

FLAVIA ANAN SAIKI

**DESENVOLVIMENTO E VALIDAÇÃO DE MARCADORES
MICROSSATÉLITES PARA *Mimosa scabrella* Benth.**

Dissertação apresentada como requisito parcial para obtenção do título de mestre no Curso de Pós-Graduação em Produção Vegetal da Universidade do Estado de Santa Catarina - UDESC.

Orientador: Prof. Dr. Altamir Frederico Guidolin

Coorientador: Prof. Dr. Adelar Mantovani

**LAGES, SC
2016**

Saiki, Flavia Anan
DESENVOLVIMENTO E VALIDAÇÃO DE MARCADORES
MICROSSATÉLITES PARA *Mimosa scabrella* Benth. /
Flavia Anan Saiki. Lages - 2016.
91 p.

Orientador: Altamir Frederico Guidolin
Co-orientador: Adelar Mantovani
Dissertação (Mestrado) - Universidade do Estado
de Santa Catarina, Centro de Ciências
Agroveterinárias, Programa de Pós-Graduação em
Produção Vegetal, Lages, 2016.

1. Marcador molecular. 2. SSR nuclear. 3.
Diversidade genética. 4. Sequenciamento de Segunda
Geração. 5. cpSSR. I. Guidolin, Altamir Frederico.
II. Mantovani, Adelar. , .III. Universidade do
Estado de Santa Catarina, Centro de Ciências
Agroveterinárias, Programa de Pós-Graduação em
Produção Vegetal. IV. Titulo.

Ficha catalográfica elaborada pelo autor, com auxílio do
programa de geração automática da Biblioteca Setorial do
CAV/UDESC.

FLAVIA ANAN SAIKI

**DESENVOLVIMENTO E VALIDAÇÃO DE MARCADORES
MICROSSATÉLITES PARA *Mimosa scabrella* Benth.**

Dissertação apresentada ao curso de Mestrado em Produção Vegetal, da Universidade do Estado de Santa Catarina, como requisito parcial para obtenção do grau de Mestre em Produção Vegetal.

Banca Examinadora:

Orientador: Prof. Dr.: Altamir Frederico Guidolin
Universidade do Estado de Santa Catarina – UDESC, Lages SC

Membro interno: Prof^a. Dra.: Denise Olkoski
Universidade do Estado de Santa Catarina – UDESC, Lages SC

Membro externo: Dr.: Gustavo Henrique Ferrero Klabunde
Universidade Federal de Santa Catarina – UFSC, Florianópolis SC

Lages, SC 26/08/2016

DADOS CURRICULARES

Flavia Anan Saiki, bióloga formada pela faculdade de Ciências Agrárias e Veterinárias/FCAV – UNESP do campus de Jaboticabal-SP. Trabalhou como técnica de laboratório de 2010 a 2012 no Laboratório de Genética Bioquímica do Departamento de Genética da Faculdade de Medicina de Ribeirão Preto da USP campus de Ribeirão Preto-SP. Em novembro de 2012 começou a trabalhar no Laboratório DNA UDESC como Professora Colaboradora do Centro de Ciências Agroveterinárias da Universidade do Estado de Santa Catarina -UDESC, atuando na área de Biologia Molecular e Determinação de Paternidade. Em julho de 2014 ingressou no curso de Mestrado do Programa de Pós-Graduação em Produção Vegetal da Universidade do Estado de Santa Catarina – UDESC.

*Às forças maiores que regem a
vida; aos meus pais, meus
maiores incentivadores e
torcedores; ao meu amado
marido, meu agente de
transformação.
Dedico...*

AGRADECIMENTOS

Resumir em palavras a gratidão que sinto é uma tarefa que por mais que eu tente, será sempre incompleta. Tornar-se mestre é uma conquista individual, mas só pude trilhar este caminho, porque foi compartilhado com pessoas e forças que me inspiraram, me apoiaram, me acolheram e principalmente me ajudaram. Chegar em um ponto da vida e perceber que eu pude e posso contar com, é uma honra. Eu devo isso a todos vocês.

Ao meu orientador, Prof. Dr. Altamir Frederico Guidolin, que acreditou em mim mais do que eu mesma.

Ao meu coorientador, Prof. Dr. Adelar Mantovani pelo apoio, por direcionar meu trabalho, e sempre me auxiliar perante as dificuldades. Aos colegas de Pós-Graduação e alunos do Grupo Uso e Conservação de Recursos Florestais: Guilherme, Thiely, Lilian, Leandro, Oieler, Bruna, Lucas e Luigi. Ver um aluno se encantar e dedicar-se a uma área de estudo é uma alegria, obrigada Bruna por me fez bater no peito novamente a alegria de ensinar.

Aos meus maravilhosos pais, Sergio e Irene. Aos meus queridos irmãos Rodrigo e Simone. Minha família, minha certeza, meu time.

Ao Thiago, meu companheiro de vida, meu amor, meu braço direito e esquerdo, minha calma, minha certeza de que sempre poderemos superar qualquer dificuldade; e aos nossos queridos filhos-cães que completam nossa família.

À equipe do Laboratório DNA UDESC e a imprescindível ajuda e cooperação da Fabiana e Denise.

Ao Prof. Dr. Maurício Sedrez dos Reis e aos pós-graduandos do NPFT-UFSC: Alison, Miguel, Marcia e Newton por todo o auxílio técnico e apoio.

Ao Prof. Jefferson Luís Meirelles Coimbra e à equipe do IMEGEM sempre dispostos a contribuir e preparar os

alunos para os desafios da pós e da vida. Às queridas Nicole, Muri e Thayse que me deram a mão no intensivão para a prova de ingresso no Mestrado.

Aos queridos profissionais do Departamento de Genética da USP Ribeirão Preto e egressos espalhados pelo mundo, que hoje tenho a honra de chamar de amigos: Prof. Aguinaldo Simões, Prof^a Eucleia (*in memorian*), Maria do Carmo, Ana Lúcia, Edna, Natália, Juliana, Fernanda, Leonardo, Rafael e Silviene.

Às minhas queridas Jades de Jaboticabal, Ana e Mariana. Aos meus amigos de Taubaté, Gabriel, Marcela, Tânia, Nathália e Thaís. Aos queridos Tito, Ju, Carlos, Mari e todo o pessoal que me acolheu em Lages.

À VIDA, que nesta jornada de superação me fez ver que só tenho a agradecer.

Namastê!

RESUMO

SAIKI, Flavia Anan. **Desenvolvimento e validação de marcadores microssatélites para *Mimosa scabrella* Benth.** 2016. 91 p. Dissertação (Mestrado em Produção Vegetal) – Universidade do Estado de Santa Catarina. Programa de Pós-Graduação em Produção Vegetal, Lages, 2016.

Microssatélites ou *simple sequence repeats* (SSR) são marcadores genéticos codominantes e altamente polimórficos usados para estudos de genética de populações. *Mimosa scabrella* Benth., popularmente conhecida como bracatinga, é uma espécie pioneira e endêmica do Brasil, ocorrendo na Floresta Ombrófila Mista, associado ao bioma Mata Atlântica. É uma espécie arbórea da família Fabaceae, de rápido crescimento, que facilita a dinâmica de sucessão ecológica. O desenvolvimento de SSR quando não há sequências é demorado e laborioso, e não há marcadores moleculares desenvolvidos para *Mimosa scabrella* Benth. Devido ao tempo reduzido, o uso de sequenciamento de segunda geração é uma ferramenta poderosa para o desenvolvimento de microssatélites nos estudos de genética de populações. O presente trabalho se propôs a desenvolver e validar os primeiros marcadores microssatélites para a espécie, avaliando matrizes e progêniens. Usando o sequenciamento de segunda geração pela plataforma Illumina, este trabalho identificou 290 locos SSR e 211 pares de primers. Um subconjunto de 11 pares de primers foi sintetizado com fluorescência no primer forward para validação por PCR e eletroforese capilar a partir do DNA de folhas de 33 adultos e 411 progêniens. A porcentagem de locos polimórficos foi de 36, quatro locos em 11, três SSR cloroplastidiais e um SSR nuclear. O número de alelos por loco polimórfico variou de 2 a 11 considerando toda a amostragem.

O sequenciamento de segunda geração possibilitou a identificação de um grande número de locos microssatélites para a bracatinga, resultando no desenvolvimento de marcadores SSR para a espécie. Os marcadores SSR avaliados e validados são aptos e úteis para análises e estudos de genética de populações. Outros marcadores SSR foram identificados, mas ainda precisam ser sintetizados e avaliados quanto ao polimorfismo e consistência entre matrizes e progênies para validação.

Palavras-chave: Marcador molecular. SSR nuclear. Diversidade genética. Sequenciamento de Segunda Geração. cpSSR

ABSTRACT

SAIKI, Flavia Anan. **Microsatellite markers development and validation for *Mimosa scabrella* Benth.** 2016. 91 p. Dissertation (Master of Plant Production) - Santa Catarina State University. Postgraduate Program in Plant Production, Lages, 2016.

Microsatellites or simple sequence repeats (SSR) are co-dominant and highly polymorphic genetic markers used for population genetic studies. *Mimosa scabrella* Benth., popularly known as “bracatinga”, is a pioneer and endemic species of Brazil, occurring in Mixed Ombrophilous Forest associated to Brazilian Atlantic Rainforest biome. It is a fast-growing tree of Fabaceae family which facilitates the dynamics of ecological succession. SSR development when there is no sequence is time and labor intensive, there are no molecular markers for *Mimosa scabrella* Benth. Due to reduced time, the use of second generation sequencing is a powerful tool for microsatellite development for population genetic studies. This current project proposed to develop and validate the first microsatellite markers for the species, evaluating mother trees and progenies. Using Second Generation Sequencing of Illumina platform, this project identified 290 SSR *loci* e 211 primer pairs. A subset of 11 primer pairs was synthetized with fluorescence in the Forward primer for PCR and capillary electrophoresis validation with leaf DNA of 33 adult and 411 progeny individuals. Polymorphic *loci* percentage was 36, four in 11 *loci*, 3 chloroplast SSR and one nuclear SSR. Allele number of polymorphic *loci* ranged from two to 11 alleles considering all sampling. Second-generation sequencing has enabled a large number of microsatellite loci identification for bracatinga, resulting in SSR markers development for the species. The assessed and validated SSR markers are suitable

and useful for analysis and population genetic studies. Other SSR markers were identified, but still need to be synthesized and evaluated for polymorphism and consistency between mother trees and progenies for validation.

Keywords: Molecular marker. Nuclear SSR. Genetic diversity. Second Generation Sequencing. cpSSR.

LISTA DE ILUSTRAÇÕES

Figura 1: a) Bracatingas, b) Destaque das inflorescências de bracatinga.....	34
Figura 2: Biplex com os marcadores Msc161 e Msc003.....	59

LISTA DE TABELAS

Tabela 1: Locos microssatélites separados pelo tamanho do motivo	52
Tabela 2: Pares de primers aprovados para análise de <i>M. scabrella</i>	52
Tabela 3: Características dos 11 pares de primers microssatélites e painéis desenvolvidos para <i>Mimosa scabrella</i> Benth.....	58
Tabela 4: Frequências alélicas dos locos polimórficos de cloroplasto.....	60
Tabela 5: Frequências alélicas do SSR nuclear Msc003.	61
Tabela 6: Estatística de diversidade do SSR nuclear Msc003 para os indivíduos adultos.....	61
Tabela 7: Estatística de diversidade do SSR nuclear Msc003 para as progênies	62

SUMÁRIO

1 INTRODUÇÃO	29
2 REVISÃO	33
2.1 BRACATINGA	33
2.2 SEQUENCIAMENTO DE SEGUNDA GERAÇÃO	37
2.3 MARCADORES MICROSSATÉLITES	39
2.4 MICROSSATÉLITES CLOROPLASTIDIAS	41
3 MATERIAL E MÉTODOS.....	44
3.1 EXTRAÇÃO DE DNA GENÔMICO E AVALIAÇÃO DE CONCENTRAÇÃO E QUALIDADE	44
3.2 SEQUENCIAMENTO DE SEGUNDA GERAÇÃO	44
3.3 QUALIDADE DOS DADOS, MONTAGEM DAS SEQUÊNCIAS E IDENTIFICAÇÃO DE MICROSSATÉLITES	45
3.4 AVALIAÇÃO DA QUALIDADE DOS PRIMERS	46
3.5 VALIDAÇÃO DOS MARCADORES	47
3.5.1 Amostragem	47
3.5.2 1 ^a fase de validação	48
3.5.3 2 ^a fase de validação	48
3.6 ANÁLISE DA DIVERSIDADE GENÉTICA	49
4 RESULTADOS	51

4.1 SEQUENCIAMENTO DE SEGUNDA GERAÇÃO.....	51
4.2 IDENTIFICAÇÃO DOS LOCOS MICROSSATÉLITES E DESENHO DE PRIMERS	51
4.3 1 ^a FASE DE VALIDAÇÃO – PCR E ELETROFORESE EM GEL DE AGAROSE	57
4.4 2 ^a FASE DE VALIDAÇÃO – PCR E ELETROFORESE CAPILAR	57
4.5 ANÁLISE DA DIVERSIDADE GENÉTICA	60
5 DISCUSSÃO	63
6 CONCLUSÃO	73
REFERÊNCIAS	74

1 INTRODUÇÃO

Estudos de conservação dos recursos genéticos são muito importantes nos dias de hoje, por fundamentar as melhores formas de preservação da diversidade dentro das espécies. A alteração antrópica do ambiente modifica as estruturas de vegetação (BERTANI; SANTOS, 2013) e, na ausência de um manejo sustentável, a velocidade com que se modifica o ambiente é muito superior à velocidade em que o mesmo consegue se restabelecer. A perda dos habitats naturais pelo domínio das atividades humanas está no centro da crise de biodiversidade atual (BOAKES et al., 2010). A fragmentação dos ambientes leva a uma diminuição do tamanho populacional e isolamento em diferentes magnitudes, que em conjunto comprometem a regeneração das espécies nas áreas fragmentadas (BERTANI; SANTOS, 2013) e podem propiciar erosão genética (SEOANE, 2007).

O fenômeno da erosão genética está associado à deterioração das reservas genéticas de muitas populações (VENCOVSKY, 1987; SEOANE, 2007). A perda e distúrbios nos habitats naturais pela atividade humana, o crescente desenvolvimento tecnológico, a exploração agropecuária irracional são os principais responsáveis por este processo (VENCOKSVKY, 1987; CARVALHO; SILVA; MEDEIROS, 2009). A redução da variabilidade está ocorrendo com diferentes intensidades, conforme a espécie e em vários casos, a situação atual parece ser bastante crítica (VENCOKSVKY, 1987; SOUZA; MARTINS, 2004).

A bracatinga, *Mimosa scabrella* Benth., é uma espécie florestal da família Fabaceae, nativa do Brasil (DUTRA; MORIM, 2016), típica da região Sul, mas também com ocorrência nos estados de Minas Gerais, São Paulo e Rio de Janeiro (CARVALHO, 2003). É uma espécie característica da Floresta Ombrófila Mista, principalmente em planaltos com altitude superior a 350 m (CARVALHO, 2003). Trata-se de

uma pioneira que costuma desenvolver-se em abundância em campos abertos, como resultado da derrubada das florestas de araucária (REITZ; KLEIN; REIS, 1978). Pode ser encontrada associada a outras espécies na vegetação secundária e também em associações puras, conhecidas como bracatingais (STEENBOCK et al., 2011).

O estabelecimento de populações de bracatinga contribui para preservação do patrimônio genético florestal, pois esta espécie pode atuar como facilitadora no processo de regeneração sucessional (BAGGIO; CARPANEZZI, 1998). O rápido estabelecimento da bracatinga em áreas deflorestadas é uma resposta evolutiva que gera bons resultados num curto espaço de tempo (KLEIN, 1981), entretanto ainda não se sabe os efeitos em longo prazo sobre os recursos genéticos dentro das populações. Se houver perda na diversidade e aumento da endogamia, a capacidade evolutiva da espécie fica comprometida (SEOANE, 2007), com consequências sobre todo o ecossistema.

Conhecer os recursos genéticos possibilita utilizá-los sem esgotá-los ou inviabilizá-los, além de promover conservação (BRASIL, 2006a). A legislação brasileira é bastante conservadora na utilização de espécies nativas. O Código Florestal Brasileiro (BRASIL, 2012), a Lei da Mata Atlântica (BRASIL, 2006b), e a lei 11.326/2006 (BRASIL, 2006c) possibilitam o corte seletivo da bracatinga para consumo dentro da propriedade rural, sem fins comerciais. Para fins comerciais, é necessária a autorização do órgão ambiental para manejo da reserva legal da propriedade, desde que não descaracterize a cobertura vegetal, não prejudique a conservação e diversidade de espécies e manejo que favoreça a regeneração de espécies nativas de acordo com o Código Florestal vigente (BRASIL, 2012).

Com a dificuldade burocrática de se utilizar a bracatinga como espécie comercial, demora na tramitação, falta de pacotes tecnológicos para este recurso genético, é natural

que os produtores rurais busquem alternativas que substituam a produção comercial de bracatinga por outras espécies (MAZUCHOWSKI et al., 2014), que em muitos casos tem o agravante de não fazer parte da flora nativa. A bracatinga possui muito usos potenciais, como na recuperação de áreas degradadas (CARNEIRO et al., 1982; URBANO et al., 2008), e também como lenha, carvão, e planta apícola (PEGORARO; CARPANEZZI, 1995; MACHADO et al., 2002). A falta de ferramentas e a diminuição do seu plantio frente a culturas de espécies não nativas reforçam a necessidade de desenvolver tecnologias para lidar com estes problemas. Neste sentido, o desenvolvimento de marcadores moleculares é fundamental para caracterização da espécie, o que possibilita um melhor delineamento de manejo e conservação.

Dentre os marcadores moleculares, os microssatélites são um dos marcadores genéticos mais populares por serem altamente polimórficos, abundantes, codominantes e bem distribuídos pelo genoma (SCHLÖTTERER, 2004). Estas propriedades os tornam adequados para estudos de genética de populações, estrutura genética e diversidade, testes de paternidade, identificação de linhagens e cultivares e mapeamento (BUSO et al., 2003; SCHLÖTTERER, 2004; ABDELKRIM et al., 2009; CAVAGNARO et al., 2011). Os marcadores microssatélites, que são baseados no uso de pares de primers na reação em cadeia da polimerase (PCR), podem ser utilizados no monitoramento e manejo de populações para evitar o processo de erosão genética. A caracterização da diversidade genética pelos marcadores microssatélites dentro e entre populações naturais e cultivadas gera informações para programas de conservação, assim como auxilia nos programas de melhoramento genético das espécies de interesse (SOUZA, 2015).

Embora exista a demanda, foram desenvolvidos poucos marcadores microssatélites (SSR) para espécies nativas, tais como: *Araucaria angustifolia* (Bertol.) Kuntze (SALGUEIRO

et al., 2005; SCHMIDT et al., 2007), *Paraptdenia rigida* (Benth.) Brenan (COSTA et al., 2012), *Dicksonia sellowiana* Hooker (NAZARENO et al., 2013), *Podocarpus lambertii* Klotzch ex Endl (VIEIRA et al., 2014). Tradicionalmente, o desenvolvimento de SSR por triagem de bibliotecas de bactérias recombinantes demanda tempo, é custosa e trabalhosa (BLAIR et al., 2009). Entretanto, com o desenvolvimento de uma nova geração de sequenciadores, esta etapa tornou-se mais simples e rápida, o que viabilizou a identificação de um maior número de marcadores microssatélites para uso nas análises genéticas de populações (ZALAPA et al., 2012), superando a maior desvantagem da utilização de microssatélites, que é a falta de sequências da espécie com que se deseja trabalhar (ABDELKRIM et al., 2009).

Considerando os usos potenciais da bracatinga, a necessidade de conservação, a diminuição de seu plantio e a falta de ferramentas para analisar a diversidade, este presente trabalho teve como objetivo desenvolver e validar os primeiros marcadores microssatélites para *Mimosa scabrella* Benth. por sequenciamento de segunda geração.

2 REVISÃO

2.1 BRACATINGA

A bracatinga (Figura 1a) caracteriza-se pelo rápido crescimento e pode atingir 20 m de altura, com diâmetro de tronco de até 40 cm (REITZ et al., 1978). É uma planta alógama (CARVALHO, 2003), com flores polinizadas principalmente por abelhas do tipo *Apis* e *Trigona* e sementes dispersas por autocoria (SOBIERAJSKI; KAGEYAMA; SEBBENN, 2006). O crescimento sofre pouca influência das condições físicas do solo, desenvolve-se naturalmente em solos pobres com pH ácido e deficientes em nutrientes importantes como o fósforo e o potássio (BURKART, 1979; BARTOSZECK, 2004). Por outro lado, a compactação do solo parece ser um fator que interfere no desenvolvimento (CATIE, 1986; CARVALHO, 2003), pois se trata de uma espécie adaptada às condições edáficas de margens de rios e bordas de capões (REITZ; KLEIN; REIS, 1978).

A formação de bracatingais, caracterizados pelo adensamento monoespecífico de 3000 a 4000 indivíduos por hectare (MACHADO et al., 2006), é favorecida pela rusticidade e precocidade desta espécie. A maturidade sexual é atingida com cerca de três anos (SOBIERAJSKI; KAGEYAMA; SEBBENN, 2006) e cada árvore pode produzir até 40.000 inflorescências (Figura 1b), (HARTER-MARQUES; ENGELS, 2003). As sementes apresentam dormência tegumentar e são impermeáveis à água (CARPANEZZI; PAGANO; BAGGIO, 1997). A queimada, prática bastante difundida no sistema agroflorestal no sul do Brasil, quebra a dormência das sementes contribuindo para uma germinação abundante (STEENBOCK; REIS, 2013).

Figura 1: a) Bracatingas, b) Destaque das inflorescências de bracatinga.

a)



b)



Fonte: Produção do próprio autor.

Do ponto de vista econômico a bracatinga possui várias aplicações, pois tem potencial de gerar uma grande variedade de produtos madeiráveis (CARNEIRO et al., 1982;

CARPANEZZI; PAGANO; BAGGIO, 1997) e pode ser utilizada na recuperação de áreas degradadas e proteção de solos, uma vez que tem a capacidade de cobrir rapidamente os solos e fixar nitrogênio (CARPANEZZI; PAGANO; BAGGIO, 1997). Também pode ser utilizada como planta apícola (PEGORARO; CARPANEZZI, 1995), forrageira (CARPANEZZI et al., 1988) e tem potencial biotecnológico, como na produção de glucomananas, biopolímeros de carboidratos utilizados na indústria farmacêutica (GANTER; REICHER, 1999; UGHINI et al., 2004; VENDRUSCOLO et al. 2005).

A bracatinga possui grande importância socioeconômica para região Sul do Brasil (SOBIERAJSKI; KAGEYAMA; SEBBENN, 2006), destacando-se principalmente em programas de reflorestamento, na recuperação de áreas degradadas (CARNEIRO et al., 1982; URBANO et al., 2008) e na produção de madeira para energia (MACHADO et al., 2002). O manejo racional da bracatinga para produção de lenha remonta ao inicio do século 20 no Paraná, podendo ser caracterizado como o mais antigo sistema deste tipo no Brasil (BAGGIO et al., 1986). Na região noroeste do planalto catarinense, o manejo dos bracatingais obtidos através de intervenções silviculturais é uma importante fonte de renda para assentados da reforma agrária (STEENBOCK, 2009). Apesar das restrições da legislação ambiental ao manejo de bracatingais naturais (STEENBOCK, 2009), nos últimos anos a madeira da bracatinga tem sido bastante procurada para usos industriais nobres, como a movelearia e laminação (ZAMARIAN, 2008), aumentando ainda mais a demanda por técnicas de produção desta espécie.

Poucos dados genéticos e de marcadores moleculares estão descritos para a bracatinga. Sobierajski, Kageyama e Sebbenn (2006) investigaram o sistema de reprodução por isoenzimas em nove populações e verificaram que a espécie se reproduz por sistema misto com predomínio de cruzamentos, e

que parte destes cruzamentos ocorre entre indivíduos aparentados. Moreira et al. (2011) analisaram a diversidade genética e o padrão reprodutivo de 14 bracatingais utilizando 8 locos de isoenzimas, e identificaram altos índices de diversidade e tendência à estruturação genética pelo sistema reprodutivo e deriva, concluindo que o manejo da paisagem natural influencia positivamente na manutenção da diversidade genética. Estudos de citogenética para o gênero *Mimosa* L. identificaram que *M. scabrella* é uma espécie tetraploide com $2n=4x=52$ (DAHMER et al, 2011). Dahmer et al. (2013) acrescentaram análises de marcadores RAPD comparados a dados citogenéticos nos estudos de variação da diversidade e observaram que em 25 populações amostradas não houve variação no número de cromossomos, todas foram $2n= 4x =52$, e para 14 populações, a análise de RAPD revelou um alto nível de variação molecular entre as populações.

A poliploidia é a variação cromossômica na qual há três ou mais conjuntos cromossômicos completos e é encontrada em muitos táxons de eucariotos (RAMSEY; SCHEMSKE, 1998). Ocorre de duas formas: alopóliploides e autopoliploides. Nos alopóliploides, ocorre a união de genomas de duas ou mais espécies diferentes. Os autopoliploides são originados pela duplicação de um mesmo genoma ou mais comumente pela união de gametas não reduzidos (SCHIFINO-WITTMANN; DALLA'GNOL, 2001). A distinção entre alo e autopoliploide afeta o pareamento e segregação dos cromossomos homólogos durante a meiose, no qual os alopóliploides formam bivalentes e nos autopoliploides formam multivalentes no pareamento (MEIRMANS; VAN TIENDEREN, 2013).

A poliploidia está associada com aumento no tamanho das células em muitos organismos e esta característica comumente afeta o tamanho das plantas, o que a torna útil para o melhoramento de plantas (TSUKAYA, 2013). Muitas das grandes culturas agronômicas são poliploides, e apesar da importância da tetraploidia para as plantas, a maior parte dos

estudos de genética com marcadores microssatélites é baseada em espécies diploides. Enquanto as ferramentas disponíveis para diploides estão cada vez mais desenvolvidas e numerosas, a pesquisa em poliploidia está muito atrás (PUYVELDE; GEERT; TRIEST, 2010). Um grande desafio na análise de dados populacionais de poliploides ainda é sobre o modo de herança e temporalizar o evento de poliploidização (MEIRMANS; VAN TIENDEREN, 2013).

2.2 SEQUENCIAMENTO DE SEGUNDA GERAÇÃO

Desde o início da década de 90, o sequenciamento de DNA é realizado basicamente por eletroforese capilar semi-automatizada com o método de terminação de cadeia de Sanger, com ddNTPs marcados (SHENDURE; JI, 2008). Entretanto, mais recentemente, novas formas alternativas de sequenciamento de DNA foram criadas, entre elas a segunda geração em relação ao sequenciamento de Sanger (SANGER et al., 1977), mais conhecida como os sequenciamentos de nova geração. Esta nova era de sequenciadores é caracterizada pela produção de dados genômicos em grande quantidade, análise massivas em paralelo e com custo reduzido por base sequenciada, comparada ao método de Sanger (LIU et al., 2012).

Os métodos de sequenciamento de nova geração diferem na bioquímica do sequenciamento, engenharia dos aparelhos e como o arranjo é gerado (SHENDURE; JI, 2008; GLENN, 2011). Entre eles, está o sequenciamento de plataforma Illumina (Illumina, San Diego, EUA), considerado de segunda geração, o qual é por síntese em base sólida (*flow cell*), conforme revisado pelos autores citados acima. Esta tecnologia de sequenciamento inclui a preparação do DNA purificado da amostra a ser sequenciada, por um processo chamado fragmentação. Neste processo, as enzimas transposases fragmentam o DNA e adicionam adaptadores.

Após este processo, ocorre uma PCR de ciclos limitados, nos quais são adicionados primers de sequenciamento, índices e oligos que auxiliam na ligação à base sólida. O DNA que foi previamente amplificado é desnaturado e liga-se à base sólida através de pequenos oligonucleotídeos aderidos na mesma, que são também complementares às sequências dos adaptadores. A DNA polimerase sintetiza a fita complementar, a partir da fita simples hibridizada com um dos dois tipos de oligonucleotídeos da base sólida, e a fita original posteriormente é retirada. A fita remanescente é clonalmente amplificada por *Bridge PCR*, na qual a fita se dobra e hibridiza com o outro tipo de oligonucleotídeo da base sólida, passando por ciclos de extensão e desnaturação, sempre utilizando os oligos da base sólida. No final são formados vários grupos de sequências iguais de DNA, os *clusters*, nos quais as fitas são linearizadas. As fitas complementares são clivadas e removidas e somente as várias cópias das fitas originais são sequenciadas em massa e em paralelo, a partir da adição dos primers de sequenciamento em ambas as extremidades. Cada ciclo de sequenciamento é feito por síntese, na base sólida, através da DNA polimerase, que utiliza nucleotídeos terminadores marcados com diferentes fluoróforos para as bases nitrogenadas. A leitura das bases é feita pela análise sequencial das imagens capturadas de acordo com a quantidade de ciclos escolhida (de 50 até no máximo 300 vezes). Em cada ciclo de sequenciamento há a adição de nucleotídeo marcado, lavagem de nucleotídeos não incorporados, leitura da nova incorporação e remoção da fluorescência, resultando nas leituras ou *reads*.

A montagem completa ou parcial do genoma consiste em um conjunto de procedimentos para organizar um grande número de *reads*, com o objetivo de representar a sequência de DNA que compõe cada cromossomo da espécie estudada. Para a montagem das sequências, a estratégia a partir do grafo de Bruijn é a mais usada quando há uma grande quantidade de *reads* curtas (ZERBINO; BIRNEY, 2008). As *words*, que são

subsequências de bases de tamanho k , também conhecidas como k -mer (CHOR et al., 2009), formam as *reads*. Os gráficos de Bruijin funcionam selecionando uma *word* inicial de tamanho k e, a partir dela, procurar uma das quatro *words* vizinhas potenciais, com uma sobreposição de tamanho $k-1$. *Words* adjacentes se sobrepõem por $k-1$ nucleotídeos, que são alinhadas e formam uma sequência linear consenso, os *contigs*, correspondentes a uma mesma região do DNA, sem ser redundante a outra sequência (ZERBINO; BIRNEY, 2008).

2.3 MARCADORES MICROSSATÉLITES

A magnitude da variabilidade detectada pelos marcadores moleculares é muito maior do que a revelada pelos marcadores clássicos (OLIVEIRA et al., 2007). Com o avanço das técnicas de biologia molecular, o DNA passou a ser o ponto central no desenvolvimento de marcadores. Com o advento da PCR (*Polymerase Chain Reaction*), mais tipos de marcadores moleculares foram desenvolvidos. A sequência de DNA de interesse passou a ser amplificada a partir de iniciadores (primers) complementares às regiões flankeadoras, e qualquer região do genoma passou a ser amplificada e analisada em vários indivíduos. Marcadores microssatélites, assim como RAPDs, ISSRs e AFLPs, começaram a ser altamente utilizados na caracterização da diversidade genética (SCHLÖTTERER, 2004).

Os marcadores microssatélites ou SSR (Simple Sequence Repeats) estão entre os mais variáveis tipos de sequência de DNA no genoma, sendo que a variação ocorre principalmente em tamanho, pela alteração no número de repetições (ELLEGREN, 2004). Marcadores SSR apresentam vantagens sobre os demais marcadores moleculares, porque são codominantes e facilmente reproduzíveis, além de serem abundantes e uniformemente dispersos no genoma (VARSHNEY; GRANER; SORRELLS, 2005). Essas

características fazem desses marcadores moleculares ferramentas eficientes para mapeamento genético, estudos de ligação, identificação de genótipos, proteção de variedades, avaliação de pureza de sementes, utilização e conservação de germoplasma, estudos de diversidade, análise gênica e de locos quantitativos, análise de pedigree, seleção assistida por marcadores e análise de bibliotecas para clonagem de genes.

Microssatélites são pequenas sequências de DNA, que ocorrem na forma de unidades repetitivas em tandem. Mono, di, tri e tetranucleotideos são as unidades de repetição mais comuns, mas penta e hexanucleotideos são também classificados como microssatélites (ELLEGREN, 2004). São encontrados em alta frequência nos genomas da maior parte dos táxons (SELKOE; TONNEN, 2006). Os marcadores microssatélites são amplificados por PCR e têm sido os marcadores mais utilizados devido a sua simplicidade técnica, rapidez na análise, grande poder de resolução e altos níveis de polimorfismo (RUFINO, 2005). O alto polimorfismo dos microssatélites nucleares possibilita uma discriminação precisa, mesmo entre indivíduos geneticamente relacionados. A provável origem deste alto polimorfismo é pela alta taxa de mutação deste tipo de marcador por eventos de *slippage* ou deslizamento da polimerase durante a replicação do DNA (ELLEGREN, 2004). Por serem altamente polimórficos, os marcadores microssatélites são extensivamente utilizados nas análises de paternidade, pela alta diversidade alélica (GILL et al., 2001).

Apesar da comprovada eficiência dos marcadores SSR, são restritos os estudos com marcadores microssatélites para espécies florestais nativas. SSR desenvolvidos para uma espécie são frequentemente testados para avaliar a transferibilidade para outra espécie próxima (ZUCCHI et al., 2003). A identificação de regiões microssatélites demanda conhecimento das sequências do genoma (MORRIS, 2016). Identificação das regiões microssatélites é um processo

laborioso e demorado (BLAIR et al., 2009; GUICHOUX et al., 2011), porque tradicionalmente os locos microssatélites eram isolados de bibliotecas genômicas, selecionadas para pequenos insertos do genoma da espécie desejada, transformados e os clones eram rastreados pela hibridização com as sondas contendo as repetições (ZANE; BARGELLONI; PATARNELLO, 2002).

Porém, essa barreira foi derrubada uma vez que com o desenvolvimento do sequenciamento de nova geração, esta etapa passou a demandar uma fração muito menor de tempo (3.100 a 100.000 vezes mais megabases (Mb) identificadas por corrida) e de custos (de 14.000 a 43.000 vezes mais barato por Mb) quando comparada a técnicas tradicionais (ZALAPA et al., 2012). O custo por megabase varia de \$1,00 (Plataforma Polonator e HelisScope) a \$2,00 (Plataformas Solexa) referente ao genoma humano (SHENDURE; JI, 2008). Com a emergência das novas tecnologias de sequenciamento, a identificação de microssatélites tornou-se mais barata, rápida e com maior alcance (GUICHOUX et al., 2011). As PCRs *multiplex* e a utilização de eletroforese capilar que diferencia os *amplicons*, não somente pelo tamanho, mas também pela marcação fluorescente, aumentam o número de locos analisados nas reações *multiplex* (BUTLER; RUITBERG; VALLONE, 2001).

2.4 MICROSSATÉLITES CLOROPLASTIDIAS

Genoma cloroplastidial (cpDNA) difere do nuclear por ser haploide, pela natureza de herança uniparental e não recombinante (NOCK et al, 2011). É uma ferramenta útil para estudos evolutivos e ecológicos por avaliar padrões uniparentais de fluxo gênico dentro e entre populações, para distinção de táxons, construção de linhagens e em barcoding (PROVAN; POWELL; HOLLINGSWORTH, 2001; MCPHERSON et al., 2013; WHEELER et al., 2014).

Em plantas, a detecção de polimorfismo nas populações no genoma cloroplastídial é difícil pelo baixo nível de substituições e baixa taxa de mutação (MCPHERSON et al., 2013). Pela baixa taxa de mutação, os marcadores microssatélites cloroplastídias (cpSSR) representam marcadores em potencial para contornar este problema e têm mostrado bastante variação intraespecífica (PROVAN; POWELL; HOLLINGSWORTH, 2001). Com o avanço das técnicas de sequenciamento, o número de sequências aumentou em quantidade e tamanho, e hoje já existem muitos genomas nucleares e cloroplastídias sequenciados por completo. Por exemplo, McPherson et al. (2013) realizou o sequenciamento do genoma de cloroplasto a partir do DNA total para a espécie *Toona ciliata* M.Roem., pela plataforma Illumina.

Análise de DNA total, ao invés de amostra com DNA de cloroplasto isolado, é um método muito mais simples e utilizado nas rotinas de aplicação (MCPHERSON et al., 2013). Os marcadores microssatélites nucleares são excelentes para a maior parte de análises genéticas, eles não são adequados para estudos de filogenética e filogeografia devido a altas taxas de substituição (PARK et al., 2016). Marcadores microssatélites cloroplastídias são utilizados em estudos de genética de populações por conseguir diferenciar grupos que divergiram recentemente, como DNA barcode, por permitir identificação de grupos genéticos ou espécies, e no estudo de hibridização (WHEELER et al., 2014). A herança uniparental permite a detecção da direção de hibridização. O cpDNA frequentemente apresenta herança na maior parte materna em angiospermas (POWELL et al., 1995). Em espécies híbridas, com diferentes haplótipos de cloroplasto maternalmente herdados, os haplótipos de cloroplasto da F₁ indicam quem é a matriz materna (WU; CAMPBELL, 2005).

Os cpSSRs são adequados para detecção de gargalos em populações naturais e para avaliar a diversidade citoplasmática e variação genética que existe em material de melhoramento de

plantas (FJELLHEIM et al., 2006). Em estudos de parentesco, os cpSSRs funcionam como marcadores parentais paternos ou maternos, atuando como medidas diretas de fluxo de pólen ou semente. Os locos, por serem não recombinantes, fornecem perfis que correspondem ao parental e não é necessário inferir a fase gamética dos locos heterozigotos como no caso dos microssatélites nucleares. Por outro lado, os cpSSRs fornecem dados de apenas um parental, e somente eles sozinhos não geram altas probabilidades de exclusão de parentesco suficientes, sendo necessário também utilizar dados de microssatélites nucleares (PROVAN; POWELL; HOLLINGSWORTH, 2001).

Os cpSSR são distribuídos principalmente em regiões não codificantes, o que aumenta as chance de maior variação do que em regiões codificantes (POWELL et al., 1995). Além disso, os marcadores cpSSR são aplicáveis na amplificação de regiões homólogas em táxons relacionados (DIEKMANN; HODKINSON; BARTH 2012). Apesar do aumento no número de estudos com cpSSRs, na maior parte são de plantas de grande importância econômica, o potencial dos cpSSRs na compreensão de processos ecológicos e evolutivos de muitas espécies ainda não foi explorado (EBERT; PEAKALL, 2009).

3 MATERIAL E MÉTODOS

O presente projeto foi conduzido em uma colaboração do Laboratório DNA UDESC da Universidade do Estado de Santa Catarina – UDESC Lages, Núcleo de Pesquisas em Florestas Tropicais – NPFT da Universidade Federal de Santa Catarina e do Laboratório de Fixação Biológica de Nitrogênio (INCT-FBN) da Universidade Federal do Paraná, para realização do sequenciamento de segunda geração Illumina - MiSeq. Após a geração dos dados brutos de sequenciamento, o processamento das amostras, genotipagem, análise *in silico* para identificação dos primers e validação foram realizadas no Laboratório DNA UDESC, da UDESC Lages.

3.1 EXTRAÇÃO DE DNA GENÔMICO E AVALIAÇÃO DE CONCENTRAÇÃO E QUALIDADE

Todas as 444 amostras avaliadas seguiram o mesmo protocolo para obtenção de DNA. Folhas congeladas de bracatinga foram maceradas com esferas de aço no homogeneizador Precellys® Evolution (Bertin Technologies, Montigny-le-Bretonneux, França). A partir do macerado, a extração de DNA prosseguiu com o método de CTAB, adaptado de Doyle e Doyle (1990). O DNA extraído foi analisado no aparelho NanoDrop 1000® (Thermo Fisher Scientific, Wilmington, EUA) em relação a concentração e razões de absorbância 260/280 e 260/230.

3.2 SEQUENCIAMENTO DE SEGUNDA GERAÇÃO

Uma amostra de DNA genômico foi extraída de uma planta coletada no campus da UDESC Lages. A exsicata foi depositada no LUSC - Herbário Lages da Universidade do Estado de Santa Catarina, com o voucher nº 8822. O DNA extraído foi utilizado para o sequenciamento pela plataforma

Illumina, foi novamente mensurado em relação a concentração aparelho Qubit (Thermo Fisher Scientific, Waltham, EUA) e normalizado para 50 ng/ μ L conforme protocolo descrito no manual para o kit NEXTERA® DNA (Illumina, San Diego, EUA) que prepara o DNA genômico para formação de bibliotecas para sequenciamento na plataforma Illumina. O DNA extraído foi submetido a uma reação de tagmentação, no qual enzimas transposases o fragmentam e adicionam adaptadores que são posteriormente usados para amplificação pela reação em cadeia da polimerase (PCR). O DNA tagmentado foi purificado pelo kit de purificação Zymo™ Purification Kit (Zymo Research, Irvine, EUA) para retirada das enzimas da tagmentação. O DNA purificado foi submetido a PCR com ciclos limitados, que adicionaram primers de sequenciamento em ambas as extremidades dos fragmentos de DNA, par de índices N704 (TCCTGAGC) e N501 (TAGATCGC), para formação das bibliotecas e posteriormente geração dos clusters para o sequenciamento e também oligos homólogos aos oligos P5 e P7 da base sólida. Depois da amplificação, a biblioteca de DNA formada foi purificada com Agencourt AMPure XP beads (Beckman Coulter, Brea, EUA). O sequenciamento foi realizado com o kit MiSeq v3 600-cycle (Illumina, San Diego, EUA) para o sequenciamento *paired-end* de 300 bp no sequenciador MiSeq da Illumina.

3.3 QUALIDADE DOS DADOS, MONTAGEM DAS SEQUÊNCIAS E IDENTIFICAÇÃO DE MICROSSATÉLITES

A qualidade dos dados brutos gerados pelo sequenciamento foi avaliada utilizando o programa FastQC (ANDREWS, 2010). Após a aprovação da qualidade das sequências, os dados de sequenciamento foram inseridos no programa CLC Genomics Workbench 7.5 (CLC Bio, Aarhus, Dinamarca) e foi realizada a trimagem, que remove os

adaptadores das sequências, leituras com tamanho inferior a 50 pares de bases (pb) e de baixa qualidade, além de unir diferentes leituras correspondentes a mesma região do DNA, não redundantes a outra sequência, formando os contigs.

Os contigs formados com cobertura superior a 15x foram exportados para o programa SSR Locator (MAIA et al., 2008) que localiza locos microssatélites e desenha primers que flanqueiam a região de interesse, conforme algoritmo baseado no Programa Primer 3 (ROZEN; SKALETSKY, 1999). A identificação dos SSR seguiu os critérios de Castoe et al. (2012), em que cada tipo de unidade de repetição do microssatélite deve ter no mínimo três repetições do motivo e ao menos 12 pb. A seleção dos primers seguiu os critérios de Dieffenbach, Lowe e Dveksler (1993): tamanho do primer de 18 a 22 pb, com ótimo de 20 pb; tamanho do amplicon de 100 a 550 pb; temperatura de anelamento de 50°C a 60°C, com ótimo de 55°C; ao menos 100 pb de distância entre microssatélites e conteúdo mínimo de 40% de CG, com o ótimo 50% .

3.4 AVALIAÇÃO DA QUALIDADE DOS PRIMERS

A qualidade dos primers foi avaliada *in silico* utilizando o programa Gene Runner (Hastings Software Inc., Hastings, EUA; www.generunner.net) e a ferramenta de análise de oligos do site da IDT DNA (IDT DNA Tech., Coralville, EUA; www.idtdna.com). A aprovação dos pares de primers consistiu na diferença igual ou inferior a 1°C entre as temperaturas de melting dos primers Forward e Reverse; ser um microssatélite perfeito, não composto e não monômero; ausência de palíndromo e $\Delta G > -5$ kcal/mole, conforme instruções da Premier Biosoft (PREMIER Biosoft, Palo Alto, EUA; www.premierbiosoft.com/tech_notes/PCR_Primer_Design.html) nas avaliações de ambos os programas para a presença de homodímeros, heterodímeros e hairpins. Os critérios de

identificação dos locos microssatélites, desenho dos primers e de qualidade dos primers foram seguidos para favorecer a montagem de painéis multiplex, nos quais mais de um marcador é avaliado numa mesma reação de PCR e eletroforese, e também para favorecer a amplificação na reação de PCR.

3.5 VALIDAÇÃO DOS MARCADORES

3.5.1 Amostragem

A população considerada compreendeu a amostragem de 444 indivíduos, sendo 33 adultos e 411 progênies. Estes 33 adultos foram amostrados do campus da Universidade do Estado de Santa Catarina (UDESC) de Lages SC, Centro de Ciências Agroveterinárias (CAV), localizada em área urbana. Estas árvores adultas são provenientes de sementes de três origens do Estado de Santa Catarina (Urupema, Atalanta e Bocaina do Sul), sementes que foram misturadas em um lote para dar origem às mudas que foram plantadas no campus.

Destes 33 adultos, sem delimitação de qual origem é cada árvore, cinco foram selecionadas aleatoriamente para produzir as progênies. As progênies foram obtidas a partir de 108 sementes de cada uma das cinco matrizes. As sementes para serem semeadas tiveram sua dormência superada em água a 80°C por 20 minutos e depois deixadas em descanso nesta mesma água por 18 horas (BIANCHETTI, 1981). Depois de dois meses de crescimento em tubetes, folhas das 411 progênies foram coletadas para análise, assim como as folhas das 33 adultas para a validação dos microssatélites. As amostras de progênies sofreram perdas, algumas sementes não germinaram, outras mudas foram dizimadas por formigas.

3.5.2 1^a fase de validação

Na primeira fase de validação, um subconjunto de 31 pares de primers aprovados pela análise *in silico* foram sintetizados na escala de 25 nmol e avaliados em amostras de DNA de 10 indivíduos adultos por marcador em reações com 1,5 ng de DNA, 0,1 µM de cada primer, 0,2 mM de DNTP, 1x Tampão, 2,5 mM MgCl₂, 0,5 U de Platinum® Taq DNA Polimerase (Invitrogen, Carlsbad, EUA), água ultrapura em um volume final de 12,5 µl. As amplificações foram realizadas no termociclador Veriti (Applied Biosystems, Foster City, EUA) de acordo com o seguinte programa: desnaturação inicial a 95°C por 2 min, seguidas de 35 ciclos de 95°C por 45 s, 54°C por 1 min e 72°C por 2 min, com extensão final a 72°C por 5 min. Os produtos amplificados foram submetidos a eletroforese horizontal em gel de agarose 2% , a 75 V com tampão TBE 1X (Tris-borato 0,09 M e EDTA 0,002 M) e corados com GelRed 1x (Biotium Inc., Hayward, EUA). Para todas as eletroforeses foi utilizado um marcador de peso molecular de 50 bp (Invitrogen, Carlsbad, EUA). Todos os primers que exibiram amplicons na região esperada foram aprovados para a segunda fase de validação.

3.5.3 2^a fase de validação

Um subconjunto de 11 pares de primers, dos 31 testados, foi sintetizado com marcação de fluoróforo no primer *forward*. Estes 11 pares de primers foram analisados pelo programa Autodimer (VALLONE; BUTLER, 2004) e Multiplex Manager (HOLLELEY; GEERTS, 2009) para formação de painéis multiplex. Os painéis foram testados comparando os resultados das genotipagens pela eletroforese capilar, quando os primers eram avaliados em conjunto e separadamente.

As reações de PCR continham 0,5 a 1 ng de DNA, 0,2 mM de DNTP, 1x Tampão, 2,5 mM Cloreto de Magnésio, 1 U de Platinum® Taq DNA Polimerase (Invitrogen, Carlsbad, EUA), concentração de primers que variou de 0,05 a 0,2 µM. As amplificações também foram realizadas no termociclador Veriti (Applied Biosystems, Foster City, EUA) de acordo com o seguinte programa: desnaturação inicial a 95°C por 10 minutos, seguidas de 25 ciclos de 95°C por 45 s, temperatura de anelamento de 55 a 56°C por 1 min e 72°C por 2 min, com extensão final a 72°C por 30 min. Para os marcadores Msc160 e Msc003 foram adicionados 10% de DMSO (Sigma Aldrich, Saint Louis, EUA). Em todas as reações de PCR foram inclusos controles negativos e positivos (amostra sequenciada).

A genotipagem foi realizada por eletroforese capilar, através da leitura da fluorescência detectada pelo Analisador Genético ABI3130 (Applied Biosystems, Foster City, EUA). Para genotipagem de cada amostra foi feita uma reação com 1,0 µL de produto da PCR, 0,3 µL de GeneScan600 LIZ® (Applied Biosystems, Foster City, EUA) e 8,7 µL de formamida Hi-Di® (Applied Biosystems, Foster City, EUA). A designação dos alelos foi determinada pela interpretação dos picos de fluorescência dos eletroferogramas gerados no software Gene Mapper ID-X versão 3.2 (Applied Biosystems, Foster City, EUA).

3.6 ANÁLISE DA DIVERSIDADE GENÉTICA

A genotipagem e análise de organismos poliploides é trabalhosa pela dosagem variável dos alelos. Considerando tetraploidia (DAHMER et al., 2013) para um indivíduo, em um marcador nuclear, ao aparecer apenas um pico no eletroferograma, foram consideradas quatro cópias do alelo (homozigoto), quatro picos foi considerado heterozigoto completo. Para dois picos e três picos, foram considerados

apenas os picos identificados, sem levar em consideração dosagem.

Todos os marcadores que apresentaram ao menos um alelo em comum entre matriz e progênie nos indivíduos e apresentaram resultados iguais entre as repetições das mesmas amostras foram analisados em relação à estatística descritiva e de diversidade. O Programa AUTOTET (THRALL; YOUNG, 2000), para organismos autotetraploides, foi utilizado para identificar o número de alelos, riqueza alélica, riqueza alélica dentro de indivíduos, riqueza genotípica, heterozigose observada, heterozigose esperada assumindo segregação cromossômica e também segregação cromatídica, considerando dupla redução máxima ($\alpha=1/7$), coeficiente de fixação (F) calculado como $1-(H_o/H_e)$, em segregação cromossômica e cromatídica e o teste de aderência ao χ^2 em relação ao equilíbrio de Hardy Weinberg em relação ao marcador nuclear que foi polimórfico. A dupla redução acontece em indivíduos autotetraploides, quando há segregação cromatídica e duas cromátides irmãs vão para a mesma célula gamética. O valor de $\alpha=1/7$ é o valor máximo de dupla redução; considerando uma cromátide, somente uma das sete cromátides restantes é a cromátide irmã (WRICKE; WEBER, 1986). Para os marcadores microssatélites cloroplastidiais que foram polimórficos, a frequência alélica, a frequência haplotípica e diversidade haplotípica foram calculados pelo programa GenAlEx 6.5 (PEAKALL; SMOUSE, 2012).

4 RESULTADOS

4.1 SEQUENCIAMENTO DE SEGUNDA GERAÇÃO

Os dados brutos de sequenciamento continham os adaptadores na sequência e algumas leituras de baixa qualidade. Para remoção destes adaptadores e leituras de baixa qualidade (leituras com menos de 50 pb), foi realizada a trimagem que gerou 1.431.034 leituras, com 237.704.384 bases nucleotídicas lidas e tamanho médio de leitura de 166 pb. Os dados trimados foram submetidos à análise de qualidade no programa FastQC, que gerou um quality value (QV) de 34. Os dados trimados foram utilizados para a montagem dos contigs. Ao todo 42.546 contigs foram obtidos, com tamanho médio de 401 pb e máximo de 59.376 pb. O valor de N50 foi 426 pb, o que significa que 50% dos dados montados têm o tamanho do contig de 426 pb ou maior.

4.2 IDENTIFICAÇÃO DOS LOCOS MICROSSATÉLITES E DESENHO DE PRIMERS

Os contigs formados foram exportados no formato FASTA para o programa SSR Locator (MAIA et al., 2008) primeiramente para a seleção de SSRs. Ao todo foram encontrados 290 locos microssatélites. Deste total, 252 locos foram considerados com repetições de estrutura perfeita com um loco adjacente a mais de 100 pb de distância downstream ou upstream, 16 com o motivo composto (totalizando 36 locos) e uma repetição de estrutura imperfeita, totalizando dois locos. A Tabela 1 mostra as quantidades de locos microssatélite, de acordo com sua unidade de repetição.

Tabela 1: Locos microssatélites separados pelo tamanho do motivo

Motivo	Mono-	Di-	Tri-	Tetra-	Penta-	Hexa-	Total
Quant.	51 17,6%	55 19,0%	94 32,4%	67 23,1%	18 6,2%	5 1,7%	290

Fonte: produção do próprio autor

Com os critérios relacionados para seleção dos primers no tópico material e métodos pela análise *in silico*, foram obtidos 211 pares de primers. Com a eliminação dos SSRs compostos e monômeros e dos reprovados pela análise *in silico*, 65 pares de primers permaneceram aptos para análise *in vitro* (Tabela 2) e, dentre estes, um subgrupo de 31 pares foi sintetizado sem marcação para validação por PCR e eletroforese em gel de agarose (Tabela 2).

Tabela 2: Pares de primers aprovados para análise de *M. scabrella* (Continua)

Loco	Sequência do Primer	Tam (pb)	Motivo
Msc001 *	F: GGCATGTAACGAATTCTTC R: ATCCTTGAGACGACTTA	519	(CCCG) ₃
Msc003 */**	F: CAATCCGTGTTATCTAGGC R: CTAGACCCTGCTCGTATG	220	(GCC) ₆
Msc009 */**	F: GAGTAAAGGGCCTGATAAA R: TACGTTGGTGTGAGA	521	(ATATA G) ₃
Msc017	F: CAGTTAGCTATCGGAGA R: ACCCAGTAGATCCAGTTGA	147	(TA) ₈
Msc020 *	F: TGCCAACAGAACCTTAGGAT R: AATGATGTCATGAAGAGGC	420	(CAAT) ₃
Msc022	F: ACTTGGTCCAGAGCATCTA R: ATCTTGTGAGCAAACATGC	391	(TA) ₇

Tabela 2: Pares de primers aprovados para análise de *M. scabrella* (Continuação)

Locos	Sequência do Primer	Tam (pb)	Motivo
Msc023*	F: CATGTTGCTCACAGATA R: GAATGATGACTTGGGTG	505	(AAAT) ₃
Msc030	F: CAGAACGTCGCTATCAAT R: TTAACGGCTAGCATCTC	504	(AAG) ₄
Msc032	F: AACTTAGCCTGAACCTCTCC R: AGGCTGGCAGAGACTAGAG	415	(TA) ₆
Msc041 */**	F: CTTAGATCAATCTCAACCCG R: AGAATTATCATGGAAGCG	426	(ATTA) ₃
Msc045 */**	F: ATGATACGTAAGCAAGGCAT R: AGTTCAAGTCACACACTCCC	286	(ATTAT) ₃
Msc063*	F: ATGACACTGATTCCCTCCAG R: CCAATATGGGAGATCAAAGA	269	(GAA) ₆
Msc064	F: AAATGACTATGGCTCCTGA R: CAAGCTCTCCTAATCCAAG	288	(AC) ₆
Msc066*	F: CAGGAATTCAACAAACCATCT R: TATTTGAGAATCCTTGAGGG	319	(AAC) ₅
Msc069	F: GGAATGATGACTATCGAGGA R: GTAAGCTCCAAGAACCACTG	239	(TA) ₆
Msc072 */**	F: CGCTCAAGCATCTCTCTAT R: TGTCGTTGTGTCTTGTGT	436	(AACT) ₄
Msc075	F: TACCCATCCTGTATATTGCC R: TTTCCATAGAGCTTCTTGC	507	(TTA) ₄
Msc080*	F: CTATGAAAGCTGGGTATGG R: TATCTACAACCACCACTCC	136	(GGT) ₆
Msc098 */**	F: CGGGAATAGAAGAGACTTCA R: CTTTGTCAATCTTCACCC	117	(GTTG) ₃
Msc099*	F: ATGAAGTGTGATTGCTGGT R: CGAATCGAACCTAATTAAC	504	(GAA) ₄

Tabela 2: Pares de primers aprovados para análise de *M. scabrella* (Continuação)

Locos	Sequência do Primer	Tam (pb)	Motivo
Msc100	F: CCATCTTGGAAAGTAGGTTG R: TCAACCTTAGGAGTGGAAAGA	133	(TCT) ₄
Msc102	F: CTACCTGATTGTGTGTGACG R: GTAGCTCAATCGCCATAATC	302	(TA) ₉
Msc105	F: AGTATATGCGACGTATGGCT R: CCTCAATGTTGGAGGAGTAA	476	(TA) ₇
Msc106	F: ACAACCAGAGGTTCACAGAC R: TTGAGAAAGAAAATCCTGGAC	397	(TA) ₆
Msc108	F: CACATGTTGCTCACGTTTAC R: TCCCATGTGTTCACACTAAA	520	(CT) ₈
Msc120 *	F: TGATGAGTAAGTATGCGTCG R: GAAATTCTTCTCCCAGGACT	395	(CCT) ₅
Msc121 *	F: TGGTAAAGAACAGGGAGAAATGA R: AAATTCTGTCCCTGTTGTTG	426	(TAAA) ₃
Msc124 */**	F: TCTTGCTCAGACCGATAAAT R: AGGATGGGTAGATGTCACAG	292	(ATCA) ₃
Msc127 *	F: GTCCGAGTTTCACCAATTAA R: TGCACAAACTTAATGGAGAC	235	(TTTA) ₃
Msc138	F: CAACGGATCTATTCTAACGG R: GAGCATTGAGAGTGGAAAGAG	174	(GAT) ₄
Msc151	F: AATATGGGCACATCAAAGTC R: TCTCCTCCTTGTATGCCCAA	131	(CAT) ₄
Msc158	F: ATTGTAGACGACGGTGAATC R: CATCCCTAATTCCCTCAATCA	496	(AT) ₉
Msc159*	F: TCCGCTGTTACTCTGAGAT R: GTGTTCCCTATCTGCTTTCG	200	(TGCTT) ₃
Msc160 */**	F: ATCGTATGAGCTCGCATC R: TCTATACAGATTGCCATCCC	101	(TCT) ₅
Msc161 */**	F: AGCTCTCTTCTCGTCCTTT R: GCCATATTCAAACGGATCTA	112	(TCT) ₅

Tabela 2: Pares de primers aprovados para análise de *M. scabrella* (Continuação)

Locos	Sequência do Primer	Tam (pb)	Motivo
Msc165	F: CCTCCAACCCTCTTCTAACT R: CACAATACGAGATGGTTCAA	247	(TG) ₇
Msc173	F: ACGACATTCATTCCTGAG R: CTCTCCTCAACCTCACTCTG	171	(TGA) ₄
Msc174*	F: CAGAGTGAGGTTGAGGAGAG R: CCCAATTAAAGCTGCAATATC	467	(ATTT) ₃
Msc175	F: TCATATGTCTTGAGCATCCA R: ACGAGACAAAGAGACCAAGAA	256	(TTG) ₄
Msc179	F: TCATCTCATCTCACCACTGA R: TTCAACTCAAGTGCATCAAG	194	(TGA) ₄
Msc184	F: CCAAGATAATTCCCTCAGACT R: CATCCTCCTTATGATTGAG	146	(AAG) ₄
Msc187	F: GAAAGGTACCCAGAGGCTAT R: ACCTGATTGTTGTGTGGTT	130	(AT) ₉
Msc193	F: GACCGGAGGGAGTATTATT R: GGAATTGCGTTACATGAGAT	137	(AT) ₈
Msc196	F: TAGTTCTGGAAGCTATGGGA R: TGCCCTTGAATAGAAACTGT	447	(GCA) ₄
Msc198	F: AGATGCAATATTAGGGCTCA R: GATCTCCTTCTACACCCCACA	190	(CCA) ₄
Msc199*	F: CACAATTCCACTCAATGTCA R: ATTGTGAAATGGTTAGGTGC	208	(CAAT) ₃
Msc200	F: CACATCGGTTCTAACGACAT R: AGGGAGAGTGAGGATAAAGG	315	(ATT) ₄
Msc204	F: ATCAGCTCGAATTAGGGTT R: TCAAAGGGATGATGATTAGG	121	(TTC) ₄
Msc206	F: ATGAGTGGGTTAGCAGAGTG R: TGAATGCAAGTCTGTGTGAT	189	(AT) ₈

Tabela 2: Pares de primers aprovados para análise de *M. scabrella* (Continuação)

Locos	Sequência do Primer	Tam (pb)	Motivo
Msc217 */**	F: TGAGTCGGAGCTATCATCTT R: AGAGGTCCAGAGATAAAGGG	161	(CTT) ₅
Msc218*	F: AAAGAAACTCAAGGTCCCTC R: TCCTCATCATCCTCAGACTC	179	(AGA) ₅
Msc219*	F: GATTATGGCTTGCTTCTGTC R: CCAATGTTACTCAAGCCAAT	175	(GCTT) ₃
Msc220	F: TCTAACTCTCTGCCTCAGC R: AAGAAGATTGTGCTGTCGAT	427	(TTC) ₄
Msc225	F: ACAAACCAAGAACACCAAAC R: GATGACTTAGAACCGGGATCA	408	(AT) ₆
Msc229 */**	F: GAGAAAGGATATGATGCAGG R: GTTGCTCGGAAATTATAAGG	140	(AATA) ₄
Msc230	F: CATATGTCAAATGCCTTCC R: ATAAATTGACTGATGGTGC	151	(TA) ₇
Msc231*	F: GTGGATTGTTAGGCATGAAT R: AATGCAAGAGTGATATGGGT	504	(CAAAA) ₃
Msc240	F: GAAATGTAGTTGGTTGACCG R: ATTCTCTTCTCCTTCGTCC	239	(CCT) ₄
Msc248*	F: GTGATTGGGATATTGAGGAA R: GGTCCCTGTCTACAATCAGA	166	(CGGC)3
Msc253*	F: TGTTGAGAGTGTGCATTCAT R: GTTGTATTGATGGAAAGGCT	321	(CATTC) ₃
Msc255	F: CAAAGCAAATCAAGTGGC R: CCAGCAGTGAECTACAACAGA	241	(TGT)4
Msc256*	F: GATTGATCTCATTGGACTGG R: TCATTCTTCATCCTTCACC	493	(AAGA)3
Msc259	F: TCCCTGATTGAAATGAAGAC R: GTTGAGAGTTGAGAAGGAAGA	219	(TG)6

Tabela 2: Pares de primers aprovados para análise de *M. scabrella* (Conclusão)

Locos	Sequência do Primer	Tam.(pb)	Motivo
Msc266*	F: GCTGCTTATCTTCCTTGG R: GTGCCTGACACATACCTCTT	219	(GAT) ₅
Msc267	F: AGAAGATCAAGGAAAGGAGC R: ATAAGTTCAATGGGAGCAGA	256	(AAG) ₄

Fonte: produção do próprio autor.

NOTAS: Primers destacados com * foram sintetizados sem marcação para a 1^a fase de validação. Primers destacados com ** foram sintetizados para a 2^a fase de validação. F = primer *forward*. R = primer *reverse*. Tam. (pb) = Tamanho esperado em pares de base.

4.3 1^a FASE DE VALIDAÇÃO – PCR E ELETROFORESE EM GEL DE AGAROSE

A 1^a fase de validação dos marcadores SSR por PCR e eletroforese em gel de agarose triou os 31 pares de primers sintetizados e 12 pares não apresentaram bandas no gel ou apresentaram em regiões não esperadas pelo previsto em sequenciamento. Dos 19 pares de primers aprovados, 11 foram sintetizados com marcação de fluoróforo no primer *forward*.

4.4 2^a FASE DE VALIDAÇÃO – PCR E ELETROFORESE CAPILAR

Os 11 pares de primers marcados foram avaliados nas 411 progêneres e 33 árvores adultas. Painéis multiplex foram montados para otimizar a obtenção de resultados (Tabela 3). Ao todo foram obtidos três painéis multiplex, todos com temperatura de anelamento de 56°C na PCR; dois quadriplex; um com os marcadores Msc098, Msc045, Msc124 e Msc217, outro com os marcadores Msc229, Msc041, Msc009 e Msc072, um bplex com os marcadores Msc161 e Msc003 (Figura 2), e um único marcador que foi analisado sozinho, o Msc160 com

temperatura de anelamento de 55°C na PCR. Todos os marcadores também foram testados em separado para comparar com os resultados dos multiplex e averiguar a consistência dos marcadores em conjunto.

A genotipagem de adultos e progênies gerou o total de 27 alelos. Os marcadores Msc160, Msc098, Msc217, Msc041, Msc072, Msc229 e Msc161 foram monomórficos para toda a amostragem.

Tabela 3: Características dos 11 pares de primers microssatélites e painéis desenvolvidos para *Mimosa scabrella* Benth.

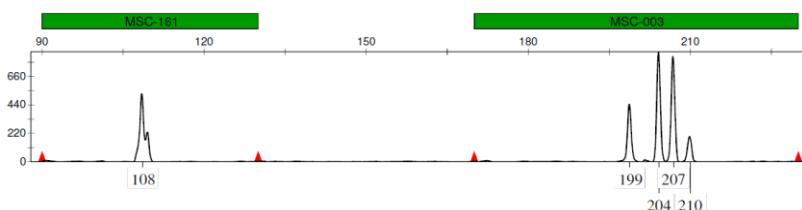
Painel	Loco	Marcação Primer F	A	Tam (pb)	Primer (µM)
1	Msc098	VIC	1	113	0,05
	Msc045	VIC	3	267-283	0,05
	Msc124	FAM	2	286-300	0,05
	Msc217	FAM	1	156	0,05
2	Msc229	PET	1	138	0,11
	Msc041	NED	1	422	0,07
	Msc009	FAM	4	499-514	0,12
	Msc072	VIC	1	430	0,07
3	Msc161	NED	1	108	0,2
	Msc003	NED	11	195-224	0,1
Mono	Msc160	FAM	1	95	0,05

Fonte: produção do próprio autor.

NOTAS: Primer F = primer forward; A (riqueza alélica) = número de alelos em cada loco; Tam. (pb) = Tamanho encontrado dos fragmentos em pares de base ou amplitude dos tamanhos dos fragmentos encontrados para os marcadores polimórficos.

Os marcadores Msc045, Msc124, Msc009 e Msc003 foram polimórficos para adultos e progêneres, com riqueza alélica (A) de 3, 2, 4 e 11 respectivamente, considerando toda população de estudo. No locos Msc003 foram obtidas amostras com um a quatro alelos. Também foi detectado heterozigose para o marcador Msc003, com genótipos totalmente heterozigotos, com 4 alelos diferentes, heterozigotos incompletos, com indivíduos que apresentaram 2 ou 3 alelos diferentes nos eletroferogramas e totalmente homozigotos com a identificação de 1 alelo. Para os heterozigotos incompletos, foram considerados como alelos somente os picos identificados, sem considerar o número de cópias de cada alelo.

Figura 2: Biplex com os marcadores Msc161 e Msc003



Fonte: Produção do próprio autor.

Dos 11 pares de primers analisados quanto ao polimorfismo, os amplicons dos marcadores Msc009, Msc045, Msc124 e Msc041 quando alinhados pela ferramenta BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), tiveram similaridade de 83 a 95% com partes dos genomas completos de cloroplasto de espécies de Fabaceae, inclusive espécies da subfamília Mimosoideae, a mesma da bracatinga. Os marcadores Msc009, Msc045, Msc124 além de terem sido polimórficos, são marcadores microssatélites de cloroplasto, reforçados por terem o mesmo único alelo entre matriz e progêneres. O marcador Msc003 não alinhou com nenhuma região pela ferramenta BLAST, mas foi polimórfico e teve indivíduos que

apresentaram quatro alelos para o marcador, o que corrobora ser um microssatélite nuclear de uma espécie tetraploide. Dados não apresentados de citometria de fluxo também corroboraram a natureza tetraploide da bracatinga. Os demais marcadores (Msc160, Msc098, Msc217, Msc072, Msc229 e Msc161) não tiveram ocorrência de alinhamento no BLAST.

4.5 ANÁLISE DA DIVERSIDADE GENÉTICA

O total de locos polimórficos foi de 36%. As frequências alélicas para os locos polimórficos de cloroplasto se encontram na Tabela 4 e na Tabela 5 as frequências alélicas para o SSR nuclear Msc003. Para o marcador Msc003 os valores de diversidade para os 33 indivíduos adultos estão descritos na Tabela 6 e para as 411 progênies na Tabela 7. O teste de aderência ao χ^2 mostrou que o loco Msc003 não está em Equilíbrio de Hardy Weinberg ($p<0,001$) tanto para adultos quanto para progênies.

Tabela 4: Frequências alélicas dos locos polimórficos de cloroplasto

Alelo	Msc009		Alelo	Msc045		Alelo	Msc124	
	A	P		A	P		A	P
	Frequência			Frequência			Frequência	
499	0,152	0,238	267	0,061	0	286	0,848	0,762
502	0,03	0	274	0,152	0,238	300	0,152	0,238
507	0,03	0	283	0,787	0,762			
514	0,788	0,762						

Fonte: Produção do próprio autor.

NOTAS: A: adultos, P: progênies.

Tabela 5: Frequências alélicas do SSR nuclear Msc003.

Alelo	Msc003	
	A	P
	Frequência	
195	0	0,0012
198	0	0,0024
199	0,0152	0,002
201	0	0,0006
202	0	0,0008
204	0,2071	0,1513
206	0	0,0016
207	0,7525	0,8082
209	0,0101	0,0174
210	0,0152	0,0138
224	0	0,0006

Fonte: Produção do próprio autor.

NOTAS: A: adultos, P: progêneres.

Tabela 6: Estatística de diversidade do SSR nuclear Msc003 para os indivíduos adultos

Loco	A	Ai	G	Ho	Cromossomo (Ce)		Cromátide (Cd)	
					He(Ce)	F(Ce)	He(Cd)	F(Cd)
Msc003**	5	2	4	0,394	0,478	0,175	0,446	0,116

Fonte: produção do próprio autor.

NOTAS: A: riqueza alélica (número de alelos em cada loco), Ai: riqueza alélica dentro de indivíduos (número médio de alelos por indivíduos em um loco), G: riqueza genotípica (número de genótipos com quatro alelos por loco), Ho: heterozigose observada, He(Ce) e F (Ce): heterozigose e coeficiente de fixação esperados em segregação cromossômica, He(Cd) e F (Cd): heterozigose e coeficiente de fixação esperados com nível $\alpha=1/7$ de segregação cromatídica. ** = Desvio significativos do HWE ($p<0,001$).

Tabela 7: Estatística de diversidade do SSR nuclear Msc003 para as progênies.

Loco	A	Ai	G	Ho	Cromossomo (Ce)		Cromátide (Cd)	
					He(Ce)	F(Ce)	He(Cd)	F(Cd)
Msc003**	11	1,786	16	0,309	0,438	0,296	0,409	0,245

Fonte: produção do próprio autor.

NOTAS: A: riqueza alélica (número de alelos em cada loco), Ai: riqueza alélica dentro de indivíduos (número médio de alelos por indivíduos em um loco), G: riqueza genotípica (número de genótipos com quatro alelos por loco), H_o : heterozigose observada, He(Ce) e F (Ce): heterozigose e coeficiente de fixação esperados em segregação cromossômica, He(Cd) e F (Cd): heterozigose e coeficiente de fixação esperados com nível $\alpha=1/7$ de segregação cromatídica. ** = Desvio significativos do HWE ($p<0,001$).

Para os marcadores microssatélites cloroplastidiais Msc009, Msc045, Msc124, foram encontrados 4 haplótipos diferentes, sendo 2 haplótipos comuns entre adultos e progênies e 2 exclusivos aos adultos. O haplótipo 1 formado pelos alelos 514 pb (Msc009), 283 pb (Msc045) e 286 pb (Msc124) foi o mais comum tanto nos adultos quanto nas progênies, com frequência de 0,788 e 0,762 respectivamente. O haplótipo 2, formado pelos alelos 499 pb (Msc009), 274 pb (Msc045) e 300 pb (Msc124) foi o segundo mais frequente entre adultos e progênies, com frequência de 0,152 e 0,238 respectivamente. O haplótipo 3 formado pelos alelos 507 pb (Msc009), 267 pb (Msc045) e 286 pb (Msc124) e o haplótipo 4 pelos alelos 502 pb (Msc009), 267 pb (Msc045) e 286 pb (Msc124) tiveram uma única ocorrência nos adultos, e ambos apresentaram a frequência de 0,030.

5 DISCUSSÃO

Com o avanço das técnicas de sequenciamento, é cada vez maior o número e o tamanho das leituras realizadas. Mesmo comparando técnicas e aparelhos das mesmas empresas, o número é variado em cada etapa e dependente das técnicas utilizadas em todo o procedimento. Por exemplo, para a plataforma Illumina, para o mesmo aparelho utilizado neste trabalho, MiSeq, existem leituras de 36 pb até 300 pb (SHENDURE; JI, 2008). Para o aparelho HiSeq da mesma plataforma, as leituras são de 100 pb, porém o número de clusters gerados é de até 3 bilhões, sendo que o MiSeq gera até no máximo 25 milhões de clusters com um output de até 15 Gb por corrida (ILUMINA, 2013). Houlston et al. (2015), utilizando o sequenciador 454 da Roche, que tem o limite de tamanho de leitura de 250 pb, obtiveram 83.643 leituras, com tamanho médio de 461 pb para o genoma montado de *Cianthus puniceus* (G. Don) Banks & Sol. ex Lindl., uma outra espécie da família Fabaceae. Para *Brachiaria ruziziensis*, Silva et al. (2013) encontraram 186.764.108 leituras, com tamanho médio de 76 pb, pelo sequenciamento single-end da Illumina. O presente trabalho gerou 1.431.034 leituras para bracatinga com tamanho médio de 166 pb para o genoma parcial montado. Até o momento, são os primeiros dados de microssatélites de bracatinga. Somente um trabalho (COSTA et al., 2012) testou a transferibilidade dos primers originalmente desenvolvidos para *Parapiptadenia rigida* (Benth.) Brenan. (Fabaceae-Mimosoideae) para *M. scabrella* e conseguiu amplificação em somente dois locos.

Para favorecer a identificação de mais regiões microssatélites, alguns autores seguem o protocolo de enriquecimento das regiões alvo porque aumenta a quantidade de DNA a ser sequenciada com as variantes de interesse. Morris et al. (2016) realizaram o enriquecimento de biblioteca genômica de *Astragalus bibullatus* Barneby & E. L. Bridges

para posteriormente sequenciar pela plataforma Illumina usando o aparelho MiSeq e encontraram mais de 3.000 pares de primers possíveis. Os métodos iniciais de desenvolvimento de marcadores microssatélites também eram baseados no enriquecimento com motivos geralmente pequenos (dinucleotídeos) de sequências predeterminadas (BILLOTTE et al., 1999). Os clones positivos eram sequenciados, utilizando o sequenciamento de Sanger, para a presença de microssatélites. Para *Butia eriospatha* (Mart. ex Drude), pela abordagem de Billotte et al. (1999), para desenvolvimento de microssatélites, Nazareno, Zucchi e Reis (2011) realizaram o sequenciamento de Sanger para 192 clones positivos e encontraram 30 locos microssatélites, para os quais todos foram desenhados primers. Utilizando a mesma abordagem citada anteriormente, Klabunde et al. (2014) sequenciaram 192 clones positivos e identificaram 36 locos microssatélites, desenhando 26 pares de primers para *Acca sellowiana* (O. Berg) Burret. Sem realizar enriquecimento prévio, o sequenciamento de segunda geração realizado no presente trabalho, identificou ao todo 290 locos microssatélites enquanto Donkpegan et al. (2015), com enriquecimento e sequenciamento de segunda geração, encontraram 635 locos microssatélites para *Afzelia bipindensis* Harms, outra espécie arbórea tetraploide da família Fabaceae. O enriquecimento prévio aumenta o número de microssatélites identificados, mas também é uma etapa extra a ser realizada, que demanda tempo e custos extras, além de gerar um viés dos tipos de motivos microssatélites detectados, reduzindo as classes de SSR esperadas (WEISING, WÖHRMANN; HUETTEL, 2015). No presente estudo, deixar de realizar o enriquecimento prévio, não impossibilitou obter um grande número de SSRs para a bracatinga.

Bell, Murphy e Gardner (2013) realizaram uma triagem inicial dos 47 locos SSR identificados para *Vachellia farnesiana* (L.) Wight & Arn., por PCR e gel de agarose, e encontraram 28 SSRs que amplificaram nos tamanhos

esperados (59,57%) para depois marcar os primers com fluorescência para avaliação de polimorfismo, dos quais 26 microssatélites foram polimórficos. Da mesma forma, no corrente trabalho, dos 31 pares de primers avaliados na primeira fase de avaliação, 19 amplificaram na região esperada (62,29%), e 11 locos foram testados quanto ao polimorfismo, utilizando primers forward sintetizados com fluorescência. Para *Jasminum sambac* (L.) Aiton, da família Oleaceae, Li e Zhang (2015) desenvolveram marcadores microssatélites pelo sequenciamento Illumina, no qual de uma triagem inicial de 100 pares de primers por PCR e eletroforese, somente 31 pares de primers foram avaliados e apenas seis foram polimórficos.

Grande parte dos trabalhos que utilizam marcadores microssatélites desenvolvidos pelo sequenciamento de segunda geração utilizam kits específicos para a montagem de PCR multiplex, melhorando a especificidade e estabilização da ligação dos primers com o DNA, favorecendo a amplificação, mesmo com regiões ricas em CG. No presente trabalho, três painéis multiplex foram montados sem a utilização de kits específicos para a multiplexação, o que não prejudicou a montagem e validação dos painéis, uma vez que as amplificações mostraram-se estáveis, com picos bem definidos, as repetições foram todas consistentes entre todas as matrizes e progênies, sendo sempre identificado ao menos um alelo em comum. Apenas para o painel com os marcadores Msc161 e Msc003, foi adicionado DMSO para melhorar a reação de PCR. O aditivo DMSO não foi utilizado na primeira fase de validação dos primers, sua utilização poderia ter aumentado o número de primers aptos para avaliação de polimorfismo com primers marcados.

Um estudo prévio com *M. scabrella* (DAHMER et al., 2013) revelou alta diversidade genética entre 14 populações de bracatinga avaliadas por marcadores RAPD (Randomly Amplified Polymorphic DNA) e encontrou um número médio de 12,6 fragmentos por primer. Neste presente trabalho, para o

marcador Msc003 ao todo foram encontrados 11 alelos considerando a única população deste estudo. O índice de fixação nas progênies foi maior do que nas árvores maternas considerando apenas o marcador Msc003. Sobierajski, Kageyama e Sebben (2006) realizaram a mesma avaliação por isoenzimas, que também mostrou índices de fixação mais altos nas progênies de bracatinga. Moreira et al. (2011), avaliando locos de aloenzimas para bracatinga, encontraram valores de polimorfismo que variaram de 75 a 100% nas 13 populações avaliadas, enquanto na única população avaliada no presente trabalho, quatro locos foram polimórficos, o que representa 36% de polimorfismo considerando os 11 locos analisados.

Sobierajski, Kageyama e Sebben (2006) indicam um quadro de endogamia biparental para *M. scabrella*, o que pode ser uma das razões que o presente trabalho identificou poucos locos polimórficos nucleares. Também deve ser destacado que a amostragem contemplou apenas uma população, sendo necessário avaliar mais matrizes e progênies de outras localidades. É importante salientar que os marcadores avaliados podem contemplar regiões conservadas do genoma da bracatinga e/ou com pouca variação, uma vez que nenhum dos 11 microssatélites avaliados quanto ao polimorfismo são dinucleotídeos, que são os que mais tendem à variação comparados às outras repetições (KATTI; RANJEKAR; GUPTA, 2001). Os locos microssatélites selecionados neste trabalho, durante o processo de montagem do genoma *de novo*, seguiram vários procedimentos para montagem dos contigs. Quanto maior a cobertura dos contigs, indica que mais vezes essa sequência única foi encontrada nas diversas leituras que compuseram o contig. Esta maior repetição favorece a seleção de regiões mais conservadas.

A sequência dos amplicons de todos os locos foram analisados pelo BLAST e a sequência dos amplicons dos marcadores Msc009, Msc045, Msc124 e Msc041 tiveram similaridade de 83 a 98% com diversas regiões do genoma de

cloroplasto de várias espécies pertencentes à família Fabaceae, inclusive para a mesma subfamília Mimosoideae, como *Prosopis glandulosa* e *Desmanthus illinoensis*. Considerando os quatro marcadores cloroplastidiais avaliados no presente trabalho, três foram polimórficos, com 9 alelos encontrados. Sakaguchi et al. (2015) avaliaram 24 marcadores SSR para *Callitris sulcata* (Parl.) Schltr. e 15 foram polimórficos, dentre eles 1 cloroplastidial e 14 nucleares. Park et al. (2016) avaliaram 41 cpSSRs para *Maianthemum bicolor* (Nakai) Cubey e 10 foram polimórficos, encontrando ao todo 21 alelos.

Poliploidia é sempre um desafio na montagem *de novo* e na montagem dos genomas de plantas, por serem dependentes do tempo do evento de ploidização, e se são auto ou alopóliploides (HAMILTON; BUELL, 2012). Estudar quaisquer espécies para as quais não existem dados já descritos de sequência de DNA é sempre uma barreira. Como os modelos de análise são em sua maior parte baseados em organismos diploides, utilizar as estratégias de organismos diploides para poliploide é um risco que não contempla a enorme dimensão dos poliploides (GRANDKE et al., 2014). Para o genoma da bracatinga não foi possível utilizar nenhuma abordagem de sequenciamento de poliploide citadas anteriormente, uma vez que não há caracterização genotípica dos parentais e não há programa de melhoramento genético que fortaleça a necessidade de conhecimento do seu genoma e das variantes entre as populações. Para as grandes culturas como o milho (SCHNABLE et al., 2009), trigo hexaploide (WINFIELD et al., 2012) e soja (ZHOU et al., 2015) existem inúmeros dados, envolvendo até re-sequenciamento, etc.. Além da necessidade alimentícia, que já justifica serem foco de estudos, o tempo de geração destas culturas é menor do que para espécies arbóreas, o que facilita a realização dos estudos.

Para a montagem das sequências da bracatinga, foram descartadas leituras com menos de 50 pb e no processo de trimagem (remoção dos adaptadores), o quality score de

trimagem foi de 99,99, o que significa que no máximo 0,01% de bases com baixa qualidade são permitidas, limite para não se descartar uma região maior com alta qualidade por uma base isolada de baixa qualidade. Na formação dos contigs, a cobertura mínima foi de 15x. Estes procedimentos de montagem levam em conta a repetibilidade das sequências, e regiões conservadas e/ou com leituras iguais, como nos homozigotos, favorecem o maior número de vezes que são observados os contigs. Leituras do mesmo cluster de sequenciamento quando alinhadas, se existirem algumas variantes, mas com pouca repetibilidade, podem ter sido eliminadas pelos critérios de qualidade na formação dos contigs. A avaliação de polimorfismo dos 11 pares de primers utilizados neste trabalho foi baseada em contigs com cobertura de 18,3x a 219,35x.

Dados de citogenética posicionam a bracatinga como tetraploide (DAHMER et al., 2011) e as evidências moleculares do presente trabalho reforçam a tetraploidia, uma vez que diferentes indivíduos foram identificados com um, dois, três ou quatro alelos diferentes, assim como observado em análises de microsatélites de *Bromelia antiacantha* Bertoloni (ZANELLA et al., 2011). A natureza da poliploidia da bracatinga, se é auto ou alopóliploide, ainda não foi determinada, mas dados de citometria de fluxo não apresentados neste trabalho corroboram a tetraploidia. Ainda não há dados sobre a meiose, mas a mitose revela cromossomos muito parecidos entre si (SEIJO, 2000). Os alelos observados no marcador Msc003 no presente trabalho reforçam a hipótese da autotetraploidia. Dados não publicados pelo Prof. Dr. Saulo Marçal, da Universidade Federal de Juiz de Fora em Minas Gerais, que analisou a mitose de amostras de populações de bracatinga dos estados de Minas Gerais, São Paulo, Rio de Janeiro e Paraná, reforçam que a espécie é tetraploide. Pelas mesmas razões que Seijo (2000), a opinião do Prof. Marçal é que a espécie é autotetraploide, mas ele

reforça a necessidade da análise meiótica para se chegar a uma conclusão definitiva.

A ausência de caracterização citogenética é barreira para o processamento de dados de genética de populações para poliploides porque os programas estatísticos são desenvolvidos com base em teorias e premissas que descrevem o comportamento biológico, e estas premissas devem ser atendidas para que uma análise seja feita. Para o presente trabalho, a análise dos dados estatísticos para o SSR nuclear considerou a *M. scabrella* como autopoliploide utilizando o programa Autotet (THRALL; YOUNG, 2000), exclusivo para autopoliploides.

Autopoliploides são caracterizados por herança polissômica, o que proporciona riqueza alélica e heterozigose maiores do que nas espécies diploides mais próximas (HOKANSON; HANCOCK, 1998). Para dados de *Parapiptadenia rigidula*, uma espécie diploide da família Fabaceae – Mimosoideae (COSTA et al., 2012), os valores de H_o variaram de 0,111 a 0,489 e H_e de 0,106 a 0,707 dependendo do loco, valores superiores aos encontrados no presente trabalho, que obteve H_o de 0,315 e H_e de 0,403 considerando dados de adultos e progêniens juntos para o marcador nuclear polimórfico Msc003. Mas Donkpegan et al. (2015) encontraram para a árvore *Afzelia bipindensis*, uma espécie tetraploide da família Fabaceae – Caesalpinoideae, H_o de 0,22 a 0,79 e H_e de 0,3 a 0,91 e para *Afzelia africana*, uma espécie diploide, H_o variou de 0 a 0,87 e H_e de 0 a 0,87 também. A comparação ao rigor deveria ser feita com espécies mais próximas ainda, mas a falta de dados para o gênero *Mimosa* L. é limitante.

Marcadores microssatélites desenvolvidos para outras espécies tetraploides como *Laurus nobilis*, mostraram sucesso na transferibilidade para *Laurus azorica*, com 15 locos transferidos dos 20 desenvolvidos para *L. nobilis* (ARROYO et al., 2010). Em *Afzelia bipindensis*, a transferibilidade foi

obtida em todos os 11 *locos* microssatélites para duas espécies tetraploidoides *A. pachyloba* e *A. bella*, e em sete *locos* para duas espécies diploides *A. africana* e *A. quanzensis* (DONKPEGAN et al., 2015). Já Palop, Palacios e González-Candelas (2000) obtiveram baixa transferibilidade de microssatélites de *Limonium narbonense*, uma espécie de lavanda tetraploide, para outras espécies do mesmo gênero. Para os marcadores SSR da bracatinga não foram feitas tentativas de transferibilidade para outras espécies próximas, mas a necessidade existe porque não há marcadores microssatélites já descritos para outras espécies do gênero *Mimosa* L.

Muitos problemas ainda devem ser esclarecidos no tratamento de dados para poliploidoides. Existe a dificuldade na identificação dos genótipos, principalmente quando se trata de heterozigotos parciais, mapas de ligação e fases do haplótipo (GRANDKE et al., 2014). Alguns autores se baseiam no trabalho de Esselink, Nybom e Vosman (2004) para estabelecer os genótipos de microssatélites para poliploidoides, levando em consideração a altura dos picos para verificar a dosagem de cada alelo (PALOP-ESTEBAN; SEGARRA-MORAGUES; GONZÁLEZ-CANDELAS, 2011; ADDISALEM et al., 2015; DONKPEGAN et al., 2015). Para o presente trabalho, a metodologia de Esselink, Nybom e Vosman (2004) não foi utilizada, assim como nos trabalhos de Arroyo et al. (2010), Dufresne et al. (2014) e Narayan, Dodd e O'Hara (2015) que consideraram a presença do alelo para formação do genótipo, sem considerar os possíveis números de cópias baseado na altura dos picos. Na análise dos poliploidoides de margaridas australianas (*Rutidosis leptorrhynchoides* F. Muell), a determinação do genótipo tetraploide foi feita comparando-se com amostras diploides por aloenzimas e analisando progênieis de cruzamentos controlados (BROWN; YOUNG, 2000). Com o agrião amarelo (*Rorippa amphibia* e *Rorippa sylvestris*), por análise de microssatélites, foi testado cruzamentos entre duas variedades tetraploidoides para formar híbridos e se estudar como

o padrão de herança influencia na formação dos genótipos (STIFT et al., 2008), o que não é possível para o corrente trabalho por não ter os dois parentais identificados e caracterizados, assim como não é ainda descrita a origem da poliploidização de *M. scabrella*.

As abordagens para tratamento de dados diploides devem ser ampliadas para serem aplicáveis para poliploides. Ainda há métodos que faltam ser desenvolvidos e outros que devem ser melhorados para o tratamento de dados de espécies poliploides (GRANDKE et al., 2014). O programa LD4X (JULIER, 2009), calcula desequilíbrio de ligação, mas não aceita genótipos incompletos, situação encontrada em muitos indivíduos heterozigotos analisados no presente trabalho. Apesar da importância dos tetraploides entre as espécies de plantas, a maior parte dos estudos de genética trabalha com espécies diploides por causa da dificuldade de analisar dados codominantes de microssatélite em tetraploides. Muitas ferramentas estão disponíveis para diploides, mas a pesquisa em poliploides ainda está muito atrás (PUYVELDE; GEERT; TRIEST, 2010).

Muitas das grandes culturas são poliploides e em algumas espécies existem múltiplos níveis de ploidia (GARCIA et al., 2013). A falta de ferramentas para tratar dados de espécies poliploides e o crescimento da geração de dados em qualidade e quantidade reforçam a necessidade de desenvolvimento para o manejo das espécies poliploides (GRANDKE et al., 2014). As tecnologias para sequenciamento de DNA estão cada vez melhores e mais acessíveis, o que possibilita explorar as potencialidades e peculiaridades de todas as espécies. Assim, espécies poliploides que não estão no foco do mercado irão se beneficiar do conhecimento gerado. Apesar de alguns efeitos comuns da poliploidia serem desejados, como o tamanho dos frutos (WU et al., 2012), a poliploidia é uma característica relativamente negligenciada

quando se trata sobre conservação genética (BROWN; YOUNG, 2000), devendo ser desenvolvida.

6 CONCLUSÃO

O sequenciamento de segunda geração possibilitou a identificação de um grande número de *locos* microssatélites para *Mimosa scabrella* Benth., resultando no desenvolvimento de marcadores SSR para a espécie.

Os marcadores SSR avaliados e validados são aptos e úteis para análises e estudos de genética de populações.

Outros marcadores SSR foram identificados, mas ainda precisam ser sintetizados e avaliados quanto ao polimorfismo e consistência entre matrizes e progênies de outras populações distantes para validação.

REFERÊNCIAS

- ABDELKRIM, J. et al. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. **Biotechniques**, v. 46, n. 3, p. 185, 2009.
- ADDISALEM, A. B. et al. Genomic sequencing and microsatellite marker development for *Boswellia papyrifera*, an economically important but threatened tree native to dry tropical forests. **AoB Plants**, v. 7, p. plu086, 2015.
- ANDREWS, S. et al. FastQC: A quality control tool for high throughput sequence data. **Reference Source**, 2010.
- ARROYO, J. M. et al. Isolation and characterization of 20 microsatellite loci for laurel species (*Laurus*, Lauraceae). **American Journal of Botany**, v. 97, n. 5, p. e26-e30, 2010.
- BAGGIO, A. J. et al. Sistema agroflorestal tradicional da bracatinga com culturas agrícolas anuais. **Boletim de Pesquisa Florestal**, Colombo, v. 12, p. 73-82, 1986.
- BAGGIO, A. J.; CARPANEZZI, A. A. **Exploração seletiva do sub-bosque**: uma alternativa para aumentar a rentabilidade dos bracatingais. Colombo: Embrapa Florestas, Circular técnica 28, 1998. 17 p.
- BARTOSZECK, A. C. P. S. et al. Dinâmica da relação hipsométrica em função da idade, do sítio e da densidade inicial de povoamentos de Bracatinga da Região Metropolitana de Curitiba, PR. **Revista Árvore**, v. 28, n. 4, p. 517-533, 2004.
- BELL, K. L.; MURPHY, D. J.; GARDNER, M. G. Isolation, via 454 sequencing, and characterization of microsatellites for

Vachellia farnesiana (Fabaceae: Mimosoideae). **Applications in Plant Sciences**, v. 1, n. 10, 2013.

BERTANI, D. F.; SANTOS, F. A. M. Regeneração de *Psychotria suterella* Müll. Arg. (Rubiaceae) em uma paisagem fragmentada de Floresta Atlântica no sudeste do Brasil. **Revista do Instituto Florestal**, v. 25, p. 193-203, 2013.

BIANCHETTI, A. Comparação de tratamentos para superar a dormência de sementes de bracatinga (*Mimosa scabrella* Bentham). **Boletim de Pesquisa Florestal**, n. 2, p. 57-68, 1981.

BILLOTTE, N. et al. Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. **Fruits**, v. 54, n. 4, p. 277-288, 1999.

BLAIR, M. W. et al. Development of microsatellite markers for common bean (*Phaseolus vulgaris* L.) based on screening of non-enriched, small-insert genomic libraries. **Genome**, v. 52, n. 9, p. 772-782, 2009.

BOAKES, E. H. et al. Extreme contagion in global habitat clearance. **Proceedings of the Royal Society of London B: Biological Sciences**, v. 277, n. 1684, p. 1081-1085, 2010.

BRASIL. Ministério do Meio Ambiente. **Programa Nacional de Conservação e Uso Sustentável do Bioma Cerrado**. Brasília, DF, 2006a. 56 p.

BRASIL. Lei nº 11.428, de 22 de dezembro de 2006b. Dispõe sobre a utilização e proteção da vegetação nativa do Bioma Mata Atlântica, e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 26 de dezembro de 2006.

BRASIL. Lei nº 11.326, de 24 de julho de 2006c. Estabelece as diretrizes para a formulação da Política Nacional da Agricultura Familiar e Empreendimentos Familiares Rurais. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 25 jul. 2006.

BRASIL. Lei nº 12.651, de 25 de maio de 2012. Dispõe sobre a proteção da vegetação nativa. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 28 mai. 2012.

BROWN, A. H. D; YOUNG, A. G. Genetic diversity in tetraploid populations of the endangered daisy *Rutidosis leptorrhynchoides* and implications for its conservation. **Heredity**, v. 85, n. 2, p. 122-129, 2000.

BURKART, A. Leguminosas Mimosoídeas. In: REITZ, R. (Ed.). **Flora Ilustrada Catarinense**. Itajaí: Herbário Barbosa Rodrigues, 1979. 304 p.

BUSO, G. S. C. et al. Marcadores microssatélites em espécies vegetais. **Revista Biotecnologia, Ciência e Desenvolvimento**, v. 30, p. 46, 2003.

BUTLER, J. M.; RUITBERG, C. M.; VALLONE, P. M. Capillary electrophoresis as a tool for optimization of multiplex PCR reactions. **Fresenius journal of Analytical Chemistry**, v. 369, n. 3-4, p. 200-205, 2001. †

CARNEIRO, R. M. et al. Importância da dormência das sementes na regeneração da bracaatinga-*Mimosa scabrella* Benth. **IPEF**, Circular técnica 149, 9 p., 1982.

CARPANEZZI, A. A. et al. **Manual técnico da bracatinga (*Mimosa scabrella* Benth)**. Curitiba: Embrapa - Centro Nacional de Pesquisa de Florestas, 1988.

CARPANEZZI, A. A.; PAGANO, S. N.; BAGGIO, A. J. Banco de sementes de bracatinga em povoamentos do sistema agroflorestal tradicional de cultivo. **Boletim de Pesquisa Florestal**, v. 35, p. 3-19, 1997.

CARVALHO, P. E. R. **Espécies Arbóreas Brasileiras**. 1 ed. Brasília: Embrapa Informação Tecnológica, 2003. 1039 p.

CARVALHO, J. M. F. C; SILVA, M. M. A.; MEDEIROS, M. J. L. **Perda e conservação dos recursos genéticos vegetais**. Campina Grande: Embrapa Algodão, Documentos, v.1, 2009. 19 p.

CASTOE, T. A. et al. Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. **PLoS One**, v. 7, n. 2, p. e30953, 2012.

CATIE- CENTRO AGRONÔMICO TROPICAL DE INVESTIGACIÓN Y ENSEÑANZA. *Mimosa scabrella*. In: **Silvicultura de especies promisorias para producción de leña en America Central**: resultados de cinco años de investigación. Turrialba: CATIE-ROCAP, p. 205-210, 1986.

CAVAGNARO, P. F. et al. Microsatellite isolation and marker development in carrot-genomic distribution, linkage mapping, genetic diversity analysis and marker transferability across Apiaceae. **BMC genomics**, v. 12, n. 1, p. 1, 2011.

CHOR, B. et al. Genomic DNA k-mer spectra: models and modalities. **Genome Biology**, v. 10, n. 10, p. 1, 2009.

COSTA, B. F. et al. Characterization of nine microsatellite loci for the tree species *Parapiptadenia rigida* (Fabaceae-Mimosoideae) and their transferability. **Genetics and Molecular Research**, v. 11, n. 3, p. 2338-2342, 2013.

DAHMER, N. et al. Chromosome numbers in the genus *Mimosa* L.: cytotaxonomic and evolutionary implications. **Plant Systematics and Evolution**, v. 291, n. 3-4, p. 211-220, 2011.

DAHMER, N. et al. “Bracatinga”(*Mimosa scabrella* Bentham), a multipurpose tree growing in Southern Brazil: chromosome number and genetic variation. **Genetic Resources and Crop Evolution**, v. 60, n. 1, p. 377-383, 2013.

DIEFFENBACH, C. W.; LOWE, T. M.; DVEKSLER, G. S. General concepts for PCR primer design. **PCR Methods Appl**, v. 3, n. 3, p. S30-S37, 1993.

DIEKMANN, K.; HODKINSON, T. R.; BARTH, S. New chloroplast microsatellite markers suitable for assessing genetic diversity of *Lolium perenne* and other related grass species. **Annals of botany**, v. 110, n. 6, p. 1327-1339, 2012.

DONKPEGAN, A. S. L. et al. Microsatellite development and flow cytometry in the African tree genus *Afzelia* (Fabaceae, Caesalpinioideae) reveal a polyploid complex. **Applications in Plant Sciences**, v. 3, n. 1, 2015.

DOYLE, J. J.; DOYLE, J. L. Isolation of Plant DNA from Fresh Tissue. **Focus**, v. 12, n. 1, p. 13-15, 1990.

DUFRESNE, F. et al. Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. **Molecular Ecology**, v. 23, n. 1, p. 40-69, 2014.

DUTRA, V. F.; MORIM, M.P. 2016. Mimosa in Lista de Espécies da Flora do Brasil. **Jardim Botânico do Rio de Janeiro**. Disponível em: <<http://floradobrasil.jbrj.gov.br/jabot/floradobrasil/FB100978>>. Acesso em: 18 Jan. 2016

EBERT, D.; PEAKALL, R. O. D. Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. **Molecular Ecology Resources**, v. 9, n. 3, p. 673-690, 2009.

ELLEGREN, H. Microsatellites: simple sequences with complex evolution. **Nature Reviews Genetics**, v. 5, n. 6, p. 435-445, 2004.

ESSELINK, G. D.; NYBOM, H.; VOSMAN, B. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting—peak ratios) method. **Theoretical and Applied Genetics**, v. 109, n. 2, p. 402-408, 2004.

FJELLHEIM, S. et al. Phylogeographical history of the widespread meadow fescue (*Festuca pratensis* Huds.) inferred from chloroplast DNA sequences. **Journal of Biogeography**, v. 33, n. 8, p. 1470-1478, 2006.

GAETA, R. T.; PIRES, J. C. Homoeologous recombination in allopolyploids: the polyploid ratchet. **New Phytologist**, v. 186, n. 1, p. 18-28, 2010.

GANTER, J. L. M. S.; REICHER, F. Water-soluble galactomannans from seeds of Mimosaceae spp. **Bioresource technology**, v. 68, n. 1, p. 55-62, 1999.

GARCIA, A. A. F. et al. SNP genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. **Scientific Reports**, v. 3, 2013.

GILL, P. et al. DNA Commission of the International Society of Forensic Genetics: recommendations on forensic analysis using Y-chromosome short tandem repeats. **Legal Medicine**, v. 3, n. 4, p. 252-257, 2001.

GLENN, T. C. Field guide to next-generation DNA sequencers. **Molecular Ecology Resources**, v. 11, n. 5, p. 759-769, 2011.

GRANDKE, F. et al. Bioinformatic Tools for Polyploid Crops. **Journal of Agricultural Science and Technology**, v. 4, p. 593-601, 2014.

GUICHOUX, E. et al. Current trends in microsatellite genotyping. **Molecular Ecology Resources**, v. 11, n. 4, p. 591-611, 2011.

HAMILTON, J. P.; BUELL, C. R. Advances in plant genome sequencing. **The Plant Journal**, v. 70, n. 1, p. 177-190, 2012.

HARTER-MARQUES, B.; ENGELS, W. A produção de sementes de *Mimosa scabrella* (Mimosaceae) no Planalto das Araucárias, RS, depende da polinização por abelhas sem ferrão. **Biociências**, v. 11, n. 1, p. 9-16, 2003.

HOLLELEY, C. E.; GEERTS, P. G. Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. **Biotechniques**, v. 46, n. 7, p. 511-517, 2009.

HOKANSON, K.; HANCOCK, J. Levels of allozymic diversity in diploid and tetraploid *Vaccinium* sect. Cyanococcus (blueberries). **Canadian Journal of Plant Science**, v. 78, n. 2, p. 327-332, 1998.

HOULISTON, G. J. et al. Simple sequence repeat markers for the endangered species *Clianthus puniceus* and *C. maximus* (Fabaceae). **Applications in Plant Sciences**, v. 3, n. 1, 2015.

ILLUMINA. An Introduction to Next Generation Sequencing Techonology. San Diego, Estados Unidos, 2013.
Disponível em:
<http://www.illumina.com/Documents/products/Illumina_Sequencing_Introduction.pdf>. Acesso em: 06 jul. 2016.

JULIER, B. A program to test linkage disequilibrium between loci in autotetraploid species. **Molecular Ecology Resources**, v. 9, n. 3, p. 746-748, 2009.

KATTI, M. V.; RANJEKAR, P. K.; GUPTA, V. S. Differential distribution of simple sequence repeats in eukaryotic genome sequences. **Molecular Biology and Evolution**, v. 18, n. 7, p. 1161-1167, 2001.

KLABUNDE, G. H. F. et al. Characterization of 10 new nuclear microsatellite markers in *Acca sellowiana* (Myrtaceae). **Applications in Plant Sciences**, v. 2, n. 6, 2014.

KLEIN, R.M. Aspectos fitofissociológicos da bracatinga (*Mimosa scabrella* Bentham.). In: Seminário sobre atualidades e perspectivas florestais “Bracatinga uma opção para reflorestamento” 4., 1981, Curitiba. **Anais ... Colombo: EMBRAPA CNPF, p.145-148.** 1981.

LANDER, E. S. et al. Initial sequencing and analysis of the human genome. **Nature**, v. 409, n. 6822, p. 860-921, 2001.

LI, Y.; ZHANG, W. Isolation and characterization of microsatellite markers for *Jasminum sambac* (Oleaceae) using Illumina shotgun sequencing. **Applications in Plant Sciences**, v. 3, n. 10, 2015.

LIU, L. et al. Comparison of next-generation sequencing systems. **Journal of Biomedicine and Biotechnology**, v. 2012, 2012.

MACHADO, S. A. et al. Evolução da área basal e do volume em bracatingais nativos submetidos a diferentes densidades iniciais e em diferentes sítios. **Floresta**, v. 32, n. 1, p. 61-74, 2002.

MACHADO, S. A. et al. Dinâmica da distribuição diamétrica de bracatingais na região metropolitana de Curitiba. **Revista Árvore**, v. 30, n. 5, p. 759-768, 2006.

MAIA, L. C. et al. SSR locator: tool for simple sequence repeat discovery integrated with primer design and PCR simulation. **International Journal of Plant Genomics**, v. 2008, 2008.

MAZUCHOWSKI, J. Z. et al. Situação atual dos bracatingais. In: MAZUCHOWSKI, J. Z.; RECH, T. D.; TORESAN, L. **Bracatinga, *Mimosa scabrella* Bentham:** Cultivo, Manejo e Usos da espécie. Florianópolis: Epagri, 2014, cap. 9, p. 243-283.

MCPHERSON, H. et al. Capturing chloroplast variation for molecular ecology studies: a simple next generation sequencing approach applied to a rainforest tree. **BMC ecology**, v. 13, n. 1, p. 8, 2013.

MEIRMANS, P. G.; VAN TIENDEREN, P. H. The effects of inheritance in tetraploids on genetic diversity and population divergence. **Heredity**, v. 110, n. 2, p. 131-137, 2013.

MOREIRA, P. A. et al. Genetic diversity and mating system of bracatinga (*Mimosa scabrella*) in a re-emergent agroforestry system in southern Brazil. **Agroforestry systems**, v. 83, n. 2, p. 245-256, 2011.

MORRIS, A. B. et al. Development and characterization of microsatellite primers in the federally endangered *Astragalus bibullatus* (Fabaceae). **Applications in Plant Sciences**, v. 4, n. 4, 2016.

NARAYAN, L.; DODD, R. S.; O'HARA, K. L. A genotyping protocol for multiple tissue types from the polyploid tree species *Sequoia sempervirens* (Cupressaceae). **Applications in Plant Sciences**, v. 3, n. 3, 2015.

NAZARENO, A. G. et al. Transferability and characterization of microsatellite markers in two Neotropical *Ficus* species. **Genetics and Molecular Biology**, v. 32, n. 3, p. 568-571, 2009.

NAZARENO, A. G.; ZUCCHI, M. I.; REIS, M. Microsatellite markers for *Butia eriospatha* (Arecaceae), a vulnerable palm species from the Atlantic Rainforest of Brazil. **American Journal of Botany**, v. 98, n. 7, p. e198-e200, 2011.

NAZARENO, A. G. et al. Microsatellite markers designed for tree-fern species *Dicksonia sellowiana*. **Biologia Plantarum**, v. 57, n. 3, p. 563-566, 2013.

NOCK, C. J. et al. Chloroplast genome sequences from total DNA for plant identification. **Plant biotechnology journal**, v. 9, n. 3, p. 328-333, 2011.

OLIVEIRA, A. C. B. et al. **Aplicação técnica de marcadores moleculares no melhoramento de plantas**. Campinas: Instituto Agronômico, Documentos, n. 81, 2007, 17p.

PALOP, M.; PALACIOS, C.; GONZÁLEZ-CANDELAS, F. Development and across-species transferability of microsatellitemarkers in the genus *Limonium* (Plumbaginaceae). **Conservation Genetics**, v. 1, n. 2, p. 177-179, 2000.

PALOP-ESTEBAN, M.; SEGARRA-MORAGUES, J. G.; GONZÁLEZ-CANDELAS, F.. Polyploid origin, genetic diversity and population structure in the tetraploid sea lavender *Limonium narbonense* Miller (Plumbaginaceae) from eastern Spain. **Genetica**, v. 139, n. 10, p. 1309-1322, 2011.

PARK, H et al. Development of Chloroplast Microsatellite Markers for the Endangered *Maianthemum bicolor* (Asparagaceae sl). **Applications in Plant Sciences**, v. 4, n. 8, p. 1600032, 2016.

PEAKALL, R. O. D.; SMOUSE, Peter E. GENALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. **Bioinformatics**, v. 28, n. 19, p. 2537-2539, 2012.

PEGORARO, A.; CARPANEZZI, A. A. Avaliação do potencial melífero da Bracatinga. **Revista do Setor de Ciências Agrárias**, v. 14, n. 2, p. 167-172, 1995.

POWELL, W. et al. Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. **Current Biology**, v. 5, n. 9, p. 1023-1029, 1995.

PROVAN, J.; POWELL, W.; HOLLINGSWORTH, P. M. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. **Trends in Ecology & Evolution**, v. 16, n. 3, p. 142-147, 2001.

PUYVELDE, K.V.; GEERT, A. V.; TRIEST, L. ATETRA, a new software program to analyse tetraploid microsatellite data: comparison with TETRA and TETRASAT. **Molecular Ecology Resources**, v. 10, n. 2, p. 331-334, 2010.

RAMSEY , J. , SCHEMSKE, D. W. Pathways, mechanisms, and rates of polyploid formation in flowering plants. **Annual Review of Ecology and Systematics**, v. 29, p. 467-501, 1998.

REITZ, R.; KLEIN, R. M.; REIS, A. **Projeto madeira de Santa Catarina**. Itajaí: Herbário Barbosa Rodrigues, 1978. 320 p.

ROZEN, S.; SKALETSKY, H. Primer3 on the WWW for general users and for biologist programmers. **Bioinformatics Methods and Protocols**, p. 365-386, 1999.

RUFINO, R. J. N. et al. Marcadores microssatélites para o cafeeiro. In: IV SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 2005, Londrina. **Anais Eletrônicos...** Londrina. CD-ROM, 2005.

SAKAGUCHI, S et al. Development of nuclear and chloroplast microsatellite markers for the endangered conifer *Callitris sulcata* (Cupressaceae). **Applications in plant sciences**, v. 3, n. 8, 2015.

SALGUEIRO, F. et al. Characterization of nuclear microsatellite loci in South American Araucariaceae species. **Molecular Ecology Notes**, v. 5,n. 2, p. 256–258, 2005.

SANGER, F. et al. Nucleotide sequence of bacteriophage Φ X174 DNA. **Nature**, v. 265, p.687-695, 1977.

SCHIFINO-WITTMANN, M. T.; DALLA GNOL, M. Gametas não reduzidos no Melhoramento de Plantas. **Ciência Rural**, v. 31, n. 1, p. 169-175, 2001.

SCHMIDT, A.B. et al. Isolation and characterization of microsatellite markers for *Araucaria angustifolia* (Araucariaceae). **Molecular Ecology Notes**, v. 7, n. 2, p. 340–342, 2007.

SCHNABLE, P. S. et al. The B73 maize genome: complexity, diversity, and dynamics. **Science**, v. 326, n. 5956, p. 1112-1115, 2009.

SCHLÖTTERER, C. The evolution of molecular markers—just a matter of fashion? **Nature Reviews Genetics**, v. 5, n. 1, p. 63-69, 2004.

SEIJO, G. Números cromosómicos de especies de *Mimosa* (Leguminosae) de Paraguay. **Bonplandia**, v. 10, p. 163-167, 2000.

SELKOE, K. A.; TOONEN, R. J. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. **Ecology letters**, v. 9, n. 5, p. 615-629, 2006.

SEOANE, C. E. S. **Efeitos da Fragmentação Florestal sobre a Genética de Populações de Guarantã**. Ministério da Agricultura, Pecuária e Abastecimento. Colombo: Embrapa Florestas, Documentos, 159, 2007. 83 p.

SHENDURE, J.; JI, H. Next-generation DNA sequencing. **Nature Biotechnology**, v. 26, n. 10, p. 1135-1145, 2008.

SILVA, P. I. T. et al. Development and validation of microsatellite markers for *Brachiaria ruziziensis* obtained by partial genome assembly of Illumina single-end reads. **BMC Genomics**, v. 14, n. 1, p. 1, 2013.

STIFT, M. et al. Segregation models for disomic, tetrasomic and intermediate inheritance in tetraploids: a general procedure applied to *Rorippa* (yellow cress) microsatellite data. **Genetics**, v. 179, n. 4, p. 2113-2123, 2008.

SOBIERAJSKI, G.R.; KAGEYAMA, P.Y.; SEBBENN, A.M. Sistema de reprodução em nove populações de *Mimosa scabrella* Benth. **Scientia Forestalis**, v.71, p.37-49, 2006.

SOUZA, D. C. L. Técnicas moleculares para caracterização e conservação de plantas medicinais e aromáticas: uma revisão. **Revista Brasileira de Plantas Medicinais**, v. 17, n. 3, p. 495-503, 2015.

SOUZA, G. A; MARTINS E.R. Análise de risco de erosão genética de populações de fava-d’anta (*Dimorphandra mollis* Benth.) no norte de Minas Gerais. **Revista Brasileira de Plantas Medicinais**, v.6, n.3, p.42-47, 2004.

STEENBOCK, W. **Domesticação de Bracatingais:** perspectivas de inclusão social e conservação ambiental. 2009. 281f. Tese (Doutorado em Recursos Genéticos Vegetais) - Universidade Federal de Santa Catarina, Florianópolis, 2009.

STEENBOCK, W. et al. Ocorrência da bracatinga (*Mimosa scabrella* Benth.) em bracatingais manejados e em florestas secundárias na região do Planalto Catarinense. **Revista Árvore**, v. 35, n. 4, p. 845-857, 2011.

STEENBOCK, W.; REIS, M. S. Domesticação da paisagem em bracatingais de assentamentos rurais no noroeste do planalto catarinense. **Ciência Florestal**, v. 23, n. 3, p. 427-437, 2013.

THRALL, P. H.; YOUNG, A. Autotet: a program for analysis of autotetraploid genotypic data. **Journal of Heredity**, v. 91, n. 4, p. 348-349, 2000.

TSUKAYA, H. Does ploidy level directly control cell size? Counterevidence from *Arabidopsis* genetics. **PLoS One**, v. 8, n. 12, p. e83729, 2013.

UGHINI, F. et al. Evaluation of xanthan and highly substituted galactomannan from *M. scabrella* as a sustained release matrix. **International Journal of Pharmaceutics**, v. 271, n. 1, p. 197-205, 2004.

URBANO, E. et al. Equações para estimar o peso de carbono fixado em árvores de *Mimosa scabrella* Bentham (Bracatinga) em povoamentos nativos. **Cerne**, v.14, n. 3, p. 194-203, 2008.

VARSHNEY, R. K.; GRANER, A.; SORRELLS, M. E. Genic microsatellite markers in plants: features and applications. **Trends in Biotechnology**, v. 23, n. 1, p. 48-55, 2005.

VALLONE, P. M.; BUTLER, J. M. AutoDimer: a screening tool for primer-dimer and hairpin Structures. **Biotechniques**, v. 37, n. 2, p. 226-231, 2004.

VENCOVSKY, R. Tamanho efetivo populacional na coleta e preservação de germoplasmas de espécies alógamas. **IPEF**, n. 35, p.79-84, 1987.

VENDRUSCOLO, C. W. et al. Xanthan and galactomannan (from *M. scabrella*) matrix tablets for oral controlled delivery of theophylline. **International Journal of Pharmaceutics**, v. 296, n. 1, p. 1-11, 2005.

VIEIRA, L. N. et al. The complete chloroplast genome sequence of *Podocarpus lambertii*: genome structure, evolutionary aspects, gene content and SSR detection. **PloS one**, v. 9, n. 3, p. e90618, 2014.

WEISING, K; WÖHRMANN, T.; HUETTEL, B. The use of high-throughput DNA sequencing for microsatellite discovery in plants. In: Hörndl, E.; Appelhans, M. S. **Next-Generation Sequencing in Plant Systematics**. Königstein: Koeltz Scientific Books, 2015. 66p.

WHEELER, G. L. et al. A review of the prevalence, utility, and caveats of using chloroplast simple sequence repeats for studies of plant biology. **Applications in plant sciences**, v. 2, n. 12, p. 1400059, 2014.

WINFIELD, M. O. et al. Targeted re-sequencing of the allohexaploid wheat exome. **Plant Biotechnology Journal**, v. 10, n. 6, p. 733-742, 2012.

WRICKE, G.; WEBER, W. E. Quantitative Genetics and Selection in Plant Breeding. 1 ed. New York: Walter de Gruyter, 1986. 406p.

WU, J. H. et al. Induced polyploidy dramatically increases the size and alters the shape of fruit in *Actinidia chinensis*. **Annals of Botany**, v. 109, n. 1, p. 169-179, 2012.

WU, C. A.; CAMPBELL, D. R. Cytoplasmic and nuclear markers reveal contrasting patterns of spatial genetic structure in a natural Ipomopsis hybrid zone. **Molecular Ecology**, v. 14, n. 3, p. 781-792, 2005.

XU, X. et al. Genome sequence and analysis of the tuber crop potato. **Nature**, v. 475, n. 7355, p. 189-195, 2011.

ZAMARIAN, E. H. C. **Características técnicas, estéticas e mercadológicas da bracatinga (*Mimosa scabrella* Bentham) na fabricação de móveis**. 2008. 118 f. Dissertação (Mestrado em Ciências Florestais) - Universidade Federal do Paraná, Curitiba, 2008.

ZALAPA, J. E. et al. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. **American Journal of Botany**, v. 99, n. 2, p. 193-208, 2012.

ZANE, L.; BARGELLONI, L.; PATARNELLO, T. Strategies for microsatellite isolation: a review. **Molecular Ecology**, v. 11, n. 1, p. 1-16, 2002.

ZANELLA, C. M. et al. Genetic structure and phenotypic variation in wild populations of the medicinal tetraploid species *Bromelia antiacantha* (Bromeliaceae). **American Journal of Botany**, v. 98, n. 9, p. 1511-1519, 2011.

ZERBINO, D. R.; BIRNEY, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. **Genome Research**, v. 18, n. 5, p. 821-829, 2008.

ZHOU, Z. et al. Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. **Nature Biotechnology**, v. 33, n. 4, p. 408-414, 2015.

ZUCCHI, M. I. et al. Genetic structure and gene flow in *Eugenia dysenterica* DC in the Brazilian Cerrado utilizing SSR markers. **Genetics and Molecular Biology**, v. 26, n. 4, p. 449-457, 2003.