



ISABELA PIRES BARROS

**DIVERSIDADE E ESTRUTURA GENÉTICA EM
POPULAÇÕES NATURAIS E DE MELHORAMENTO DE
*Eucalyptus urophylla***

LAVRAS-MG

2021

ISABELA PIRES BARROS

**DIVERSIDADE E ESTRUTURA GENÉTICA EM
POPULAÇÕES NATURAIS E DE MELHORAMENTO DE
*Eucalyptus urophylla***

Monografia apresentada à Universidade Federal de Lavras, como parte das exigências do Curso de Ciências Biológicas, para a obtenção do título de Licenciada.

Prof. Dr. Evandro Novaes

Orientador

LAVRAS-MG

2021

ISABELA PIRES BARROS

**DIVERSIDADE E ESTRUTURA GENÉTICA EM POPULAÇÕES
NATURAIS E DE MELHORAMENTO DE *Eucalyptus urophylla***

Monografia apresentada à Universidade Federal de Lavras, como parte das exigências do Curso de Ciências Biológicas, para a obtenção do título de Licenciada.

Aprovada em 30 de novembro de 2021.

Dr. Evandro Novaes

Dr.a Lucimara Cruz de Souza, UFLA

Dr. Paulo Henrique Muller da Silva, IPEF

Prof. Dr. Evandro Novaes

Orientador

LAVRAS-MG

2021

*À minha família por todo o apoio e amor.
Com todo o meu amor, carinho e gratidão.*

AGRADECIMENTOS

À Universidade Federal de Lavras, por possibilitar minha formação.

À FAPESP pelo apoio financeiro ao projeto (Proc. 15/15651-2)

Ao professor Evandro Novaes, por ter me acolhido, me orientado e por toda a experiência que isso me trouxe. E aos demais professores, com quem tive o prazer em aprender.

Ao Leonardo Oliveira, por ter me ajudado imensamente neste trabalho.

Ao IPEF e ao Paulo Henrique Müller da Silva por cederem os dados para a realização deste trabalho.

Aos amigos de graduação por caminharem comigo desde o início e por compartilharem tudo de bom e de ruim que a graduação tem a oferecer.

Às minhas irmãs, Isadora e Vitória, que são tudo pra mim e me inspiram a ser uma pessoa cada dia melhor. Vocês são a melhor coisa da minha vida.

Aos meus pais, por nunca medirem esforços para me ajudar, por serem sempre tão presentes, sempre tão atenciosos e maravilhosos! Sou extremamente grata por toda a confiança que vocês têm em mim.

Ao meu namorado, por todo o amor, apoio e companheirismo.

Obrigada por acreditarem no meu sonho e por torná-lo possível. Eu amo vocês incondicionalmente.

Muito obrigada!

RESUMO

Eucalyptus urophylla é uma importante espécie de interesse econômico, cultivada para a produção de madeira. Dessa forma, o conhecimento de suas características genéticas, tanto em populações naturais como de melhoramento, é imprescindível para o melhoramento genético dessa espécie. Diante disso, este trabalho teve como objetivo avaliar a diversidade genética entre e dentro de sete populações de *E. urophylla*; estimar a distância genética entre os indivíduos e inferir a estrutura genética entre e dentro das populações amostradas. Para isso, 19 marcadores microssatélites foram empregados na genotipagem de 254 indivíduos de sete populações de *E. urophylla*, sendo quatro constituídas por populações naturais amostradas nas ilhas da Indonésia, duas por populações de melhoramento e uma constituída por clones comerciais. Utilizou-se o programa de análises estatísticas R para a obtenção dos parâmetros, que indicou que todos os *locos* são polimórficos, a heterozigosidade média esperada foi de $He = 0,8165$ e a observada $Ho = 0,7456$. Quanto a estruturação foi estimado o $Fst = 0,030$, o que indica uma baixa estruturação entre as populações e um $Fis = 0,101$, que aponta para um baixo nível de endogamia. Também foi calculado a distância genética de Nei, o que revelou que as populações que se encontram mais distantes geneticamente também são as mais distantes geograficamente. As distâncias genéticas foram representadas na forma de um dendrograma, em que os grupos tiveram uma baixa consistência entre os nós, em 10.000 *bootstraps*, demonstrando uma pequena estruturação genética. Realizou-se também análises de componentes principais (PCA), em que se evidenciou novamente a baixa estruturação das populações. Essa falta de estruturação foi confirmada nas análises realizadas pelo programa *Structure*. Esses resultados apontam para a existência de fluxo gênico entre as populações naturais e um curto tempo de melhoramento genético nas populações introduzidas no Brasil, uma vez que essas ainda são muito semelhantes às populações naturais.

Palavras-chave: Genética de populações. *Eucalyptus urophylla*. Diversidade genética.

RESUMO EM LÍNGUA ESTRANGEIRA

Eucalyptus urophylla is an important species of economic interest, harvested for wood production. That way, the knowledge about its genetic characteristics, both in natural and enhanced populations, is essential for this species's genetic enhancement. That said, this work targeted to evaluate the genetic diversity between the subjects, among and inside seven *E. urophylla* populations; to estimate genetic distance among subjects and to infer the genetic structure between and inside sampled populations. As so, 19 microsatellite markers were used to genotype 254 subjects of seven *E. urophylla*'s populations, being four constituted of Indonesian islands natural population samples, two of enhanced populations and one constituted of commercial clones. It was used in the statistical analysis program R to obtain the parameters that identified all *loci* as polymorphic, the average expected heterozygosity was $He=0,8165$ and was observed $Ho=0,7456$. About the structure, it was estimated $Fis=0,030$, which indicates a low at structurations among populations and a $Fis=0,101$ which points to a low endogamy level. It was additionally calculated genetic distance between Nei, revealing that populations that are more genetic far away are also geographically distant. Genetic distances were represented at a dendrogram, in which the group had a low consistency throughout the knots, at 10.000 bootstraps, demonstrating a low genetic structuration. It was performed as well the Principal Components Analysis (PCA), in which it became evident again the low population genetic structuring. This lack of structuring was confirmed in the analysis carried out by the *Structure* program. These results point to the existence of genetic flow amongst natural populations and a short time of genetic enhancement at the introduced population in Brazil, since these ones are still similar to natural populations.

Keywords: Populations Genetic. *Eucalyptus urophylla*. Genetic diversity.

SUMÁRIO

1. INTRODUÇÃO	8
2. OBJETIVOS	10
3. REFERENCIAL TEÓRICO	11
3.1. <i>Eucalyptus urophylla</i>	11
3.2. Marcadores genéticos um enfoque em marcadores microsstélites	12
3.3. Diversidade e estrutura genética de populações	14
4. MATERIAIS E MÉTODOS.....	18
4.1. Populações de estudo.....	18
4.2. Extração de material genético	19
4.3. Marcadores microssatélites	19
4.4. Diversidade e estrutura genética	20
5. RESULTADOS E DISCUSSÃO	22
5.1. Diversidade genética.....	22
5.2. Distância e estrutura genética.....	25
6. CONSIDERAÇÕES FINAIS	34
7. REFERÊNCIAS BIBLIOGRÁFICAS	35
APÊNDICE.....	42

1. INTRODUÇÃO

A demanda por madeira é muito grande no país, o que pode gerar pressão de corte em espécies nativas, acarretando em diversos problemas ambientais como perda de diversidade, risco de extinção de espécies e impactos ambientais graves aos ecossistemas (Relatório Ibá, 2020). No Brasil, que é o segundo maior produtor de madeira tropical do mundo, o bioma mais impactado pelo extrativismo de madeira é a Amazônia (SFB & IPAM, 2011). Como forma de reduzir esses problemas, a silvicultura foi inserida no país por volta do século XX e vem se baseando principalmente em espécies exóticas como o *Pinus* e o eucalipto (FERREIRA & SANTOS, 1997).

Em 2019, a produção florestal foi responsável por 1,2% do PIB do Brasil. Além disso, houve a geração de 1,3 milhões de empregos diretos (Relatório IBÁ, 2020). As espécies de eucalipto são as mais importantes pois representam aproximadamente 73% dos plantios florestais do país, destacando-se a *Eucalyptus urophylla*. (Relatório IBÁ, 2019; ALMEIDA, 2012). A madeira pode ser utilizada para a produção de celulose, papel, chapas duras, carvão, dentre outras utilidades. A espécie de *E. urophylla* se destaca no país e se caracteriza como uma das espécies mais plantadas, principalmente devido à sua resistência ao fungo *Cryphonectria cubensis* que causa o cancro do eucalipto (FERREIRA, 1999).

Por isso, é importante conhecer as características genéticas das populações naturais dessa espécie, pois elas constituem recursos genéticos importantes para o melhoramento do eucalipto, especialmente no Brasil, em que as unidades plantadas são oriundas de um curto tempo de melhoramento genético, ou seja, ainda são muito próximas às populações naturais (SILVA, 2019).

Nesse contexto, estudos são necessários, uma vez que essas informações são essenciais para os programas que visam o seu melhoramento. Para isso, é possível utilizar os marcadores moleculares em nível de DNA para estimar a variabilidade genética dentro de uma população e entre populações distintas (MORA *et al*, 2016; LU *et al*, 2018; BALDONI *et al*, 2020). Trabalhos como o realizado por Guimarães (2019), demonstram como essa técnica viabiliza a quantificação da diversidade e estruturação genética em espécies vegetais.

Devido às suas propriedades como a codominância, multi-alelismo, grande taxa de polimorfismo e transferabilidade, os microssatélites se destacam dentre os marcadores genéticos. Com eles é possível estimar parâmetros importantes para o estudo da genética de populações, como o nível de heterosigozidade esperada e observada; riqueza alélica; estrutura genética e a distância genética das populações (ZOLET *et al*, 2017; SOARES, 2014). Por isso, os microssatélites têm sido bastante utilizados para a determinação do grau de diversidade e de estruturação de populações de eucalipto (COSTA, 2017; PAYN, 2008). Dessa forma, marcadores microssatélites são uma ótima opção para o estudo dos aspectos genético-populacionais de espécies florestais.

2. OBJETIVOS

- Avaliar a diversidade genética entre e dentro de sete populações de *E. urophylla*, sendo que quatro são populações naturais; duas de melhoramento e uma de clones comerciais;
- Estimar a distância genética entre os indivíduos de *E. urophylla*;
- Inferir sobre a estrutura genética entre e dentro das populações amostradas.

3. REFERENCIAL TEÓRICO

3.1. *Eucalyptus urophylla*

Eucalyptus urophylla S.T. Blake é uma espécie arbórea de grande porte pertencente à família *Myrtaceae* (NICOLLE, 2015). Seu habitat é localizado em regiões de clima tropical, com médias de temperaturas anuais variando entre 16 e 27° C e precipitação de 1000 a 2000 mm ao ano. A espécie ocorre em algumas ilhas do Arquipélago de Sonda, como na ilha de Flores, Adonara, Wetar, Alor, Lembata, Pantar, e Timor da Indonésia, se caracterizando como um raro caso de ocorrência natural de eucalipto fora da Austrália (ELDRIDGE et al., 1994).

Em Timor são encontrados espécimes com altura de 45 m e 2 m de diâmetro, entretanto é mais comum atingirem até 40 m de altura e 80 cm de diâmetro. Naturalmente são encontradas em altitudes que variam desde o nível do mar até 2500 m. A localização geográfica mais precisa de ocorrência da espécie vai de 7°30' a 10° S de Latitude e 122° a 127° E de Longitude e sua distribuição tem por volta de 476 km no sentido leste/oeste e 230 km no sentido norte/sul (DVORAK; HODGE, PAYN, 2008).

Nas regiões de baixa altitude em Timor, há a interseção de habitats entre *E. urophylla* e *E. alba*. Nessas regiões as hibridações naturais são comuns devido à coincidência dos períodos reprodutivos de ambas as espécies. A medida que a altitude aumenta, há o predomínio de *E. urophylla* e populações puras são encontradas (ELDRIDGE et al., 1994; DVORAK; HODGE; PAYN, 2008).

A reprodução tem início no segundo ano do plantio, porém é com a idade de quatro anos que a reprodução aumenta significativamente, ocorrendo nas estações secas. A polinização é realizada por insetos e as sementes atingem o amadurecimento seis meses depois do florescimento (SELN & MITLÖHNER, 2011). As flores dessa espécie são hermafroditas e protândricas, para se evitar autofecundação. Além disso, há um mecanismo de autoincompatibilidade genética. Estudos indicam que as taxas de crescimento do tubo polínico em situações de autofecundação são menores e as anormalidades são maiores. (ASSIS et al., 2004; PRYOR, 1976).

3.2. Marcadores genéticos um enfoque em marcadores microsstélites

O primeiro estudo que sugeriu a possibilidade do uso de características visíveis como marcadores morfológicos foi publicado no início do século XX, quando houve a descoberta do conceito de genes ligados, em que percebeu-se que o gene para cor das pétalas das flores de ervilha era ligado ao gene da forma do grão de pólen (ZOLET *et al*, 2017; FERREIRA, GRATTAPAGLIA, 1998).

Os marcadores morfológicos foram um marco histórico para a pesquisa em genética e são usados até os dias atuais. Entretanto, algumas características só aparecem em determinadas fases de vida dos organismos, logo, são de uso limitado e podem demorar para aparecer, além de poderem sofrer alterações dependendo do ambiente (ZOLET *et al*, 2017).

Com o passar do tempo, novos marcadores foram descobertos, como os bioquímicos, que analisam alguns compostos produzidos pelos seres vivos como os metabólitos secundários, enzimas e outros compostos, House e Bell (1994), publicaram um estudo com populações naturais de *E. urophylla* em que usaram de isoenzimas para determinarem a diversidade genética. Porém, essa categoria de marcadores é pouco utilizada atualmente. Com o desenvolvimento das técnicas de marcadores moleculares a nível de DNA houve uma revolução e um avanço rápido em estudos que utilizam das técnicas (GROVER, 2016).

Os marcadores moleculares são importantes ferramentas utilizadas com bastante recorrência para várias espécies vegetais e animais, como em trabalhos de Motamayor *et al* (2009) e Baldoni *et al* (2020). Neles, marcadores microsatélites foram usados para obtenção de informações genéticas para o estudo de populações de espécies da floresta amazônica. Além disso, com os marcadores moleculares é possível estimar a variabilidade genética, o parentesco entre indivíduos, realizar estudos em genética de populações e, com isso, racionalizar as ações tanto para conservação de espécies naturais como para melhoramento genético das cultivadas (SOARES, 2014; SCHNEIDER, 2020; SILVA, 2014).

Ademais, os marcadores moleculares podem ser do tipo dominante ou codominante. Os tipos dominantes não possibilitam a discriminação entre indivíduos heterozigotos e de homozigoto com duas cópias do alelo, já os marcadores codominantes possibilitam a identificação separada desses dois genótipos (ZOLET *et al*, 2017).

Para o estudo de populações, um dos marcadores mais usados, são os marcadores microssatélites. Basicamente, os microssatélites são regiões do genoma com repetições em tandem (um atrás do outro) de pequenos motivos que possuem de um a seis nucleotídeos. Essas sequências simples repetidas (SSRs ou microssatélites) são frequentes ao longo de todo o DNA, tanto de eucariotos quanto procariotos. Os SSRs são geralmente encontrados nos genomas nucleares (nuSSR), mas também são identificados em mitocôndrias (mtSSR) e cloroplastos (cpSSR) (SANSALONI, 2008; KALIA et al., 2011).

No caso dos eucariotos, essas repetições são mais frequentes, com alto polimorfismo e mais distribuídas no genoma do que nos procariotos. Os microssatélites possuem algumas vantagens em relação à outras classes de marcadores, tais como: abundância nos genomas de eucariotos, são codominantes, multialélicos, fornecem muitas informações e se caracterizam por terem as maiores taxas de polimorfismos entre os marcadores moleculares (ZOLET *et al*, 2017; FERREIRA, GRATTAPAGLIA, 1998;).

Essa alta taxa de polimorfismo e multi-alelismo se deve à alta taxa de mutação que ocorre durante a replicação e/ou recombinação (ELLEGREN, 2004). Dessa forma, os SSR tem sido bastante utilizados nas pesquisas em genética de população, pois suas propriedades são úteis na determinação de variações genéticas, tanto entre indivíduos de uma única população, quanto entre populações (CHEZHIAN; YASODHA; GHOSH, 2010, GOVINDARAJ; VETRIVENTHAN; SRINIVASAN, 2015).

Para realizar a genotipagem dos microssatélites, os fragmentos de DNA contendo esses *locos* de interesse são amplificados por meio da PCR (reação em cadeia da polimerase). Nas regiões adjacentes aos microssatélites há regiões conservadas, que podem ser usadas como base para a escolha de primers. Dessa forma, pode-se identificar o tamanho dos fragmentos amplificados, por meio de eletroforese e, assim, determinar as diferenças no número de repetições das SSRs entre os indivíduos, avaliando também se são homocigotos ou heterocigotos. A interpretação desses dados se dá por meio do acréscimo ou decréscimo do número de pares de bases das regiões de microssatélites (ZOLET *et al*, 2017).

A partir da genotipagem dos *locos* microssatélites em uma ou mais amostras de indivíduos, é possível estimar diversos parâmetros sobre a estrutura e diversidade genética das populações amostradas. Esses parâmetros incluem: coeficiente de parentesco; tamanho efetivo de populações; o número total de alelos; o número médio

de alelos por loco; a riqueza alélica, a heterozigosidade observada e esperada sob equilíbrio de Hardy Weinberg; os índices de fixação de Wright, dentre outros dados importantes para a compreensão das dinâmicas populacionais (PORTH; ELKASSABY, 2014). Esses marcadores microssatélites já foram utilizados em pesquisas com espécies florestais (PAYN *et al*, 2008; TRIPIANA *et al*, 2007).

3.3. Diversidade e estrutura genética de populações

No âmbito da genética de populações, o termo população se refere ao conjunto de indivíduos da mesma espécie que vivem numa região restrita, em que haja cruzamento entre os indivíduos. Entretanto, delimitar as populações não é uma tarefa banal, pois os indivíduos se encontram dispersos no espaço e quase sempre não é possível saber ao certo a extensão do fluxo gênico dos indivíduos. Assim, mesmo indivíduos que se encontram mais distantes podem contribuir com alelos na população (CRUZ *et al*, 2020). Esses padrões de distribuição e acasalamentos de uma população são a base para a definição da estrutura e diversidade genética de uma população (HAMRICK & GODT, 1996).

A diversidade genética é, segundo Wang, Barkley e Jenkins (2009), o acúmulo de modificações dos nucleotídeos, genes, cromossomos ou genomas dos seres vivos. Todas as espécies possuem um *pool* gênico característico, que é constituído por uma variedade de alelos, que são combinados de diversas formas ao longo das gerações, podendo ser mais ou menos diverso a depender das forças evolutivas atuando em suas populações. Populações com maiores índices de diversidade genética tendem a possuir vantagens evolutivas, pois há maior probabilidade de haver indivíduos com características adaptativas que podem ser valiosas para enfrentarem as dificuldades impostas pelo ambiente no processo da seleção natural (Hartl & Clarck, 2010).

O modelo mais básico da variação genética em populações foi desenvolvido, independentemente, em 1909, por Hardy e Weinberg e é chamado de equilíbrio de Hardy-Weinberg (*EHW*). Nesse modelo, há uma série de premissas para que uma população esteja em *EHW*. Para estarem em *EHW*, as populações não podem estar sofrendo a ação de nenhuma força evolutiva (deriva, seleção, mutação e migração) e, para isso, os cruzamentos devem ocorrer ao acaso, as populações devem ser grandes para que se possa desconsiderar o efeito da deriva genética, a reprodução deve ser sexuada, sem sobreposição de gerações e as frequências alélicas devem se distribuir

igualmente entre machos e fêmeas. Nesse caso, após uma única geração de cruzamentos ao acaso, tanto as frequências alélicas como genotípicas permanecem estáveis e podem ser descritas pelo modelo do *EHW*. Entretanto, se houverem forças que modifiquem as frequências alélicas, a população sai do *EHW*. Assim, desvios significativos nas frequências esperadas no *EHW* indicam que há forças evolutivas atuando sobre essa população (Hartl & Clarck, 2010).

Logo, a variação genética de uma população pode ser compreendida como o resultado da combinação de alguns processos evolutivos como fluxo gênico, deriva genética e seleção. A deriva genética é um processo no qual as frequências alélicas se modificam de forma aleatória no decorrer das gerações. Inevitavelmente, a deriva reduz a variação genética, pois há a perda de alelos ao longo das gerações. A tendência é que quanto menor for a população, maior é a velocidade da perda de diversidade causada pela deriva genética. Quando uma população é subdivida em subpopulações, se não houver fluxo gênico entre elas, com o passar do tempo a deriva genética vai aumentando o distanciamento genético. Entretanto, quando há fluxo gênico, ou seja, troca de gametas entre indivíduos pertencentes a subpopulações diferentes, há a homogeneização genética o que as mantém similares geneticamente (Hartl & Clarck, 2010). Essas diferenças genéticas caracterizam a estrutura de populações.

Dessa maneira, o entendimento da distribuição da variabilidade genética entre e dentro das populações pode indicar a ação relativa das forças evolutivas que aumentam a distância genética entre as populações (deriva genética e seleção) versus o fluxo gênico, que diminui essa distância. Essa distribuição da variabilidade genética no espaço é conhecida como estrutura genética populacional. A estrutura genética é inferida por meio dos dados de frequências alélicas e genotípicas existentes entre os indivíduos da população. Com esses parâmetros é possível extrair outras informações valiosas como a taxa de diferenciação genética, a trajetória evolutiva, as taxas de migração, dispersão, endogamia, diversidade genética. Essas estimativas podem ser aplicadas tanto para o manejo das espécies ameaçadas de extinção quanto para o melhoramento genético (Robinson, 1998; Hartl & Clarck, 2010; CRUZ et al., 2020).

Geralmente, as árvores de grande porte de florestas tropicais possuem baixa estrutura genética, ou seja, a variabilidade genética se distribui majoritariamente dentro das populações. Apesar de serem organismos sésseis, a taxa de fluxo gênico

é alta, pois a dispersão de pólen e sementes é muito eficiente, o que permite que indivíduos distantes combinem seus genomas, reduzindo o distanciamento genético. Essa eficiência na dispersão de propágulos aumenta o tamanho das populações, reduzindo a taxa de deriva genética. Além disso, devido a sua longevidade, as forças evolutivas tendem a atuar de forma mais lenta, por unidade de tempo (e não gerações), nas espécies arbóreas. Com isso a perda de alelos, seja por deriva genética ou por seleção, ocorre de forma mais lenta (LOVELESS & HAMRICK, 1984; COLLEVATTI *et al.*, 2001; GANDARA, 2009).

Uma das consequências da divisão e estruturação de subpopulações é a redução dos acasalamentos ao acaso e aumento da taxa de cruzamento entre indivíduos mais aparentados, o que reduz a heterozigosidade. Wright, definiu o índice de fixação, que quantifica as consequências do endocruzamento gerado pela subdivisão populacional. O índice de fixação F de Wright é um indicador de diferenciação genética, em que é acrescentado subscritos para identificar os níveis hierárquicos que serão comparados. Com essa estatística, é possível calcular os níveis médios de endogamia no âmbito de população (F_{IS}) e total (F_{IT}), bem como a diferenciação genética entre as populações (F_{ST}). Esses índices de fixação indicam a estrutura genética das populações e, portanto, são muito utilizados nas pesquisas de genética de populações (HARTL & CLARCK, 2010).

De acordo com Hartl e Clark (2010), o F_{IS} quantifica o desvio das proporções de Hardy-Weinberg, indicando a probabilidade de que ambos os alelos de um gene de um indivíduo tenham a mesma origem por descendência, ou seja, o F_{IS} mede o grau de endogamia na população. Por fim, o F_{ST} é denominado de coeficiente de ancestralidade, com ele é possível estimar o nível de diferenciação genética entre as populações.

Esses coeficientes são muito utilizados em trabalhos científicos com enfoque populacional. Payn *et al* (2008), lançou mão do F_{ST} para investigar a distribuição genética de 19 populações de *E. urophylla* do leste da Indonésia. Assim como Mora *et al* (2017) utilizou o índice de diferenciação para analisar uma população de *E. cladocalyx* introduzida no Chile.

Um outro dado importante para se determinar é a distância genética, pois é a partir dela que se determina o quão próximas ou distantes as populações são entre si e, com base nesses resultados é possível a construção de dendrogramas. Essa representação gráfica, também conhecido como diagrama de ramos, facilita a

visualização da divergência genética (FERREIRA & GRATTAPAGLIA, 1998). Para isso, utilizando de dados genéticos, como as diferenças entre as frequências alélicas e genotípicas, estima-se as medidas de dissimilaridade e, a partir disso se prevê o nível da divergência genética entre os indivíduos ou entre as populações, por meio das medidas de distância genética.

4. MATERIAIS E MÉTODOS

4.1. Populações de estudo

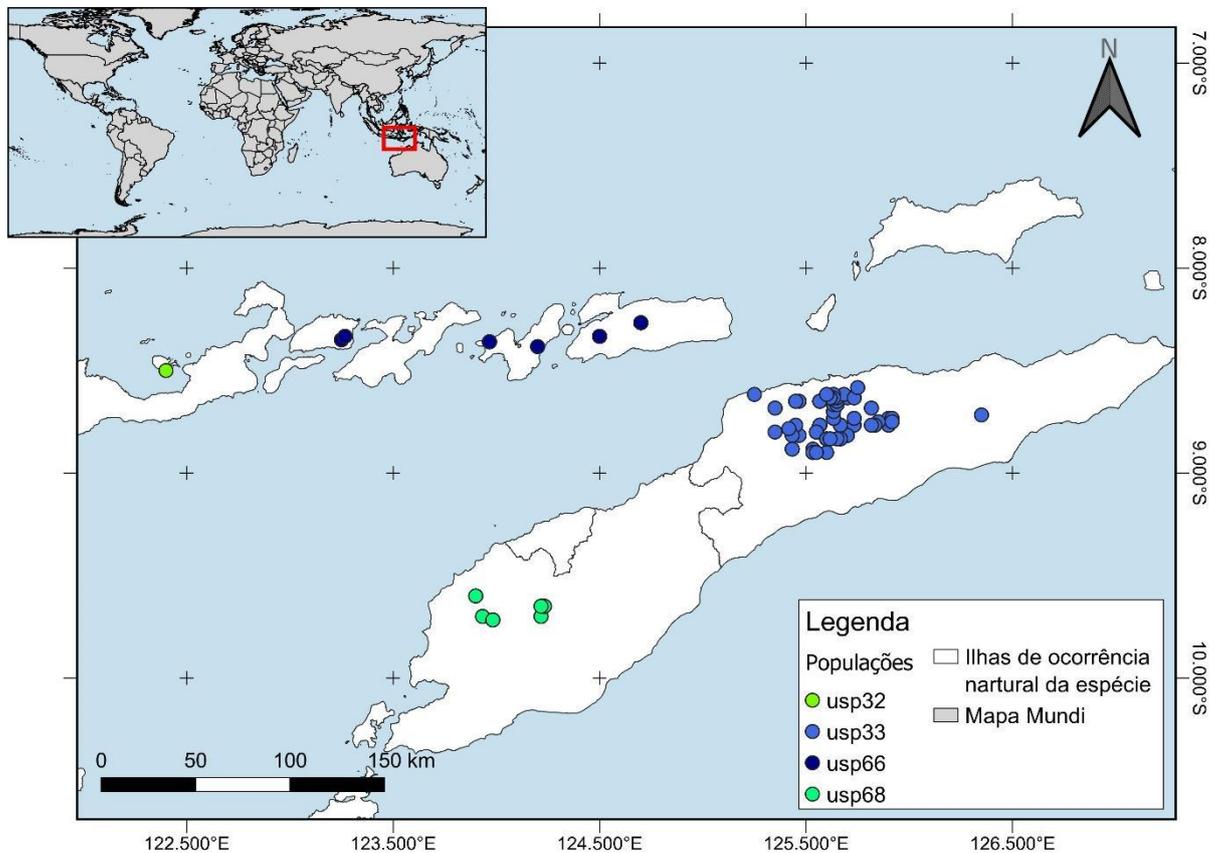
As populações de *E. urophylla* utilizadas neste estudo são constituídas por sete populações distintas. Duas delas são representadas por populações de melhoramento, ou seja, por indivíduos provenientes de programas de melhoramento em andamento. Quatro populações de introdução são compostas por indivíduos obtidos por meio de sementes oriundas de populações naturais, na região de origem da espécie em ilhas da Indonésia. Para comparação da diversidade e estrutura genética das populações naturais e de melhoramento, considerou-se dez clones comerciais como uma sétima população.

As populações de melhoramento são identificadas como Melhoramento 1 e Melhoramento 2 e são representadas respectivamente pelos 65 e 79 melhores indivíduos de cada programa. A avaliação baseou-se na sobrevivência, diâmetro a altura do peito (*DAP*) e altura total (*Ht*) aos três anos de idade após o plantio em cada população. A população Melhoramento 1 se encontra na cidade de Lençóis Paulistas e a Melhoramento 2 na cidade de Anhembi.

As populações de introdução foram identificadas como A08A 032, A08B 033, A08E 066 e A08G 068 e cada população é representada pelos 25 melhores indivíduos. Para facilitar a análise no *software* R v.4.0.5 (R CORE TEAM, 2021) e visualização dos resultados, essas populações foram renomeadas de Usp32a, Usp33a, Usp66a e Usp68a, respectivamente. Os indivíduos que compõem essas populações foram consideradas naturais, pois foram coletadas em seus respectivos habitats. As localizações geográficas das populações naturais encontram-se na Figura 1.

Os clones comerciais são identificados como AEC144, CNB10, FIB0075, GG100, H13, IPB15, JAR2646, KLA2864, VER361 e VM04. Cabe ressaltar que esses clones podem ser constituído por híbridos.

Figura 1 - Localização geográfica das populações naturais com marcadores com diferentes cores



Fonte: Do autor (2021)

4.2. Extração de material genético

Obteve-se o material genético das árvores por meio de suas folhas frescas. A extração de DNA foi realizada com protocolo baseado no detergente CTAB e a amplificação dos marcadores microssatélites de acordo com o protocolo descrito por Faria et al. (2010, 2011). A detecção da fluorescência foi realizada via eletroforese capilar com a plataforma automática ABI 3100 XL, também conforme Faria et al. (2010, 2011).

4.3. Marcadores microssatélites

Foram utilizados 19 marcadores microssatélites: EMBRA2; EMBRA3; EMBRA10; EMBRA11; EMBRA12; EMBRA21; EMBRA28; EMBRA32; EMBRA38; EMBRA45; EMBRA63; EMBRA128; EMBRA157; EMBRA204; EMBRA210;

EMBRA681; EMBRA915; EMBRA1144; e EMBRA1349. Esses marcadores foram obtidos e descritos por Grattapaglia *et al.* (2015).

4.4. Diversidade e estrutura genética

Os parâmetros de diversidade genética foram obtidos por meio da genotipagem dos indivíduos. A genotipagem baseada nos 19 microssatélites, foi registrada em uma planilha, gentilmente fornecida pelo Instituto de Pesquisas e Estudos Florestais (IPEF).

Para a estimativa dos parâmetros genéticos, foi utilizado o programa de análises estatísticas R com a utilização do pacote *adegenet* v.2 (JOMBART, 2008). Com esse pacote foram obtidos o número de alelos por loco (A); o número de alelos privados (PA); os valores de heterosigozidade observada (H_o) e esperada sob equilíbrio de Hardy-Weinberg (H_e) por loco e população. A estimativa da riqueza alélica (AR) foi realizada com o pacote *hierfstat* v. 0.5.7.

Para o cálculo das estatísticas F de Wright foi utilizado o pacote *hierfstat* v. 0.5.7 e, a interpretação do valor de F_{st} é feita de acordo com a definição de Wright (HARTL & CLARK, 2010), da seguinte forma: valores de 0 a 0,05 representam uma pequena diferenciação genética; de 0,05 a 0,15 indicam uma moderada diferenciação genética; de 0,15 a 0,25 uma grande diferenciação genética e acima de 0,25 significa uma diferenciação genética muito grande. Já em relação ao F_{is} , valores iguais a 0 representam uma população sem endogamia; valores iguais a 1 representam uma população totalmente endogâmica e valores negativos representam a redução da homozigose para o loco de estudo.

As análises de distância e estruturação genética foram realizadas com as populações naturais separadas do conjunto de dados total, visando o entendimento da estrutura genética dessas populações. As distâncias genéticas entre as árvores foram visualizadas por meio de dendrograma e gráfico com os dois primeiros componentes do PCA. Já a análise de estruturação genética entre as árvores foi realizada por meio do programa Structure.

As distâncias entre as populações foram estimadas pela distância genética de Nei (1983) com o pacote *poppr*, v. 2.9.3. Para construção do dendrograma, a matriz de distâncias foi analisada pelo método de agrupamento UPGMA (Unweighted Pair Group Average). As consistências dos nós do dendrograma foram obtidas por meio

de 10000 bootstrap. Dessa forma, foram gerados dendrogramas com as distâncias entre as populações (Figura 4), bem como entre todos os indivíduos (Apêndice A) e somente entre os indivíduos das populações naturais (Apêndice B).

Além disso, a distância entre as árvores também foi avaliada por meio de um gráfico gerado a partir da análise de componentes principais (PCA). Essa análise foi realizada com o pacote *hierfstat*, que utiliza a matriz de frequências alélicas individuais.

Por último, foram realizadas análises de estruturação genética com o software Structure (PRITCHARD *et al*, 2000), com o intuito de calcular o número de subpopulações (K) que melhor explicam os dados genotípicos. O programa Structure utiliza a inferência Bayesiana para alocar os indivíduos nas subpopulações, com a premissa de que os grupos formados se aproximem ao máximo do Equilíbrio de Hardy-Weinberg. Para isso, as análises foram realizadas com o burn-in de 100.000 Cadeias de Markov e Monte Carlo (MCMC), 100.000 MCMC após o burn-in e com 5 repetições para cada valor de K , com K variando de 1 a 10. As análises foram realizadas para o conjunto de dados total e separadamente para as populações naturais.

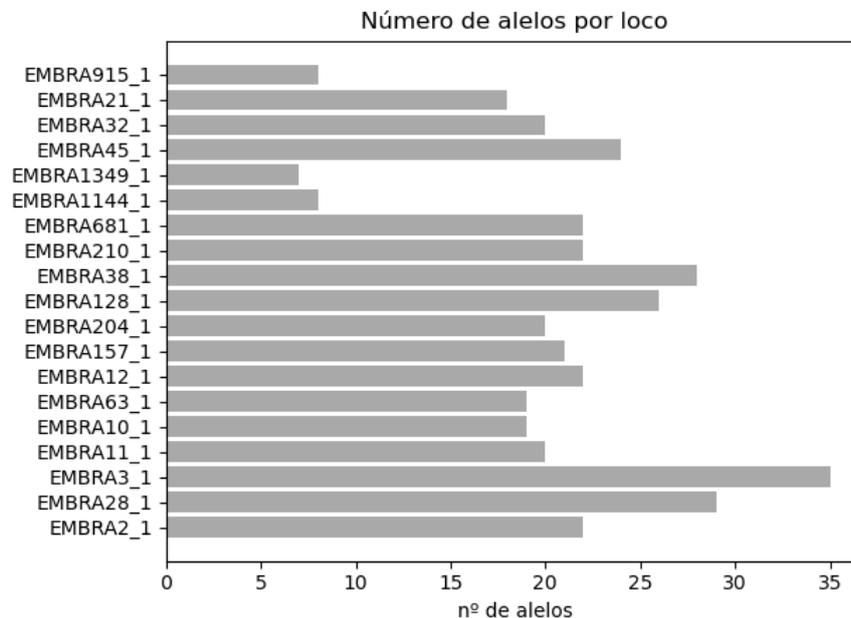
O programa Structure não indica qual é o número de subpopulações (K) que melhor explica o conjunto de dados. Para isso, outras análises são necessárias. A metodologia utilizada para determinação do número de clusters utilizada foi aquela implementada no programa Structure Selector (Li YL, 2018), que se baseia nos índices definidos por Puechmaille (2016).

5. RESULTADOS E DISCUSSÃO

5.1. Diversidade genética

Para cada marcador microssatélite foi determinado o número de alelos por loco, como pode ser visto na Figura 2. Os valores obtidos variaram de sete alelos para o loco EMBRA1349_1 a 35 alelos para o marcador EMBRA3_1.

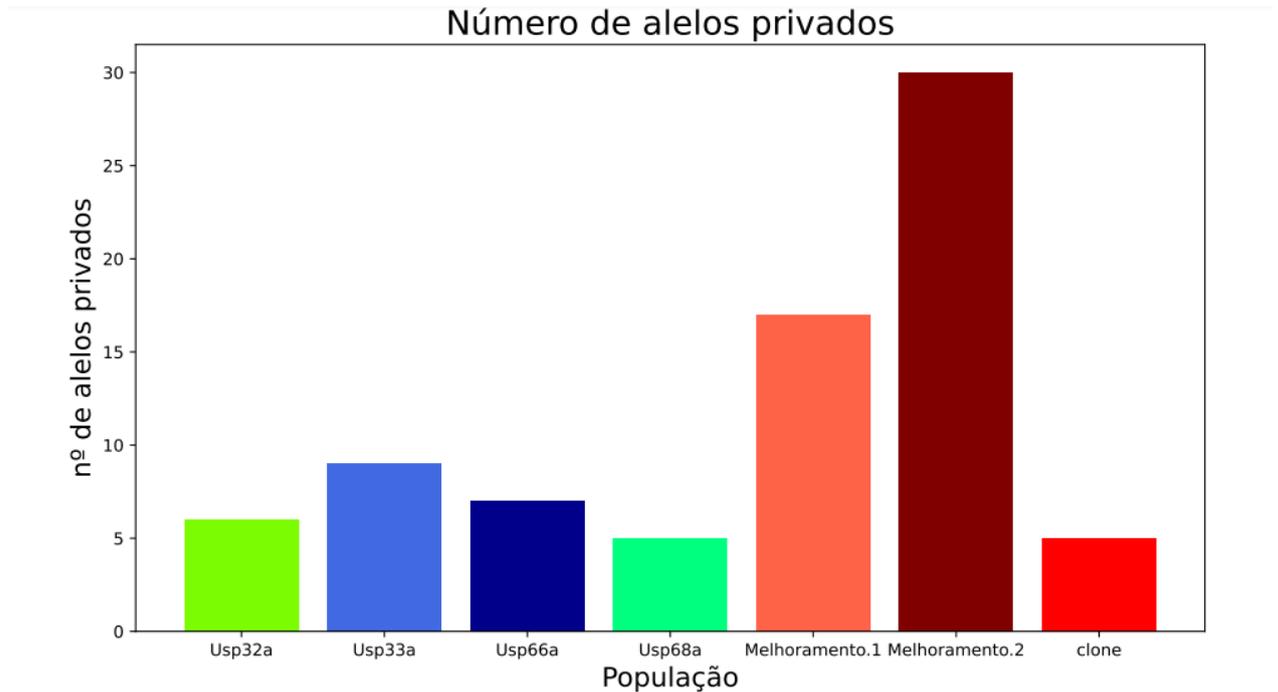
Figura 2 - Número total de alelos por loco



Fonte: Do autor (2021)

Além disso, determinou-se também o número de alelos privados para cada população. Os dados resumidos na Figura 3, indicam que as populações representadas pelos clones comerciais e Usp68a possuem a menor quantidade de alelos exclusivos entre as sete populações, com apenas cinco. Por outro lado, a população Melhoramento 2 conta com a maior quantidade, com 30 alelos.

Figura 3 - Número de alelos privados por população.



Fonte: Do autor (2021)

A diversidade genética das populações foi comparada com diversas estimativas, como a média de heterozigosidade esperada no equilíbrio de Hardy-Weinberg; a média de heterozigosidade observada; o número de alelos total; a riqueza alélica e o número de alelos privados (tabela 1). A maior heterozigosidade esperada foi encontrada na população de melhoramento Melhoramento 2, com 0,8563 e o menor na Usp32a com 0,7445. Já com relação à heterozigosidade observada, os valores variam de 0,8368 na população de clones a 0,6716 na população Usp68a. Quanto ao número total de alelos encontrados em cada população, o menor valor foi de 149 na Usp32a e o maior foi 310 na Melhoramento 2. Os valores de riqueza alélica foram estimados tomando como base o número de dez indivíduos, pois é o tamanho da menor população do conjunto. Para esse índice as populações dos clones e a Melhoramento 2 obtiveram a maior média, já o menor valor foi encontrado para a Usp32a.

De maneira geral, observou-se uma diversidade genética grande em todas as unidades populacionais quando comparada ao que já foi visto para outras espécies de eucalipto (Mora *et al*, 2016; Jones *et al*, 2002), com um número de alelos alto para todos os lócos estudados (valores de 7 a 35), com He média de 0,8165 e Ho média

de 0,7456. Em estudos pregressos com *E. urophylla*, resultados semelhantes foram obtidos em Tripiana *et al* (2007), em que foram utilizados dez marcadores microsatélites e obteve-se valores de 10 a 29 alelos por loco e *He* variando de 0,51 a 0,72. Já nos estudos de Payn *et al* (2008), utilizou-se doze marcadores microsatélites, e foram encontrados de 6 a 56 alelos por marcador e a *Ho* entre 0,737 e 0,610. Nesse mesmo estudo percebeu-se que a população localizada em Timor possuía a maior taxa de diversidade, resultado que se repetiu neste trabalho, uma vez que a população Usp33a é proveniente da mesma localidade e apresentou a maior diversidade genética dentre as populações naturais. Os autores atribuíram essa maior diversidade ao grande tamanho da população e a possibilidade de ter sido a primeira ilha a ser colonizada. Em populações mais antigas a diversidade tende a ser maior (Payn *et al*, 2007).

Já Tripiana *et al* (2007), atribuíram essa alta taxa de diversidade à história evolutiva da espécie, ao grande tamanho efetivo, bem como pelas características reprodutivas da população, uma vez que o pólen e as sementes dessa espécie podem viajar por longas distâncias mediadas por animais (Hamrick *et al*. 1992; McGoldrick e Mac Nally, 1998; Tripiana *et al*, 2007).

Tabela 1 - Métricas de diversidade estimadas por população.

POPULAÇÃO	n	He	Ho	A	RA	PA
Usp32a	25	0,7445	0,7411	149	6,01	6
Usp33a	25	0,8178	0,7389	217	8,15	9
Usp66a	25	0,8166	0,7200	206	8,04	7
Usp68a	25	0,7961	0,6716	180	7,22	5
Melhoramento 1	65	0,8544	0,7360	281	8,66	17
Melhoramento 2	79	0,8563	0,7748	310	8,84	30
Clones	10	0,8297	0,8368	170	8,95	5
Média	36,28	0,8165	0,7456	216	7,98	11,28

Em que **n**: tamanho da amostra; **He**: Heterosigozidade esperada média sob equilíbrio de Hardy-Weinberg; **Ho**: Heterosigozidade observada média; **A**: Número de alelos; **RA**: riqueza alélica média; **PA**: Alelos privados.

Fonte: Do autor (2021)

Entretanto, no que se refere aos parâmetros como a quantidade de alelos privados; a heterozigiosidade esperada e observada; a quantidade de alelos total, bem como a riqueza alélica, os índices das populações de melhoramento (Melhoramento 1 e Melhoramento 2) mostraram-se ligeiramente maiores que o das populações

naturais (Usp32a, Usp33a, Usp66a e Usp68a), o que indica uma maior diversidade genética nas populações de melhoramento. Mesmo utilizando o método de rarefação, com o intuito de reduzir a interferência do tamanho da amostra, os valores da riqueza alélica foram maiores nas populações de melhoramento. Em um estudo recente, com as mesmas duas populações de melhoramento (Melhoramento 1 e Melhoramento 2) e com os mesmos locos microssatélites, Silva *et al* (2019) chegaram a conclusões semelhantes quando compararam aos resultados obtidos por Payn *et al* (2008) para populações naturais.

Esse resultado pode ser devido ao fato do melhoramento genético de *E. urophylla* ainda estar em seus estágios iniciais e, portanto ser muito parecida com as populações naturais, como aponta Silva *et al* (2019). Além disso, as populações de melhoramento são geralmente formadas por uma mistura de procedências de várias populações naturais diferentes, o que aumenta a diversidade genética (SILVA *et al*, 2019).

5.2. Distância e estrutura genética

Com a estatística F , determinou-se o F_{st} médio com o valor de 0,030 e o valor de F_{is} médio de 0,101, como mostra a tabela 2.

Tabela 2 - Parâmetros por loco obtidos com o pacote hierfstat.

LOCO	Ho	He	Ht	Dst	Fst	Fis
EMBRA2	0.576	0.691	0.711	0.021	0.029	0.166
EMBRA28	0.860	0.907	0.923	0.016	0.017	0.052
EMBRA3	0.872	0.890	0.911	0.021	0.023	0.020
EMBRA11	0.822	0.869	0.880	0.011	0.012	0.054
EMBRA10	0.857	0.851	0.872	0.021	0.024	-0.007
EMBRA63	0.772	0.810	0.836	0.027	0.032	0.046
EMBRA12	0.882	0.898	0.911	0.013	0.015	0.018
EMBRA157	0.601	0.885	0.905	0.020	0.023	0.321
EMBRA204	0.862	0.883	0.896	0.013	0.014	0.024
EMBRA128	0.536	0.870	0.900	0.030	0.033	0.385
EMBRA38	0.891	0.847	0.904	0.057	0.063	-0.052
EMBRA210	0.915	0.883	0.898	0.015	0.017	-0.036
EMBRA681	0.680	0.828	0.874	0.046	0.052	0.179
EMBRA1144	0.729	0.738	0.775	0.037	0.047	0.013
EMBRA1349	0.694	0.719	0.725	0.007	0.009	0.035
EMBRA45	0.516	0.899	0.918	0.019	0.021	0.427
EMBRA32	0.681	0.826	0.856	0.029	0.034	0.176
EMBRA21	0.741	0.895	0.913	0.018	0.020	0.172
EMBRA915	0.681	0.686	0.750	0.064	0.086	0.006
Média	0.746	0.836	0.861	0.026	0.030	0.101

Em que *Ho*: Heterosigozidade observada média; *He*: Heterosigozidade esperada média; *Ht*: Heterosigozidade esperada total; *Dst*: Diversidade genética entre as amostras; *Fst*: índice de fixação entre populações; *Fis*: Índice de fixação entre indivíduos.

Fonte: Do autor (2021)

Dessa forma, o valor encontrado de *Fst*, indica uma baixa estruturação, ou seja, que não há diferenciação genética suficiente entre as populações para que elas sejam consideradas geneticamente distintas. Esse resultado já foi obtido para *E. urophylla* por Payn *et al* (2008) bem como por Tripiana *et al* (2007) (*Fst* = 0,04). Esses resultados indicam que pode estar havendo fluxo gênico entre as quatro subpopulações naturais, apesar de algumas se localizarem em ilhas diferentes. Além disso, já foram relatados baixos índices de *Fst* para outras espécies de eucalipto, como para *E. pauciflora* (*Fst* = 0,03, Gauli *et al*, 2014) e para *E. obliqua* (*Fst* = 0,015, Blomfield, 2011). Ademais, o valor de *Fis* indica uma baixa taxa de endocruzamentos.

Com a distância de Nei (tabela 3) é possível observar que, entre as todas as populações naturais, Usp32a e Usp33a são as mais distantes tanto geneticamente quanto geograficamente. Já os clones comerciais possuem as maiores distâncias genéticas com as outras populações, encontrando-se acima de 0.30. A população

Melhoramento 2, por sua vez possui as menores distâncias com as populações Melhoramento 1 e com os clones. A população da Melhoramento 1, encontra-se mais próxima das populações de introdução do que dos clones, com exceção da Usp32a.

Com o dendrograma da figura 4, é possível visualizar com mais clareza as relações de proximidade genética entre as populações. De maneira geral, há uma correlação entre a distância geográfica com a distância genética. Os resultados quanto às populações naturais estão de acordo com os resultados obtidos por Payn *et al.* (2008), em que as populações presentes no Timor (Usp33a) apresentam maior diferenciação genética em relação às populações das demais ilhas.

Tabela 3 - Distância de Nei entre cada população de estudo.

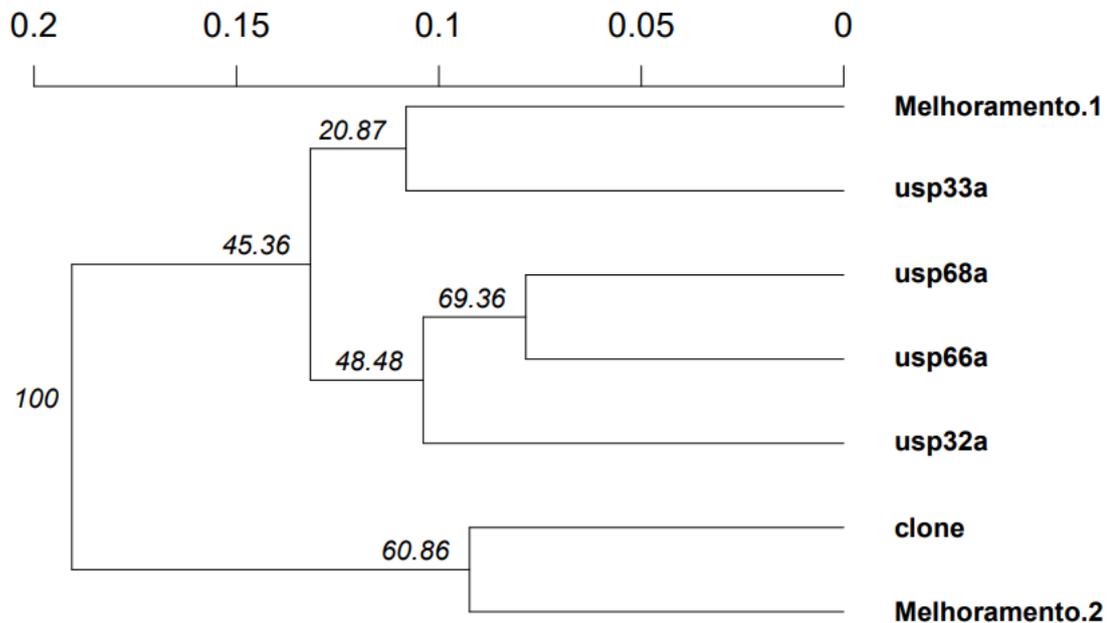
POPULAÇÃO	Usp32a	Usp33a	Usp66a	Usp68a	Melhor. 1	Melhor. 2
Usp33a	0.310					
Usp66a	0.207	0.173				
Usp68a	0.208	0.246	0.157			
Melhoramento 1	0.380	0.216	0.233	0.238		
Melhoramento 2	0.449	0.283	0.292	0.389	0.197	
clone	0.641	0.429	0.368	0.453	0.312	0.185

Fonte: Do autor (2021)

Esses resultados indicam que a população Melhoramento 2 é um intermediário entre a população Melhoramento 1 e os clones comerciais.

Com base na distância genética de Nei, também foram gerados dendrogramas para as árvores. Observa-se pelo dendrograma contendo apenas as populações naturais (Apêndice B), que a distância genética entre os indivíduos não é o suficiente para discriminá-los em grupos correspondentes às populações de origem. Esse resultado reforça a possibilidade de haver trocas de material genético (fluxo gênico) entre elas. Para todo o conjunto de dados, a separação dos indivíduos nas populações de origem também não foi possível, o que corrobora com a inferência de que as populações de melhoramento ainda se encontram muito próximas às populações naturais.

Figura 4 - Dendrograma baseado na distância de Nei das populações de estudo.



Fonte: Do autor (2021)

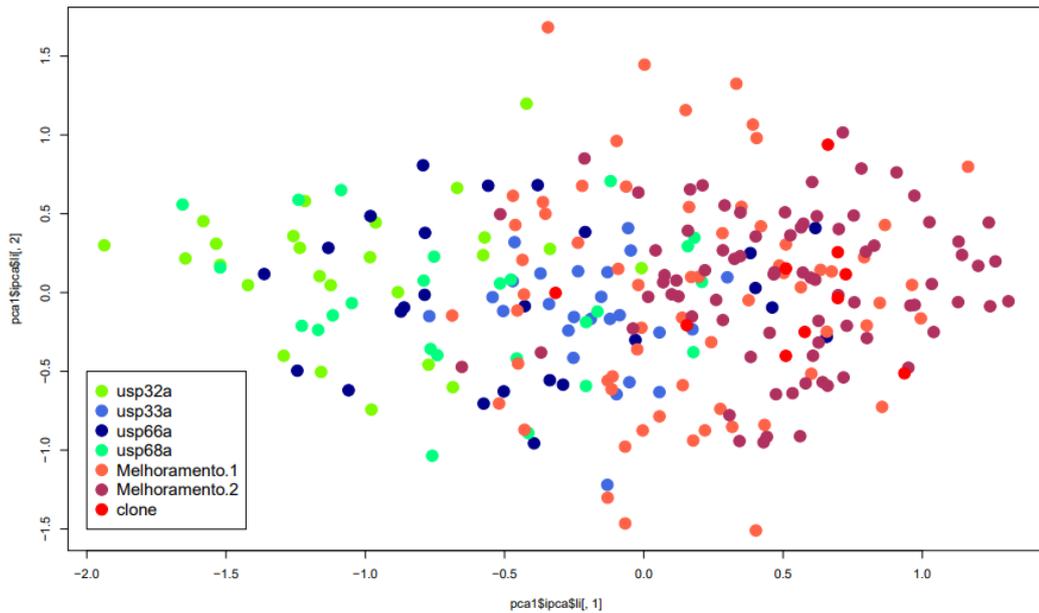
Vale ressaltar que há apenas um nó com consistência considerável (100%) nas análises de bootstrap, que divide os clones e Melhoramento 2 das outras populações, para os demais nós a consistência foi baixa. Essa baixa consistência nos nós do dendrograma é mais um indicativo de que as populações não podem ser separadas em grupos distintos com consistência. Isso também reflete a baixa diferenciação genética entre as populações, indicando a ocorrência de fluxo gênico entre elas.

Quando se observa o gráfico da figura 5, nota-se mais uma vez que não é possível distinguir as populações por meio da PCA. No gráfico com os dois primeiros componentes, não há uma separação muito clara entre as populações, o que era esperado, uma vez que o índice de fixação F_{st} geral também é muito baixo (0.03).

No entanto, pode-se observar uma certa similaridade com o dendrograma das populações (Figura 4), em que nos extremos encontram-se indivíduos das populações Usp32a e Melhoramento 2. Quando se observa a Figura 6, também não é possível realizar a distinção entre as populações naturais, entretanto, consegue-se notar que Usp32a se encontra mais distante de Usp33a. De fato, essas também são as duas

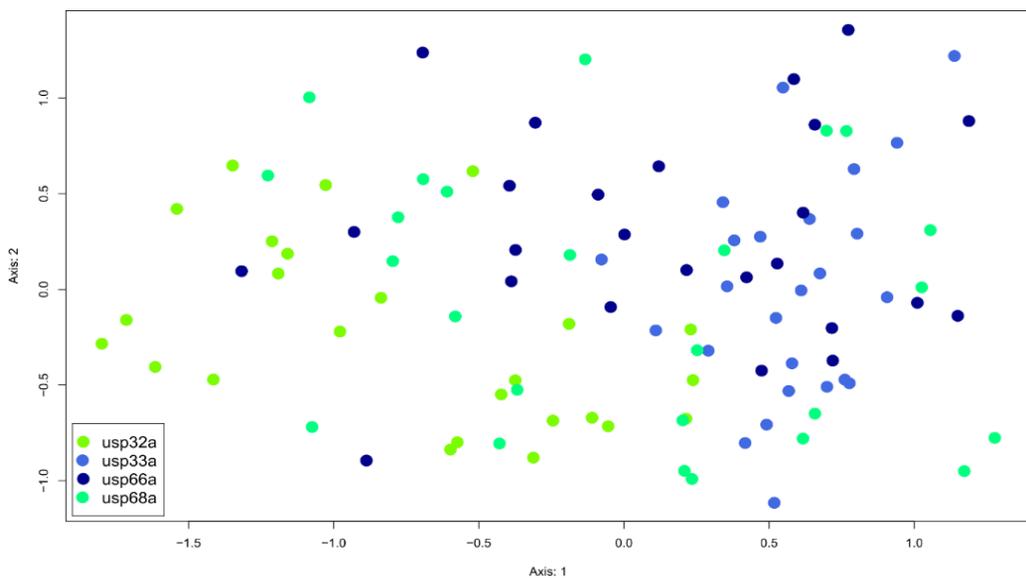
populações geograficamente mais distantes (Figura 1), o que dificulta o fluxo gênico entre elas.

Figura 5 - Análise de componentes principais (PCA) da frequência individual de alelos de todos os indivíduos.



Fonte: Do autor (2021)

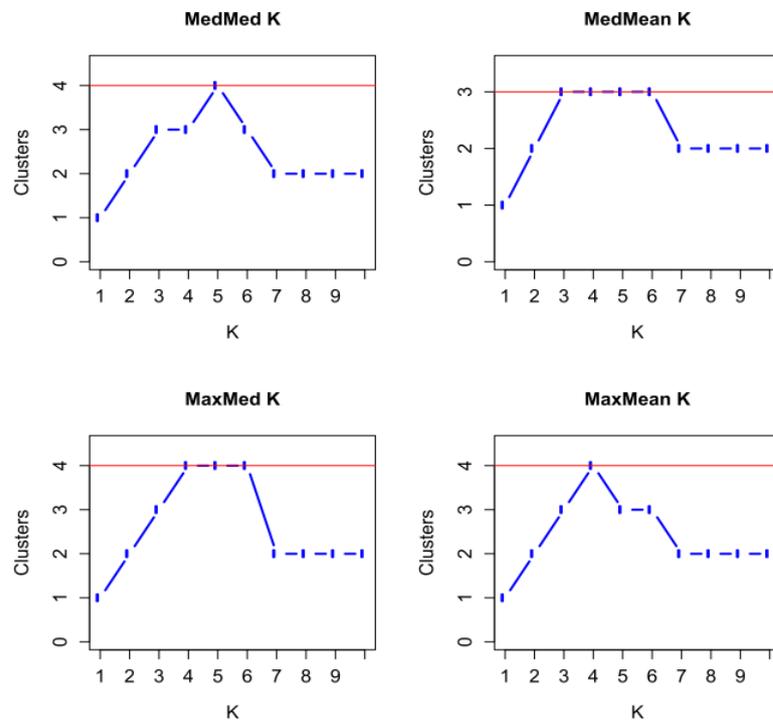
Figura 6 - Análise de componentes principais (PCA) da frequência individual de alelos das populações naturais.



Fonte: Do autor (2021)

A análise dos resultados do *Structure*, com o programa *Structure Selector*, indica o número de clusters que explica o conjunto total dos dados é de $K=4$, o que significa que há possivelmente quatro grupos geneticamente distintos. Já quando se leva em consideração apenas as populações naturais, o valor de K encontrado foi o de dois grupos, como mostram as figuras 7 e 8.

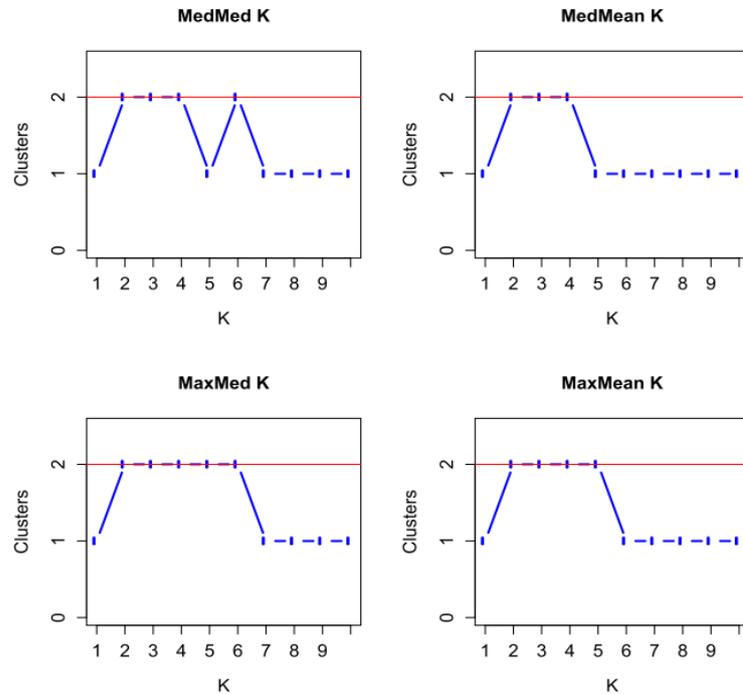
Figura 7- Valores das métricas de *MedMeanK*, *MaxMeanK*, *MaxMedK* e *MedMedK* para todo o conjunto de dados



A imagem indica que o modelo com $K=4$ é o que melhor explica os dados.

Fonte: Do autor (2021)

Figura 8 - Valores das métricas de *MedMeanK*, *MaxMeanK*, *MaxMedK* e *MedMedK* para todo as populações naturais,

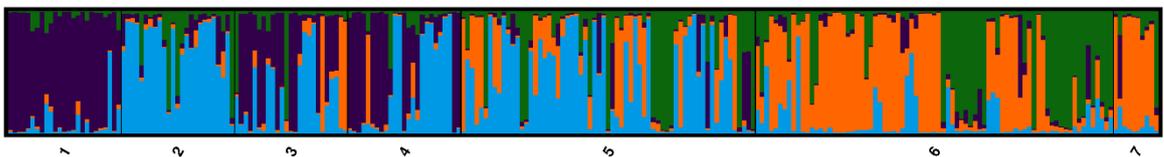


A imagem indica que o modelo com $K=2$ é o que melhor explica os dados.

Fonte: Do autor (2021)

Ao definir o valor de K , realizou-se a plotagem de um gráfico de barras com o percentual de quanto de cada cluster há em cada indivíduo. O gráfico para todo o conjunto de dados e para as populações naturais encontram-se nas Figuras 9 e 10, respectivamente.

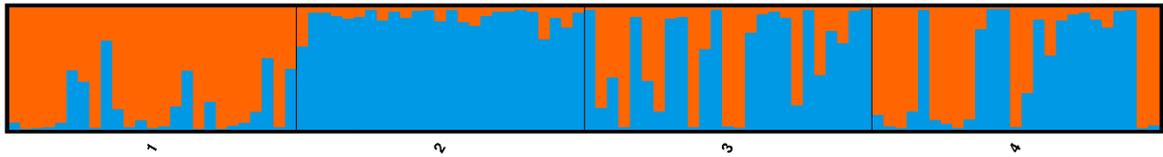
Figura 9 - Gráfico de barras da proporção de correspondência de cada indivíduo a cada cluster ($K=4$) para todo o conjunto de dados.



Em que 1:Usp32a, 2:Usp33a,3:Usp66a,4:Usp68a,5:Melhoramento 1, 6:Melhoramento 2, 7:Clones.

Fonte: Do autor (2021)

Figura 10 - Gráfico de barras da proporção de correspondência de cada indivíduo a cada cluster ($K=2$) para as populações naturais.



Em que 1:Usp32a, 2:Usp33a,3:Usp66a,4:Usp68a.

Fonte: Do autor (2021)

Com esses gráficos, percebe-se que apesar do *Structure* definir os agrupamentos genéticos, em nenhum dos dois há uma divisão consistente que indique a existência de populações muito distintas umas das outras. Entretanto, percebe-se que os resultados indicam que as populações Usp66a e Usp68a, aparentam ser intermediárias entre Usp33a e Usp32a, que são as mais distantes. Esse resultado é consistente com a distância geográfica, cuja maior distância ocorre entre as populações Usp33a e Usp32a, com as outras duas populações estando entre elas (Figura 1). A população de melhoramento da empresa Melhoramento 1 se encontra entre a Melhoramento 2 e as populações naturais. Já a população composta pelos clones, por sua vez, é mais próxima da Melhoramento 2 do que de Melhoramento 1. Esses resultados são consistentes com os observados no dendrograma da Figura 4.

Com essa última análise, percebe-se que, mais uma vez há uma divisão semelhante as já encontradas anteriormente, tanto com relação à estatística de Wright, distância de Nei, quanto a PCA, corroborando com esses demais resultados. Desse modo, também com os resultados do programa *Structure*, percebe-se que não é possível dividi-las em populações diferentes.

Os resultados obtidos, vão contra ao que se espera para populações divididas em ilhas distintas, uma vez que há uma barreira formada por oceanos, o que atrapalha o fluxo gênico, especialmente a dispersão de sementes por meio da gravidade e favorece a diferenciação dos agrupamentos por deriva genética (Bottin *et al*, 2005; Payn *et al* 2007). Todavia, Tripliana *et al* (2007), apontam que esse fato pode ser explicado por migração de material genético (dispersão de pólen) mediado por insetos transportados por meio de ventos, ocorrer em uma frequência suficiente para evitar a diferenciação e reduzir o efeito da deriva. Além disso, as distâncias entre as ilhas são

inferiores a 35 km e um estudo de Rutherford *et al* (2018), relataram a existência de fluxo gênico para outras espécies de eucalipto até uma distância de 70 km.

Ademais, Payn *et al* (2008), calculou o número de migrantes e verificou um valor menor do que o das populações inseridos na mesma ilha, mas também apontou para a possibilidade de ter ocorrido o transporte de pólen, por animais levados pelo vento numa frequência suficiente para evitar a diferenciação. Isso explicaria o motivo das populações mais distantes geograficamente serem as mais distantes geneticamente, pois o fluxo gênico reduz a medida que a distância aumenta.

6. CONSIDERAÇÕES FINAIS

Conhecer as características genéticas de uma espécie florestal de interesse comercial é a base para os programas de melhoramento. *E.urophylla* é uma dessas espécies de interesse do setor florestal e o presente estudo melhora o entendimento dos níveis de diversidade e estruturação genética em populações naturais e de melhoramento da espécie.

Com este trabalho foi possível perceber que há uma grande diversidade genética dentro das subpopulações e que as distâncias encontradas não são tão significativas a ponto de ser possível separá-las em populações distintas. Isso porque não há estrutura genética evidente entre as populações, o que sugere um curto tempo de melhoramento nas populações de introdução e a existência de fluxo gênico entre as populações naturais. Apesar de estarem ilhas separadas, as populações naturais parecem estar conectadas por fluxo gênico, o que pode ser explicado pela proximidade das ilhas da Indonésia, onde as populações foram amostradas.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ASSIS, T.; WARBURTON, P.; HARWOOD, C. **Artificially induced protogyny: an advance in the controlled pollinization of Eucalyptus**. *Australian Forestry, Canberra*, v. 68, n. 1, p. 27-33, 2004.

BALDONI, Aisy Botega et al. **Genetic diversity of Brazil nut tree** (*Bertholletia excelsa* Bonpl.) in southern Brazilian Amazon. *Forest Ecology and Management*, v. 458, p. 117795, 2020.

BLOOMFIELD, Justin A. et al. **Molecular genetic variation in a widespread forest tree species Eucalyptus obliqua** (Myrtaceae) on the island of Tasmania. *Australian Journal of Botany*, v. 59, n. 3, p. 226-237, 2011.

BRASILEIRO, Serviço Florestal. **Instituto de Pesquisa Ambiental da Amazônia, Florestas nativas de Produção Brasileiras** (Report, Brasília, IPAM, 2011).

BOTTIN, Lorraine et al. **Genetic diversity and population structure of an insular tree, Santalum austrocaledonicum in New Caledonian archipelago**. *Molecular ecology*, v. 14, n. 7, p. 1979-1989, 2005.

CHEZHIAN, P.; YASODHA, R.; GHOSH, M. **Genetic diversity analysis in a seed orchard of Eucalyptus tereticornis**. *New Forests, Heidelberg*, v.40, n.1, p.85-99, 2010.

COLLEVATTI, R. G.; GRATTAPAGLIA, D.; HAY, J. D. **Population genetic structure of the endangered tropical tree species Caryocar brasiliense, based on variability at microsatellite loci**. *Molecular Ecology*. v. 10, 2001.

COSTA, Joana et al. **Microsatellite analysis of population structure in Eucalyptus globulus**. *Genome*, v. 60, n. 9, p. 770-777, 2017.

CRUZ, Cosme Damião; FERREIRA, Fábio Medeiros; PESSONI, Luiz Alberto. **Biometria aplicada ao estudo da diversidade genética**. 2. Ed. Visconde do Rio Branco: Suprema, v. 620, 2020.

DE ALMEIDA, Eduardo J. et al. **Reação de clones de eucalipto a *Meloidogyne enterolobii***. 2012.

DVORAK, W. S.; HODGE, G. R.; PAYN, K. G. **The conservation and breeding of *Eucalyptus urophylla*: a case study to better protect important populations and improve productivity**. Southern Forests: a journal of forest science, Londres, v. 70, n. 2, p.77-85, 2008.

ELDRIDGE, R, J. DAVIDSON, C. HARWOOD and G. VAN WYK, 1993 **Eucalypt Domestication and Breeding**. Oxford Science Publications, Clarendon Press, Oxford 1994.

ELLEGREN, Hans. **Microsatellites: simple sequences with complex evolution**. Nature reviews genetics, v. 5, n. 6, p. 435-445, 2004.

FALEIRO, F. G. **Marcadores-genético moleculares aplicados a programas de conservação e uso de recursos genéticos**. Planaltina (DF): Embrapa, 2007. 102 p.

Faria, D.A., Mamani, E.M.C., Pappas, M.R., Pappas, G.J., and Grattapaglia, D. 2010. **A selected set of EST-derived microsatellites, polymorphic and transferable across 6 species of *Eucalyptus***. J. Hered. 101: 512–520.

Faria, D.A., Mamani, E.M.C., Pappas, G.J., and Grattapaglia, D. 2011. **Genotyping systems for *Eucalyptus* based on tetra-, penta-, and hexanucleotide repeat EST microsatellites and their use for individual fingerprinting and assignment tests**. Tree Genet. Genomes, v.7 n. 1, p. 63–77, 2011.

FERREIRA, M.; DOS SANTOS, P. E. T. **Genetic improvement of *Eucalyptus* in Brazil-brief review and perspectives**. In: IUFRO Conference on Silviculture and Improvement of Eucalypts, Salvador (Brazil), 24-29 Aug 1997. Embrapa, 1997.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3. ed. Brasília: Embrapa, 1998. 220 p

FLORES, Thiago Bevilacqua et al. **Eucalyptus no Brasil: zoneamento climático e guia para identificação**. Piracicaba: IPEF, 2016.

GAULI, Archana et al. **Molecular genetic diversity and population structure in Eucalyptus pauciflora subsp. pauciflora (Myrtaceae) on the island of Tasmania**. Australian Journal of Botany, v. 62, n. 3, p. 175-188, 2014.

GANDARA, F. B. **Diversidade genética de populações de cedro** (*Cedrela fissilis* Vell. Meliaceae) no Centro-Sul do Brasil. 2009. 88 f. Tese (Doutorado em Ciências) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, 2009.

GOVINDARAJ, Mahalingam; VETRIVENTHAN, Mani; SRINIVASAN, M. **Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives**. Genetics research international, v. 2015, 2015.

GRATTAPAGLIA, Dario et al. **A novel genome-wide microsatellite resource for species of Eucalyptus with linkage-to-physical correspondence on the reference genome sequence**. Molecular ecology resources, v. 15, n. 2, p. 437-448, 2015.

GROVER, Atul; SHARMA, P. C. **Development and use of molecular markers: past and present**. Critical reviews in biotechnology, v. 36, n. 2, p. 290-302, 2016.

GUIMARÃES, Rejane Araújo et al. **Assessing genetic diversity and population structure in a *Dipteryx alata* germplasm collection utilizing microsatellite markers**. Crop Breeding and Applied Biotechnology, v. 19, n. 3, p. 329-336, 2019.

HAMRICK, James L.; GODT, Mary Jo W.; SHERMAN-BROYLES, Susan L. **Factors influencing levels of genetic diversity in woody plant species.** In: **Population genetics of forest trees.** Springer, Dordrecht, 1992. p. 95-124.

HAMRICK, J. L.; GODT, MJ W. Effects of life history traits on genetic diversity in plant species. **Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences**, v. 351, n. 1345, p. 1291-1298, 1996.

HARTL, D. L.; CLARK, A. G. **Princípios de genética de populações.** 4. ed. Artmed: Porto Alegre, 2010. 660 p.

HOUSE, A. P. N.; BELL, J. C. **Isozyme variation and mating system in Eucalyptus urophylla.** *Silvae Genet*, v. 43, p. 167-176, 1994.

IBÁ, IBÁ **Relatório Anual.** Ano base 2018. 2019. Disponível em ><https://iba.org/datafiles/publicacoes/relatorios/iba-relatorioanual2019.pdf><, acesso em 05 de novembro de 2021.

Jerome Goudet and Thibaut Jombart (2020). hierfstat: **Estimation and Tests of Hierarchical F-Statistics.** R package version 0.5-7. <https://CRAN.R-project.org/package=hierfstat>

Jombart, T. (2008) **adegenet: a R package for the multivariate analysis of genetic markers.** *Bioinformatics* 24: 1403-1405. doi: 10.1093/bioinformatics/btn129.

JONES, Rebecca C. et al. **Microsatellite and morphological analysis of Eucalyptus globulus populations.** *Canadian Journal of Forest Research*, v. 32, n. 1, p. 59-66, 2002.

KALIA, R.K.; RAI, M.K.; KALIA, S.; SINGH, R.; DHAWAN, A.K. **Microsatellite markers: an overview of the recente progress in plants.** *Euphytica*, Dordrecht, v. 177, n. 3, p. 309-334, 2011.

Kamvar ZN, Tabima JF, Grünwald NJ. (2014) **Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction.** PeerJ 2:e281. doi: 10.7717/peerj.281

Kamvar ZN, Brooks JC and Grünwald NJ (2015) **Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality.** Front. Genet. 6:208. doi: 10.3389/fgene.2015.00208

LOVELESS, M. D.; HAMRICK, J. L. **Ecological determinants of genetic structure in plant populations.** Annual Review of Ecology and Systematics. v. 15, 1984.

LI, Yu-Long; LIU, Jin-Xian. **StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods.** Molecular ecology resources, v. 18, n. 1, p. 176-177, 2018.

LU, Wanhong et al. **Genetic diversity and structure through three cycles of a Eucalyptus urophylla ST Blake breeding program.** Forests, v. 9, n. 7, p. 372, 2018.

LV, Jiabin et al. **Genetic diversity analysis of a breeding population of Eucalyptus cloeziana F. Muell.(Myrtaceae) and extraction of a core germplasm collection using microsatellite markers.** Industrial Crops and Products, v. 145, p. 112157, 2020

MCGOLDRICK, Joshua M.; MAC NALLY, R. **Impact of flowering on bird community dynamics in some central Victorian eucalypt forests.** Ecological Research, v. 13, n. 2, p. 125-139, 1998.

MELO, A. T. O. **Fluxo gênico e estrutura genética espacial de *Cabralea canjerana* (Vell.) Mart. (Meliaceae) em fragmentos florestais de Mata Atlântica.** 2012. 87 f. Dissertação (Mestrado em Genética e Melhoramento de Plantas) – Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Goiânia, 2012.

MOTAMAYOR, Juan Carlos et al. **Geographic and genetic population differentiation of the amazonian chocolate tree (*Theobroma cacao* L.),** 2009.

MORA, Freddy et al. **Genetic diversity and population structure of a drought-tolerant species of Eucalyptus, using microsatellite markers**. Journal of Plant Biochemistry and Biotechnology, v. 26, n. 3, p. 274-281, 2017.

NEI, Masatoshi; TAJIMA, Fumio; TATENO, Yoshio. **Accuracy of estimated phylogenetic trees from molecular data**. Journal of molecular evolution, v. 19, n. 2, p. 153-170, 1983.

NICOLLE, D. **Classification of the eucalypts (Angophora, Corymbia and Eucalyptus)** Version 2. 2015. 56p. Disponível em: <<http://www.dn.com.au/Classification-Of-The-Eucalypts.pdf>>. Acesso em: 11 nov. 2020.

PAYN, Kitt G.; DVORAK, William S.; MYBURG, Alexander A. **Chloroplast DNA phylogeography reveals the island colonisation route of Eucalyptus urophylla (Myrtaceae)**. Australian Journal of Botany, v. 55, n. 7, p. 673-683, 2007.

PAYN, Kitt G. et al. **Microsatellite diversity and genetic structure of the commercially important tropical tree species Eucalyptus urophylla, endemic to seven islands in eastern Indonesia**. Tree Genetics & Genomes, v. 4, n. 3, p. 519-530, 2008.

PORTH, I.; EL-KASSABY, Y.A. **Assessment of the genetic diversity in forest tree populations using molecular markers**. Diversity, Basiléia, v. 6, n. 2, p. 283-295, 2014.

PRITCHARD, Jonathan K.; STEPHENS, Matthew; DONNELLY, Peter. **Inference of population structure using multilocus genotype data**. Genetics, v. 155, n. 2, p. 945-959, 2000.

PRYOR, L.D. **Biology of Eucaliptos**. Canberra: Edward Arnold, 1976. 78 p. (Studies in Biology, 61).

R Core Team (2021). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

ROBINSON, I. P. **Aloenzimas na genética de populações de plantas**. In: ALFENAS, A.C. Eletroforese de isoenzimas e proteínas afins; fundamentos e aplicações em plantas e microrganismos. Viçosa: UFV, 1998, cap. 7, p. 329-380.

RUTHERFORD, Susan et al. **Speciation in the presence of gene flow: population genomics of closely related and diverging Eucalyptus species**. Heredity, v. 121, n. 2, p. 126-141, 2018.

SANSALONI, C. P. **Desenvolvimento, caracterização e mapeamento de microssatélites de tetra epentanucleotídeos em Eucalyptus spp.** 2008. 114 f. Dissertação (Mestrado em Biologia Molecular) - Universidade de Brasília, Brasília, 2008.

SCHNEIDER, T. et al. **Genetic Dissimilarity in Conyza sumatrensis Revealed by Simple Sequence Repeat (SSR) Markers**. Planta Daninha, v. 38, 2020.

SEIN, Chaw Chaw; MITLÖHNER, Ralph. **Eucalyptus urophylla ST Blake: Ecology and silviculture in Vietnam**. CIFOR, 2011.

SIGRIST, M. S. **Divergência genética em Curcuma longa L. utilizando marcadores microssatélites e agromorfológicos**. 2009. 82 f. Dissertação (Mestrado em Área de Genética, Melhoramento Vegetal e Biotecnologia.) - Instituto Agrônomo de Campinas, Campinas, 2009.

SILVA, Paulo Henrique Müller da et al. **Selecting for stable and productive families of Eucalyptus urophylla across a country-wide range of climates in Brazil**. Canadian Journal of Forest Research, v. 49, n. 1, p. 87-95, 2019.

SILVA, Susana Maria Melo et al. **Genetic parameters for Hymenaea courbaril L. conservation in Southwestern AMAZON**. Ciência Florestal, v. 24, n. 1, p. 87-95, 2014.

SOARES, Sabrina Delgado et al. **Diversidade genética em população de melhoramento de mogno africano (*Khaya ivorensis* A. Chev.)**. Dissertação. (Mestrado em Genética e Biologia Molecular). Universidade Federal de Goiás. Goiânia. 2014.

SOBRAL, L. et al. **Acertando o alvo 2: consumo de madeira amazônica e certificação florestal no Estado de São Paulo**. Belém: Imazon, 2002. 72 p.

STEANE, Dorothy A. et al. **A comparative analysis of population structure of a forest tree, *Eucalyptus globulus* (Myrtaceae), using microsatellite markers and quantitative traits**. *Tree Genetics & Genomes*, v. 2, n. 1, p. 30-38, 2006.

WANG, M.L.; BARKLEY, N.A.; JENKINS, T.M. **Microsatellite markers in plants and insects**. Part I: Applications of Biotechnology. *Genes, Genomes and Genomics*, Bethesda, v. 3, n. 1, p. 54-67, 2009.

WEIR, B. S. **Genetic data analysis II: methods for discrete population genetic data**.

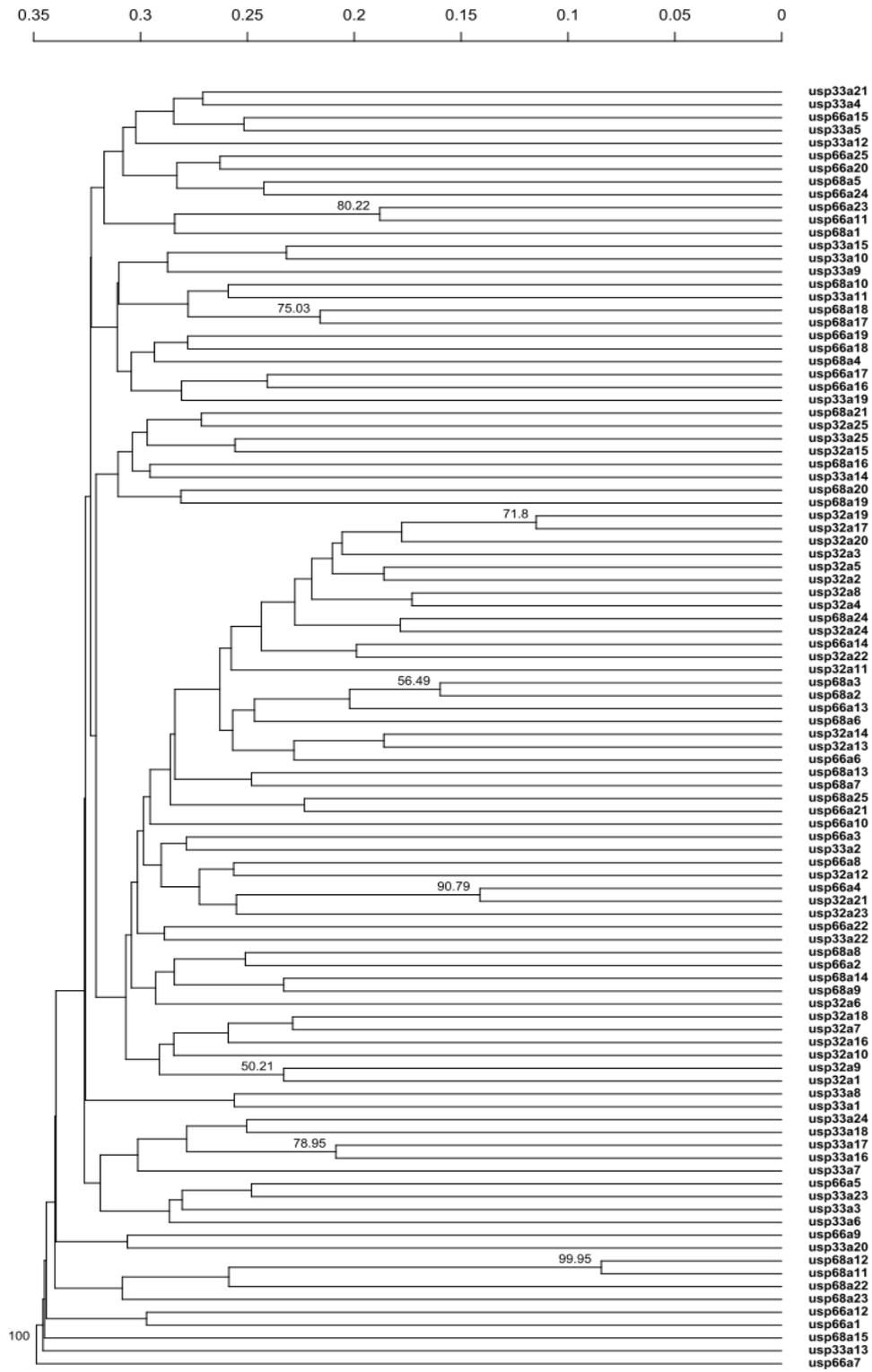
Sunderland (MA): Sinauer Associates, Sunderland. 1996. 376 p

WHITE, G.; W. POWELL, W. **Isolation and characterization of microsatellite loci in *Swietenia humilis* (Meliaceae): an endangered tropical hardwood species**. *Molecular Ecology*. v. 6, 1997.

ZOLET, Andreia Carina Turchetto et al. **Marcadores moleculares na era genômica: metodologias e aplicações**. 2017.

ZUCCHI, M. I. **Análise da estrutura genética de *Eugenia dysenterica* DC utilizando marcadores RAPD e SSR**. 2002. 130 f. Tese (Doutorado em Agronomia) – Escola Superior Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2002.

APÊNDICE



APÊNDICE B – Dendrograma baseado na distância de Nei dos indivíduos das populações naturais