Dist x MedSel				-0,94003**++					
Correlação Dist x RecSel				0,20791					

Médias seguidas por letras iguais nas colunas não diferem pelo teste de Tukey a 5% de probabilidade

** , * - Significativo a 1 e 5% de probabilidade pelo teste t

++, + - Significativo a 1 e 5% de probabilidade pelo teste de Mantel baseado em 5000 simulações

Tabela 5. Similaridade genética entre parentais doadores e recorrente similaridades genéticas entre populações da segunda geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento e Correlação de Pearson / Teste de Mantel, médias das gerações RC2: Dist= Distância Genética, Sim= Similaridade, Recup= Recuperação do genoma recorrente de todas as plantas.

População	RCs	Distância Genética	No. Plants	Similaridade com Recorrente			Recuperação do Genoma Recorrente		
				Média	Desvio Padrão	Amplitude	Média	Desvio Padrão	Amplitude
PR1XPD1	RC2	0,238	30	98,64%	0,38%	98,07% - 99,71%	94,48%	1,78%	90,37% - 98,81%
PR2XPD1	RC2	0,256	48	97,74%	0,68%	95,78% - 99,22%	91,70%	2,52%	83,85% - 97,12%
PR3XPD1	RC2	0,286	48	98,29%	0,53%	96,94% - 99,22%	94,54%	1,69%	90,19% - 97,41%
Média Grupo1		0,260	42	98,22% a			93,57% a		
PR4XPD1	RC2	0,335	41	97,04%	1,30%	93,08% - 98,81%	92,54%	2,89%	85,18% - 96,99%
PR5XPD1	RC2	0,307	31	97,29%	0,66%	95,76% - 98,45%	92,22%	2,08%	87,83% - 97,03%
PR6XPD1	RC2	0,309	46	96,89%	1,01%	94,32% - 99,34%	90,82%	2,87%	84,19% - 98,06%
Média Grupo2		0,317	39	97,08% ab			91,86% a		
PR7XPD1	RC2	0,452	26	96,65%	1,12%	94,81% - 98,39%	93,29%	2,41%	89,39% - 96,81%
PR8XPD1	RC2	0,408	30	96,54%	0,92%	94,30% - 98,70%	92,49%	2,24%	87,18% - 97,72%
PR9XPD1	RC2	0,499	86	95,25%	2,50%	82,94% - 98,97%	91,02%	3,57%	75,74% - 98,35%
Média Grupo3		0,453	47	96,14% b			92,27% a		
Média Geral		0,343	43	97,15%			92,57%		
Correlação Dist x Sim				-0,900358****					
Correlação Dist x Recup				-0,34507					

Médias seguidas por letras iguais nas colunas não diferem pelo teste de Tukey a 5% de probabilidade

**, * - Significativo a 1 e 5% de probabilidade pelo teste t

++, + - Significativo a 1 e 5% de probabilidade pelo teste de Mantel baseado em 5000 simulações

Tabela 6. Similaridade genética entre as plantas selecionadas da segunda geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento e Correlação de Pearson / Teste de Mantel, médias das gerações RC2: Dist= Distância Genética, MedSel= Média das similaridades das cinco plantas selecionadas e RecSel= Média da recuperação do genoma recorrente das cinco plantas selecionadas.

População	RCs	Distância Genética	No. Plants	Similaridade com Recorrente			Recuperação do Genoma Recorrente		
				Média	Desvio Padrão	Amplitude	Média	Desvio Padrão	Amplitude
PR1XPD1	RC2	0,238	5	99,18%	0,31%	98,93% - 99,71%	96,98%	1,04%	96,27% - 98,81%
PR2XPD1	RC2	0,256	5	98,74%	0,30%	98,45% - 99,22%	95,61%	0,89%	94,80% - 97,12%
PR3XPD1	RC2	0,286	5	98,93%	0,27%	93,08% - 99,22%	97,08%	0,28%	96,76% - 97,41%
Média Grupo1		0,260		98,95% a			96,56% a		
PR4XPD1	RC2	0,335	5	98,58%	0,19%	98,33% - 98,81%	96,45%	0,41%	95,92% - 96,99%
PR5XPD1	RC2	0,307	5	98,24%	0,22%	97,87% - 98,45%	95,43%	1,16%	93,85% - 97,03%
PR6XPD1	RC2	0,309	5	98,63%	0,44%	98,22% - 99,34%	96,10%	1,19%	94,89% - 98,06%
Média Grupo2		0,317		98,48% ab			96,00% a		
PR7XPD1	RC2	0,452	5	98,10%	0,24%	97,82% - 98,39%	96,26%	0,52%	95,67% - 96,81%
PR8XPD1	RC2	0,408	5	97,84%	0,53%	97,35% - 98,70%	95,93%	1,07%	94,95% - 97,72%
PR9XPD1	RC2	0,499	5	97,37%	0,71%	97,05% - 98,97%	95,77%	1,45%	94,90% - 98,35%
Média Grupo3		0,453		97,77% b			95,99% a		
Média Geral		0,343		98,40%			96,18%		
Correlação Dist x MedSel				-0,91702****					
Correlação Dist x RecSel				-0.31262					

Médias seguidas por letras iguais nas colunas não diferem pelo teste de Tukey a 5% de probabilidade

**, * - Significativo a 1 e 5% de probabilidade pelo teste t

++, + - Significativo a 1 e 5% de probabilidade pelo teste de Mantel baseado em 5000 simulações

Tabela 7. Similaridade genética entre parentais doadores e recorrente similaridades genéticas entre populações da terceira geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento e Correlação de Pearson / Teste de Mantel, médias das gerações RC3: Dist= Distância Genética, Sim= Similaridade, Recup= Recuperação do genoma recorrente de todas as plantas.

População	RCs	Distância Genética	No. Plants	Similaridade com Recorrente			Recuperação do Genoma Recorrente		
				Média	Desvio Padrão	Amplitude	Média	Desvio Padrão	Amplitude
PR1XPD1	RC3	0,238	86	98,46%	0,35%	97,56% - 99,15%	95,52%	1,34%	91,71% - 98,28%
PR2XPD1	RC3	0,256	48	98,82%	0,46%	97,39% - 99,65%	97,24%	1,04%	94,60% - 99,14%
PR3XPD1	RC3	0,286	81	99,22%	0,66%	94,19% - 99,80%	97,93%	1,22%	90,79% - 99,73%
Média Grupo1		0,260	72	98,84% a			96,90% a		
PR4XPD1	RC3	0,335	80	98,38%	0,90%	93,76% - 99,72%	96,80%	1,69%	88,69% - 99,69%
PR5XPD1	RC3	0,307	85	99,04%	0,43%	96,86% - 99,83%	97,91%	0,83%	94,30% - 99,90%
PR6XPD1	RC3	0,309	99	98,38%	0,50%	95,83% - 99,11%	95,96%	1,35%	90,50% - 98,25%
Média Grupo2		0,317	88	98,60% a			96,89% a		
PR7XPD1	RC3	0,452	79	98,35%	1,01%	92,89% - 99,70%	97,23%	1,69%	86,96% - 99,57%
PR8XPD1	RC3	0,408	82	98,35%	0,51%	97,20% - 99,49%	96,98%	0,97%	94,80% - 99,22%
PR9XPD1	RC3	0,499	78	98,48%	1,27%	92,34% - 99,91%	97,20%	2,03%	89,81% - 99,81%
Média Grupo3		0,453	80	98,39% a			97,14% a		
Média Geral		0,343	80	98,61%			96,97%		
Correlação Dist x Sim				-0,45451					
Correlação Dist x Recup							0.24366		

Médias seguidas por letras iguais nas colunas não diferem pelo teste de Tukey a 5% de probabilidade

**, * - Significativo a 1 e 5% de probabilidade pelo teste t

++, + - Significativo a 1 e 5% de probabilidade pelo teste de Mantel baseado em 5000 simulações

Tabela 8. Similaridade genética entre as plantas selecionadas da terceira geração de retrocruzamento com o genitor recorrente, em nove populações de retrocruzamento e Correlação de Pearson / Teste de Mantel, médias das gerações RC3: Dist= Distância Genética, MedSel= Média das similaridades das cinco plantas selecionadas e RecSel= Média da recuperação do genoma recorrente das cinco plantas selecionadas.

População	RCs	Distância Genética	No. Plants	Similaridade com Recorrente			Recuperação do Genoma Recorrente		
				Média	Desvio Padrão	Amplitude	Média	Desvio Padrão	Amplitude
PR1XPD1	RC3	0,238	5	99,09%	0,05%	99,03% - 99,15%	97,83%	0,34%	97,44% - 98,28%
PR2XPD1	RC3	0,256	5	99,40%	0,24%	99,00% - 99,65%	98,81%	0,20%	98,62% - 99,14%
PR3XPD1	RC3	0,286	5	99,69%	0,09%	99,59% - 99,80%	99,44%	0,17%	99,33% - 99,73%
Média Grupo1		0,260		99,40% a			98,69% a		
PR4XPD1	RC3	0,335	5	99,48%	0,20%	99,18% - 99,72%	99,05%	0,37%	98,80% - 99,69%
PR5XPD1	RC3	0,307	5	99,64%	0,11%	99,55% - 99,83%	99,31%	0,36%	99,04% - 99,90%
PR6XPD1	RC3	0,309	5	98,99%	0,13%	98,85% - 99,11%	97,96%	0,17%	97,86% - 98,25%
Média Grupo2		0,317		99,37% a			98,77% a		
PR7XPD1	RC3	0,452	5	99,50%	0,14%	99,33% - 99,70%	99,35%	0,18%	99,12% - 99,57%
PR8XPD1	RC3	0,408	5	99,38%	0,12%	99,20% - 99,49%	98,97%	0,25%	98,56% - 99,22%
PR9XPD1	RC3	0,499	5	99,78%	0,08%	99,71% - 99,91%	99,55%	0,16%	99,43% - 99,83%
Média Grupo3		0,453		99,55% a			99,29% a		
Média Geral		0,343		99,44%			98,92%		
Correlação Dist x MedSel				0,47229					
Correlação Dist x RecSel							0,57665		

Médias seguidas por letras iguais nas colunas não diferem pelo teste de Tukey a 5% de probabilidade

**, * - Significativo a 1 e 5% de probabilidade pelo teste t

++, + - Significativo a 1 e 5% de probabilidade pelo teste de Mantel baseado em 5000 simulações

5. CONCLUSÃO

A avaliação da similaridade genética com maior número de marcadores fornece informações mais relevantes para o gerenciamento dos programas de RC do que a avaliação da recuperação do genoma recorrente utilizando apenas marcadores polimórficos entre os parentais.

A utilização de genitores doadores geneticamente mais próximos dos genitores recorrentes não pode reduzir de três para duas gerações de RC para a recuperação do genoma recorrente, em programas de retrocruzamento assistido por marcadores moleculares.

6. REFERÊNCIAS

Arif IA, Bakir MA, Khan HA, Farhan AHAI, Homaidan AAI, Bahkali AH, Sadoon MAI, Shobrak M (2010) A brief review of molecular techniques to assess plant diversity. **International Journal of Molecular**, 11: 2079-2096.

Beckett JB. (1971) Classification of male-sterile cytoplasm in maize (*Zea mays* L.). **Crop Sci.** 11:724–724.

Bellon G, Faleiro FG, Junqueira NTV, Fuhrmann E (2014) Variabilidade genética de genótipos elite de maracujazeiro, obtidos em programas de retrocruzamento envolvendo espécies silvestres e comerciais com base em marcadores RAPD. **Biosci. J.** 6: 1692-1697.

Borém A, Miranda GV (2005), Melhoria de Plantas, **Editora UFV**, 4: 1-546.

Bouchez A, Hospital F, CAUSSEE M, Gallais A, Charcosset A (2002) Marker-assisted introgression of favorable alleles at quantitative trait loci between maize elite lines. **Genetics** 162: 1945-1959.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. (2007) [TASSEL: Software for association mapping of complex traits in diverse samples.](#) **Bioinformatics** 23:2633-2635.

Caixeta ET, Pestana KN, Pestana RK (2015) Melhoria do Cafeeiro: Ênfase na aplicação dos marcadores moleculares. **Embrapa Café – Capítulo em livro científico**, Tópico Especiais em Produção Vegetal V: 154-179.

Chen L, Liu YG (2014) Male sterility and fertility restoration in crops. **Plant Biol** 65: 579 – 606.

Cheng X, Yonghong R, Yianqiao J, Zifeng G, Yan Z, Chuanxiao X, Junjie F, Hongwu W, Guoying W, Yunbi X, Ping L, Cheng Z (2017). Development of a maize 55 K SNP array with improved genome coverage for molecular breeding. **Molecular Breeding**, 37: 1-20.

Conab (2021). Produção de grãos deve chegar a 268,3 milhões de toneladas, diz Conab. **Agência Brasil**. Disponível em: <https://www.conab.gov.br/info-agro/safras/serie-historica-das-safras?start=20>. Acesso em 06 de maio de 2021.

Cruz CD (2014) Programa Genes - versão Windows: aplicativo computacional em genética e estatística. **Viçosa UFV**.

Engelsing MJ, Coimbra JLM, Vale NM, Barili LD, Stingenen JC, Guidolin AF, Bertildo JG (2012) Adaptabilidade e estabilidade em milho: rendimento de grãos x severidade de cercosporiose. **Revista de Ciências Agroveterinárias** 11: 106-117.

Guimarães CT, Schuster I, Magalhães JV, Souza Junior CL (2016) Marcadores moleculares no melhoramento de plantas. In: Borem, A., Caixeta, E.T. (eds). **Marcadores Moleculares, Viçosa, Ed. UFV 3: 129-177.**

He J, Zhao X, Laroche A, Lu Z.X, Liu H, Li, Z (2014) Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. **Frontiers in Plant Sciences 5: 1-8.**

Hehir-Kaw J, Egmont-Petersen M, Janssen IM, Smmets D, Van Kessel AG, Veltman JA (2007) Genome-wide copy number profiling on high-density BAC, SNP and oligonucleotide microarray: a platform comparison based on statistical power analysis. **DNA Res 14: 1-11.**

Holland PM, Abramson RD, Watson R, Gelfand DH (1991) Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of *Thermus aquaticus* DNA polymerase, **Proc. Natl. Acad. Sci. USA 88: 7276-7280.**

Hospital F (2001) Size of donor chromosome segments around introgressed loci and reduction of linkage drag in marker-assisted backcross programs, **Genetics 158: 1363-1379.**

ISAA (2019) Facts about Biotech/GM Crops in 2018: Report Reveals Biotech Crops Continue to Provide Solutions to Hunger Malnutrition and Climate Change, Disponível em: <https://www.isaaa.org/resources/publications/briefs/55/executivesummary/default.asp>, **ISAA Brief No.55**, Acesso em 08 de maio de 2021.

McGuigan FEA, Ralston SH (2002) Single nucleotide polymorphism detection: allelic discrimination using Taqman. **Psychiat Genet 12: 133-366.**

Mesquita AGG, Guimarães CT, Parentoni SN, Paiva E (2005) Recuperação do Genitor Recorrente em Milho Utilizando Retrocruzamento Assistido por Marcadores Microsatélites. **Revista Brasileira de Milho e Sorgo 4: 275-285.**

MESQUITA, AGG (2002) **Retrocruzamento assistido por marcadores SSRs em milho.** 87 f. Tese Doutorado em Genética e Melhoramento de Plantas. Lavras, UFLA

Miranda RA, (2019) Clube dos 100 milhões: Progresso técnico e a evolução da produção de milho no Brasil. Embrapa Milho e Sorgo – Boletim Grão em Grão. **Jornal Eletrônico da Embrapa Milho e Sorgo, Sete Lagoas-MG 13: 1-5.**

Mumm RH (2013) A look at product development with genetically modified crops: examples from maize. **Journal of Agricultural and Food Chemistry 61:8254-8259.**

Mumm RH, Walters DS (2001) Quality control in the development of transgenic crop seed products. **Crop Sci 41:1381-1389.**

Openshaw SJ, Jarboe SG, Beavis WD (1994) Marker assisted selection in backcross breeding. In: Symposium Analysis of Molecular Marker Data, Corvallis, Oregon. Proceedings...Corvallis: **American Society for Horticultural Science/ Crop Science of America**, p.41-43.

Peng T, Sun X, Munn RH (2013) Optimized breeding strategies for multiple trait integration: I. Minimizing linkage in single event introgression, **Molecular Breeding** 33: 89-104.

Pinto MO, Souza IRP, Paula ALSP, Guimarães PE, Trindade RS, Guimarães LJM (2018) Introgressão da resistência ao mosaico-comum em milho assistida por marcadores snp. **Embrapa Milho e Sorgo**, Sete Lagos, MG, Brasil.

Podder M, Ruan J, Tripp BW, Chu ZE e Tebbut SJ (2008) Robust SNP genotyping by multiplex PCR and arrayed extension. **BMC Med Genomics** 1: 1-15.

Prince JA, Feuk L, Howell WM, Jobs M, Emahazion T, Blennow K e Brookes AJ (2001) Robust and accurate single nucleotide polymorphism genotyping by dynamic allele-specific hybridization (DASH): Design criteria and assay validation. **Genome Res** 11: 152-1162.

Ramos LN, (2018) **Híbridos de milho geneticamente modificados comparados a híbridos de milho convencionais em terras altas do Brasil central**. 98f Tese (Doutorado) Faculdade de Agronomia e Medicina Veterinária, Universidade de Brasília.

Stevanovic M, Camdzija Z, Pavlov J, Markovic K, Vancetovic J, Drinic SM, Filipovic M. (2016) The application of protein markers in conversion of maize inbred lines to the cytoplasmic male sterility basis. **Genetika** 48:8-691.

Tang J, Hu Y, Ji H, Chen W, Ji L, Zhen Y. (2001) The SSR marker of the restoring gene Rf4 for CMS-C cytoplasmic male sterility in maize. **J Henan Agric Univ.** 35:1–03.

Torkamaneh D, Laroche J e Belzile F (2016) Genome-wide SNP calling from genotyping by sequencing (GBS) data: A comparison of seven pipelines and two sequencing technologies. **PLoS One** 1-14.

Tsuchihashi Z, Dracopoli, NC (2002). Progress in high throughput SNP genotyping methods. *Pharmacogenomics Journal*. 2: 103-110.

Turchetto-Zolet AC, Turchetto C, Guzman F, Silva-Arias GA, Sperb-Ludwing F e Veto NM (2017). Polimorfismo de nucleotídeo único (SNP): metodologia de identificação, análise e aplicações, **Marcadores Moleculares na Era Genômica: Metodologia e Aplicações – SBG** 1: 132-166.

Wang J, Chapma SC, Bonnett DG, Rebetzke GJ, Crouch J (2007) Application of population genetic theory and simulation models to efficiently pyramid multiple genes via marker-assisted selection. **Crop Sci** 47:582–590.

Wang N, Yuan Y, Wang H, Yu D, Liu Y, Gowda AZM, Nair SK, Hao Z, Lu Y, Vicente FS, Prasanna BM, Li X, Zhang X (2020) Applications of genotyping-by-sequencing (GBS) in maize genetics and breeding. **Nature Research** 10: 16-308.

Wilhelm, WW, Johson, BE, Schepers JS (1995) Yield, quality, and nitrogen use of inbred corn with varying numbers of leaves during detasseling. **Crop Sci.** 35: 209-212.