

UNIVERSIDADE FEDERAL DO PAMPA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS

LUANA OLIVEIRA DE OLIVEIRA

CARACTERIZAÇÃO MOLECULAR DE ACESSOS DE NOGUEIRA-PECÃ
Carya illinoensis (WANGENH.) K. KOCH NO RIO GRANDE DO SUL USANDO
MARCADORES S-SAP

Dissertação de Mestrado

São Gabriel
2021

LUANA OLIVEIRA DE OLIVEIRA

**CARACTERIZAÇÃO MOLECULAR DE ACESSOS DE NOGUEIRA-PECÃ
Carya illinoensis (WANGENH.) K. KOCH NO RIO GRANDE DO SUL USANDO
MARCADORES S-SAP**

Dissertação apresentada ao Programa de Pós-graduação *Stricto Sensu* em Ciências Biológicas da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Ciências Biológicas.

Orientador: Dr. Valdir Marcos Stefenon
Co-orientador: Igor Poletto

**São Gabriel
2021**

Ficha catalográfica elaborada automaticamente com os dados fornecidos
pelo(a) autor(a) através do Módulo de Biblioteca do
Sistema GURI (Gestão Unificada de Recursos Institucionais).

A481o Oliveira, Luana Oliveira

CARACTERIZAÇÃO MOLECULAR DE ACESSOS DE NOGUEIRA-PECÃ
Carya illinoensis (WANGENH.) K. KOCH NO RIO GRANDE DO SUL
USANDO MARCADORES S-SAP. / Luana Oliveira de Oliveira. – 2021.

115 p. : il.

Dissertação (Mestrado) – Universidade Federal do Pampa,
Mestrado em Ciências Biológicas, 2021.
Orientação: Valdir Marcos Stefenon.

1. Juglandaceae. 2. Elementos transponíveis. 3. Marcadores Moleculares. I.
II. Título.

LUANA OLIVEIRA DE OLIVEIRA

CARACTERIZAÇÃO MOLECULAR DE ACESSOS DE NOGUEIRA-PECÃ *Carya illinoensis* (WANGENH.) K. KOCH NO RIO GRANDE DO SUL USANDO MARCADORES S-SAP

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Biológicas da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Ciências Biológicas.

Dissertação defendida e aprovada em: 10 de setembro de 2021.

Banca examinadora:

Prof. Dr. Valdir Marcos Stefenon

Orientador

UFSC

Prof. Dr^a. Adelar Montovani

UDESC

Prof. Dr^a. Lia Rejane Reiniger

UFSM



Assinado eletronicamente por **Adelar Mantovani, Usuário Externo**, em 21/09/2021, às 16:19, conforme horário oficial de Brasília, de acordo com as normativas legais aplicáveis.



Assinado eletronicamente por **Valdir Marcos Stefenon, Usuário Externo**, em 21/09/2021, às 17:50, conforme horário oficial de Brasília, de acordo com as normativas legais aplicáveis.



Assinado eletronicamente por **Lia Rejane Silveira Reiniger, Usuário Externo**, em 22/09/2021, às 10:55, conforme horário oficial de Brasília, de acordo com as normativas legais aplicáveis.



A autenticidade deste documento pode ser conferida no site https://sei.unipampa.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0619505** e o código CRC **B09260AB**.

“O conhecimento serve para encantar as pessoas não para humilhá-las”

Mario Sergio Cortella.

AGRADECIMENTOS

A minha família e a Deus por ter me concedido esta família, a minha mãe Maria de Lourdes, um ser exemplar movida pelo amor, é à base de tudo em minha vida, ao meu pai Jaci Oliveira (In memoriam), que ainda sinto junto a mim pelo amor que me deu em vida, aos meus irmãos Elderson, Eduardo e Ederson pelo apoio e companheirismo de sempre.

Aos professores da UNIPAMPA pelos ensinamentos e pelos desafios impostos que ajudaram na minha formação. Em especial agradeço ao meu orientador Valdir Stefenon pelos ensinamentos, pela paciência, pelos conselhos e pela confiança, é um exemplo de bom profissional e de ser humano. Agradeço aos meus colegas de laboratório, aos técnicos da UNIPAMPA e toda equipe de limpeza, da portaria, que ajudam tudo isso funcionar. Em especial aos meus colegas e amigos Daniele e Dalvan pela amizade, pelo apoio, pelo ombro amigo sempre que precisei, sou grata por estarmos sempre juntos mesmo separados fisicamente por uma pandemia, sem o apoio de vocês isso não seria possível.

Ao meu namorado Marco Boiani pelo apoio incondicional, por me incentivar sempre, por confiar em mim, por me amar tanto e aturar meus pequenos surtos, te amo. Enfim agradeço a todos que fizeram parte da minha trajetória acadêmica e contribuíram para realização desta etapa da minha vida.

Muito Obrigada!

RESUMO

O cultivo de noqueira-pecã vem aumentando em diversos países, baseado na demanda crescente pela noz. Esse movimento também se observa no Brasil, onde muitos produtores vêm investindo na cultura, especialmente na região sul do país. O arranjo produtivo dessa frutífera vem sendo consolidado, não só pelo aumento de áreas de plantio, mas também pela crescente inserção de agricultores e de empresas que fomentam a melhoria de equipamentos, a assistência técnica no surgimento de pequenas agroindústrias e outras redes de fabricação. No entanto, seu cultivo carece de informações tecnológicas que possam subsidiar os produtores, técnicos e envolvidos na cadeia produtiva na tomada de decisões frente aos desafios que ocorrem na produção e na comercialização. A presente dissertação teve por objetivo sistematizar a evolução do estado da arte/ciência relativa a pecanicultura e caracterizar a diversidade genética de acessos selecionados para constituir o banco de germoplasma em implantação na Universidade Federal do Pampa. Apesar do avanço de estudos com marcadores moleculares e seu uso para avaliação genética e de diversidade em plantas, especificamente, para a noqueira-pecã pouco ainda é pesquisado no Brasil. Este trabalho está dividido em dois capítulos. O primeiro trata-se de uma revisão do estado da arte sobre os estudos com marcadores moleculares voltadas para a esta espécie, ressaltando as lacunas e as perspectivas dessas tecnologias voltadas para a pecanicultura. O segundo capítulo apresenta uma análise que investigou diferenciações genéticas entre acessos de noqueira-pecã no RS, utilizando a técnica S-SAP. Com base nos dados obtidos nessa dissertação, pode-se inferir que configuram avanços para o conhecimento científico da *Carya illinoensis*, uma vez que foi possível caracterizar a diversidade genética de acessos da espécie utilizando marcadores moleculares usando a técnica S-SAP, obtendo uma diversidade genética de $H=0,46$ para os acessos analisados, tal resultado expressa uma diversidade alta se tratando de marcadores dominantes. As informações obtidas até o momento nos permitem auxiliar no aprofundamento de análises da diversidade genética, podendo indicar genes seletivos para melhoramento genético, sendo a diversidade crucial para isto. Nesse sentido, este resultado evidencia a importância do avanço de estudos genéticos voltados para a espécie, contribuindo para a seleção de genótipos superiores e possibilitando a criação de novos bancos de germoplasma no sul do Brasil.

Palavras-Chave: Juglandaceae, elementos transponíveis, marcadores moleculares.

ABSTRACT

The cultivation of pecan nuts has been increasing in several countries, based on the growing demand for the fruit. This movement can also be observed in Brazil, where many producers have been investing in the culture, especially in the southern region of the country. The productive arrangement of this fruit has been consolidated, not only by the increase in planting areas, but also by the growing insertion of farmers and companies that foster the improvement of equipment, technical assistance in the emergence of small agribusinesses and other manufacturing networks. However, its cultivation lacks technological information that can support producers, technicians, and those involved in the production chain in making decisions in face of the challenges that occur in production and commercialization. The present dissertation represents an effort to organize and present information about the evolution of the study of this crop, besides characterizing the genetic diversity of accessions selected to compose the germplasm bank being implemented at UNIPAMPA. Despite the advance of studies with molecular markers and their use for genetic evaluation and diversity in plants, specifically for the pecan walnut, little is still researched in Brazil. This work is divided into two chapters. The first is a review of the state of the art of studies using molecular markers for this species, highlighting the gaps and perspectives of these technologies for pecan cultivation. The second chapter presents an analysis that investigated genetic differentiation between pecan accessions in RS, using the S-SAP technique. Based on the data obtained in this dissertation, it can be inferred that they represent advances for the scientific knowledge of *Carya illinoensis*, since it was possible to characterize the genetic diversity of accessions of the species using molecular markers using the S-SAP technique, obtaining a genetic diversity of $H = 0.46$ for the accessions analyzed, this result expresses a high diversity when dealing with dominant markers. The information obtained so far allows us to assist in further analysis of genetic diversity, indicating that these genes may be related to selective genes, knowing that genetic diversity is crucial for genetic improvement. In this sense, this result highlights the importance of advancing genetic studies focused on the species, contributing to the selection of superior cultivars and enabling the creation of new germplasm banks in southern Brazil.

Keywords: Juglandaceae, transposable elements, molecular markers.

LISTA DE FIGURAS

Capítulo 1

Figura 1 – Planta de noqueira-pecã *Carya illinoensis*. a: visão geral de noqueira-pecã cultivar Imperial; b: visão da noz da cv. Imperial; c: visão da amêndoa da cv. Imperial 14

Figura 2 – Zoneamento edafoclimático da noqueira-pecã para o Rio Grande do Sul..... 18

Capítulo 2

Figura 2.1 – Análise dos componentes principais da noqueira-pecã com base nos dados genéticos moleculares..... 57

Figura 2.2 – Dendrograma de UPGMA baseada no índice de similaridade de Jaccard de 30 acessos de noqueira-pecã..... 58

LISTA DE TABELAS

Capítulo 2

Tabela 1 – Primers AFLP e LTRs utilizados para análise dos acessos de noqueira-pecã ... 54

Tabela 2 – Dados e combinações de primers para diversidade genética (H) total e índice de Shannon (*I*) 56

LISTA DE SIGLAS E ABREVIATURAS

°C - Graus Celsius

AFLP - *Amplified fragment length polymorphism* (Polimorfismo de comprimento de fragmento amplificado)

BSA - *Bovine Serum Albumin* (Albumina de soro bovino)

CTAB - *Cationic hexadecyl trimethyl ammonium bromide* (Catiônico de hexadecil trimetil amônio)

DNA - Ácido desoxirribonucleico

dNTP - Trifosfato desoxinucleotídeos

ISSR - *Inter Simple Sequence Repeats* (Repetições de sequência inter simples)

MgCl₂ - Cloreto de magnésio

ng - Nanograma

nm - Nanômetro

m - metros

cm - centímetros

ha - hectare

NGS - *Next generation Sequencing* (Sequenciamento de nova geração)

pb - Pares de base

PCR - *Polymerase Chain Reaction* (Reação em cadeia da polimerase)

RAPD - *Random amplified polymorphic DNA* (DNA polimórfico amplificado ao acaso)

RFLP - *Restriction fragment Length Polymorphism* (Polimorfismo de comprimento de fragmento de restrição)

LTR – Longas repetições terminais

SNP - *Single Nucleotide Polymorphism* (Polimorfismo de nucleotídeo simples)

SSR - *Simple sequence repeats* (Sequências simples repetidas)

S-SAP - *Sequence-Specific Amplified Polymorphism* (Polimorfismo amplificado de sequência específica)

ETs - Elementos Transponíveis

UV - Ultravioleta

μL - Microlitro

μM - Micromolar

SUMÁRIO

1	INTRODUÇÃO	14
1.1	A Nogueira-pecã	14
1.2	História do cultivo da noqueira-pecã no Brasil	15
1.3	Importância Econômica	16
1.4	Marcadores moleculares	18
1.5	Elementos transponíveis	20
1.6	Classificação dos elementos transponíveis	21
	Retrotransposons com LTRs	22
	Retrotransposons sem LTRs	22
	Elementos de Classe II	22
	Elementos de Classe III	23
1.7	Retrotransposons como marcadores moleculares em plantas	23
1.8	Mecanismos de ação dos transposons	24
1.9	Localização genômica de retrotransposons	25
1.10	Estudos diversos sobre retrotransposons em plantas	26
2	OBJETIVOS	28
2.1	Objetivo Geral	28
2.2	Objetivos Específicos	28
3	HIPÓTESE	28
	REFERÊNCIAS	29
	CAPÍTULO 1	37
	MOLECULAR MARKERS IN <i>Carya illinoensis</i> (Jugladaceae): FROM GENETIC CHARACTERIZATION TO MOLECULAR BREEDING	37
	CAPÍTULO 2	49
	CARACTERIZAÇÃO MOLECULAR DE ACESSOS DE NOGUEIRA-PECÃ	49

<i>Carya illinoensis</i> (WANGENH.) K. KOCH NO RIO GRANDE DO SUL USANDO MARCADORES S-SAP	49
1. INTRODUÇÃO	50
2. MATERIAL E MÉTODOS	52
2.1 Amostragem do material vegetal.....	52
2.2 Extração de DNA	53
2.3 Análise genética e molecular.....	54
2.4 Análise estatística.....	55
3. Resultados e discussões	56
REFERÊNCIAS	60
CONSIDERAÇÕES FINAIS.....	67

1 INTRODUÇÃO

1.1 A Nogueira-pecã

A noqueira-pecã pertence à família Juglandaceae da ordem Fagales que é composta por oito gêneros de plantas, destacando-se além de *Carya* as *Juglans*. Esses dois gêneros são os mais conhecidos no sul do Brasil, pela produção de nozes e madeira. O gênero *Juglans* ao qual pertence à espécie *Juglans regia* L., originária da Europa, tem suas amêndoas muito apreciadas no período natalino (Backes; Irgang, 2004). No Brasil, este gênero, embora pouco cultivado, é usado para fins ornamentais. O gênero *Carya*, tem como destaque *Carya illinoensis* (Wangenh.) K. Koch, conhecida como noqueira-pecã, é bem popular no sul do Brasil e bastante comercializada (MARCHIORI, 1997).

O Rio Grande do Sul possui o maior pomar de noqueira-pecã da América Latina, sendo *Carya illinoensis* a principal espécie plantada, e a sua produção é exclusiva para produção de nozes para consumo *in natura*, e ao final do seu ciclo produtivo a madeira das árvores pode ser utilizada como lenha (GATTO et al., 2008). A noqueira-pecã pode chegar a uma altura de 50 m, possui tronco retilíneo copa elíptica e com folhagem caducifólia. As folhas são do tipo alternas-compostas e imparipinadas, possuem de 9 a 17 folíolos sésseis ou quase sésseis, oblongo-lanceolados, de 5 a 20 cm de comprimento por 2,5 a 7,5 cm de largura, tendo ápice agudo ou acuminado, de coloração verde-amarelada-escura e glabros na face superior, são mais claros e raramente pubescentes no interior (MARCHIORI, 1997) (Figura 1a).

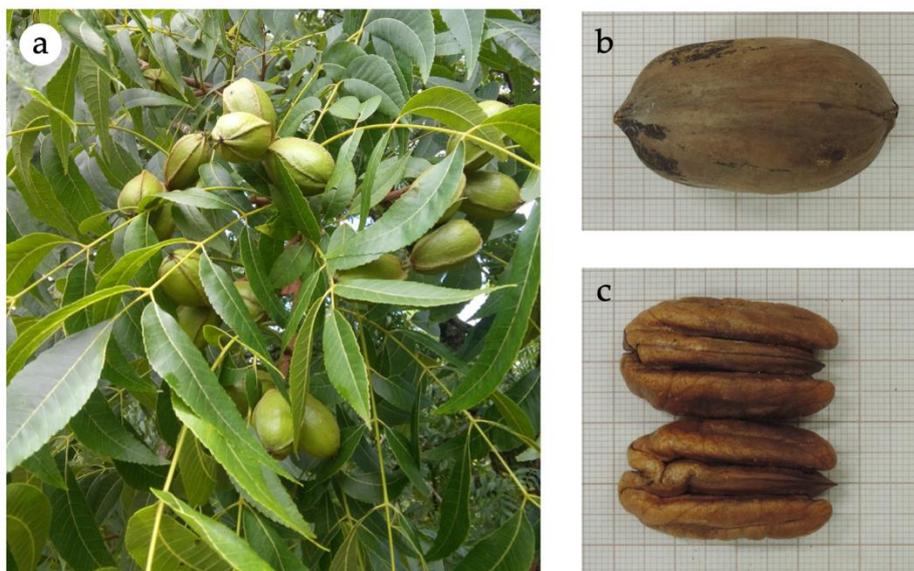


Figura 1: Planta de noqueira-pecã *Carya illinoensis*. **a:** visão geral de noqueira-pecã cultivar Imperial; **b:** visão da noz da cv. Imperial. **c:** visão da amêndoa da cv. Imperial. Fonte: Nagel et al. (2020).

Esta espécie é classificada como monoica, possui flores masculinas e femininas separadas, os frutos são do tipo drupa, normalmente agrupados em cachos de três a sete unidades, com epicarpo que se separa do fruto na maturação (figura 1b; figura 1c). A parte comestível do fruto varia de 40 a 60%, o período de maturação do fruto ocorre geralmente de março a maio, dependendo das condições edafoclimáticas, fitossanitárias, nutricionais e especialmente genéticas (DUARTE; ORTIZ, 2001).

Para que as plantas apresentem boa produtividade de nozes, é necessário que o pomar seja constituído de plantas pertencentes ao grupo das cultivares protândricas, ou seja, àquelas em que os órgãos masculinos chegam primeiro à maturidade “Desirable, Western, Cheyenne, Caddo, Cape Fear, Success, Barton, Cherokee”, e das cultivares protogínicas, em que os órgãos femininos chegam primeiro à maturidade “Choctaw, Wichita, Kiowa, Shawnee, Apache, Mohawk, Shoshoni, Chickasaw, Tejas, Elliott, Stuart, Comanche e Schley”. (RASEIRA, 1990).

Uma das principais formas de propagação utilizada para as cultivares é a enxertia, que pode ser iniciada a partir dos quatro até os oito anos, contando a partir do plantio (BACKES; IRGANG, 2004). Tanto as amêndoas, quanto as cascas das nozes apresentam compostos fenólicos com elevada capacidade antioxidante, o que pode constituir uma importante fonte para consumo na dieta humana. Em 100 gramas de amêndoas, são encontradas 718 calorias, 9,7 gramas de proteínas, 2,4g de fibras, 74g de gordura, e doses significativas de nutrientes e vitaminas, tais como ferro, potássio, magnésio, betacaroteno, tiamina, riboflavina, niacina e ácido ascórbico. Dentre seus benefícios, a amêndoa consumida *in natura* ajuda a manter os níveis de lipídios saudáveis, e por isso é tão apreciada na dieta humana (MCWILLIAMS, 2013).

1.2 História do cultivo da pecã no Brasil

A espécie *Carya illinoensis* teve sua chegada ao Brasil nos anos de 1900, e apesar de não ter informações precisas da sua introdução, o que se sabe é que pode ter sido trazida por pioneiros da imigração norte americana no país. Entre esses imigrantes, esta o Sr. Ezekiel Pyles, que se estabeleceu em Santa Bárbara D'Oeste, no estado de São Paulo, ele teria trazido consigo nozes que deram origem aos primeiros plantios em solo brasileiro. (ORTIZ e CAMARGO, 2005). No estado do Rio Grande do Sul, a cidade de Anta Gorda foi a primeira a estabelecer o cultivo de noqueira-pecã, que deu início aos seus primeiros pomares em 1943. A procedência das primeiras mudas para plantio era de Kentucky (EUA), que existem no

município até hoje e são consideradas as ancestrais da maior parte das noqueiras-pecã encontradas no município e na região (FRONZA; POLETTTO; HAMANN, 2013).

Apesar da maioria dos pomares implantados com incentivos fiscais nos anos 60, não estarem mais ativos por possuírem problemas fitossanitários, o cultivo ainda é praticado em toda a região sul, parte do sudeste e no Mato Grosso do Sul, tendo a maior concentração de pomares encontrada na depressão central do estado Rio Grande do Sul, Brasil (ORTIZ e CAMARGO, 2005). Algumas cultivares foram trazidas do México e dos EUA na forma de ramos e enxertadas no Brasil, já outras foram trazidas como sementes. Houve a tentativa de cultivo de muitas cultivares, porém algumas não se adaptaram e foram sendo eliminadas dos pomares. Ao longo de muitas décadas, as plantas para produção foram selecionadas pelos agricultores, com base principalmente em aspectos morfológicos. Como resultado desta seleção uma grande variedade de formas e tamanhos de noz-pecã é encontrada e cultivada em toda a área sul do Brasil. (POLETTTO et al., 2015).

1.3 Importância Econômica

A noqueira-pecã possui uma boa capacidade de adaptação, e com isso a cultura se espalhou pelos continentes e é cultivada em vários países como Uruguai, Argentina, Chile, Peru, China, Índia, Egito, África do Sul e Austrália. Os Estados Unidos lideram a produção global de noz- pecã, representando 53% do total, com 200.300 toneladas produzidas (com casca) na colheita de 2019. O México é o segundo maior produtor do mundo com 1500.00 toneladas produzidas em 2019, (INC- *International Nut and Dried Fruit Council*, 2019). A China, por sua vez, é a maior importadora de noz-pecã.

Atualmente, a maior parte da área plantada com noqueira-pecã no mundo situa-se na América do Norte, principalmente no seu local de origem, nos Estados Unidos e México. Seu cultivo comercial vem se ampliando e ocorrendo em vários países de diferentes continentes, como África, Austrália, inclusive na América do Sul, principalmente no Uruguai, Argentina, Chile, Peru e Brasil (BRILHARVA et al., 2018).

No contexto sul-americano, o cultivo brasileiro de noqueira-pecã se destaca como um dos principais países produtores de noz. A estimativa de área plantada no Brasil supera os 8 mil ha, ocorrendo em propriedades inferiores a 15 ha, geralmente de base familiar, sendo que mais de 70% da área está no RS, seguido por PR e SC (MARTINS et al., 2018).

O rendimento dos pomares implantados no país, pode chegar a mais de 2 mil kg ha⁻¹, quando em condições ambientais favoráveis e manejo adequado. No entanto, abaixo do

rendimento médio dos pomares americanos que chegam a produzir 33,3% a mais do que no Brasil (DUARTE; ORTIZ, 2001).

Nesse contexto, surgiu a demanda por um zoneamento edafoclimático para nogueira-pecã no Sul do Brasil, a exemplo de empreendimentos anteriores (FILIPPINI-ALBA et al., 2014; 2018a; 2018b). Trata-se de um instrumento suporte para o sistema produtivo e políticas públicas, no qual se identificam as zonas aptas e inaptas para cultivos, nesse caso para nogueira-pecã, com base em informações acerca do clima, do solo e de aspectos ambientais, de maneira a aprimorar a exploração dos recursos, proteger a natureza e viabilizar a sustentabilidade dos empreendimentos socioeconômicos.

Foi elaborado um estudo edafoclimático da nogueira-pecã para o RS e as classes definidas foram: P = Preferencial (ótimas condições de solo); R = Recomendada (há algum grau de restrição das propriedades dos solos); PR = Pouco recomendada (o grau de restrição das propriedades dos solos aumenta); PRf = PR com única restrição pela fertilidade do solo (aspecto passível de correção via tecnológica); NR = Não recomendada (condições inadequadas das propriedades do solo); (CANÇADO et al., 2012).

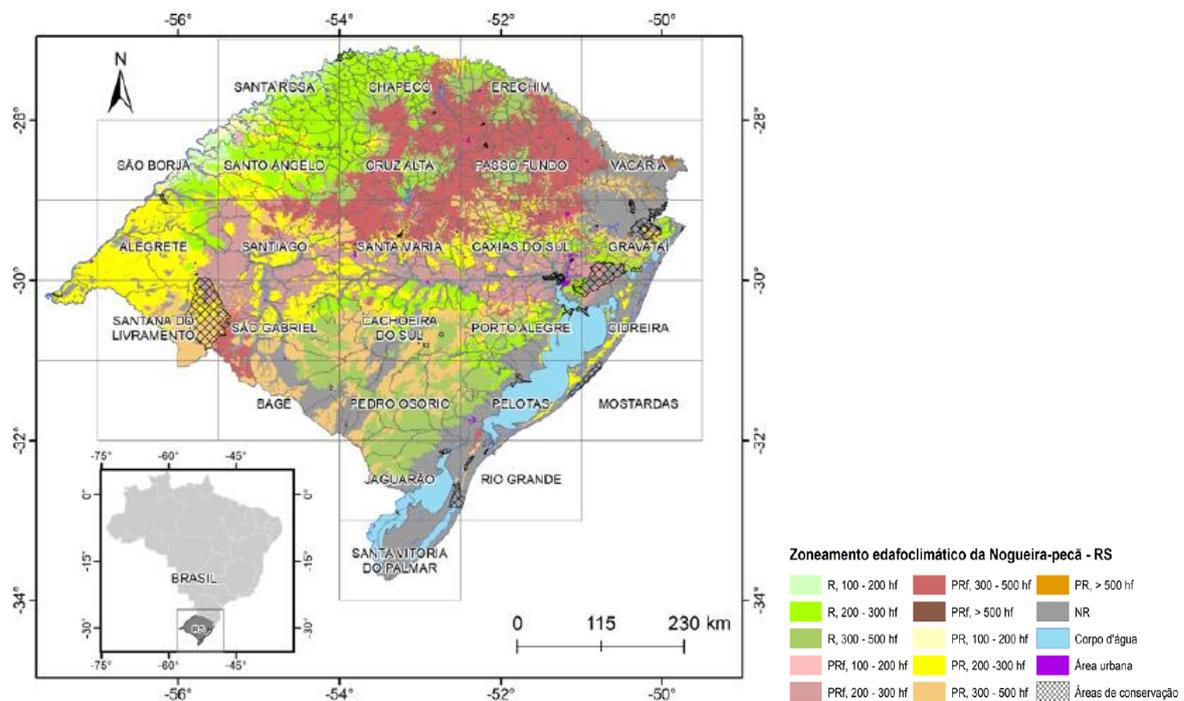


Figura 2: Zoneamento edafoclimático da nogueira-pecã para o Rio Grande do Sul e disposição das folhas cartográficas 1:250.000.

Fonte: Labplan – Embrapa Clima Temperado, Zoneamento Edafoclimático da Nogueira-pecã para a Região Sul do Brasil.

Na Figura 2, as cores verdes indicam áreas com aptidão edáfica recomendada (R), que migram para tons cor de rosa pouco recomendada com restrição na fertilidade do solo (PRf), e amarelos, pouco recomendada (PR) respectivamente. Os autores verificaram que no Rio Grande do Sul a disponibilidade de horas de frio (hf) é favorável para o cultivo da noqueira-pecã em todo o território. Nesse sentido, as restrições devem-se aos solos arenosos no litoral ou com camada inferior aos 50 cm, ou ainda aos solos mal drenados no setor central e Fronteira Oeste como mostra a figura 2.

As perspectivas de produção e as possibilidades de cultivo, bem como a comercialização das nozes, foram avaliadas por Lima (2011), em um sítio localizado no município de Itaqui-RS. A autora constatou que o produtor rural pode oportunizar o desenvolvimento de sua propriedade, por meio da renda obtida com a comercialização futura das amêndoas. Ressaltando que o cultivo de noz-pecã exige elevado custo inicial, mas apresenta resultados econômicos positivos a médio e longo prazo, o que torna a cultura uma alternativa viável para pequenos produtores rurais, em termos de diversificação de culturas e fonte de renda. Filippin (2011) corrobora enfatizando ser viável, economicamente a implantação da noqueira-pecã em áreas de reserva florestal e em sistema agroflorestais nas áreas de preservação permanente no Extremo Oeste Catarinense.

1.4 Marcadores moleculares

Marcadores moleculares podem ser definidos como todo e qualquer fenótipo molecular oriundo de um gene expresso ou de segmento específico do DNA. Os marcadores moleculares de DNA são técnicas da biotecnologia moderna, muito utilizados em diferentes áreas da pesquisa científica com humanos, animais, vegetais ou microrganismos. Seu potencial para desvendar os segredos contidos na molécula de DNA desperta o interesse em pesquisadores, não só da área de genética, mas também de inúmeros outros campos do conhecimento (CANÇADO et al., 2012).

A partir da década de 1960, métodos com base na detecção de polimorfismos genéticos por meio do padrão eletroforético de isoenzimas foram utilizados em análises de diversidade genética, uma vez que tais enzimas são codificadas por diferentes alelos em um ou mais locos gênicos. Entretanto, a expressão das enzimas pode ser muito influenciada pelas condições ambientais e pelo estágio de desenvolvimento das plantas. Consequentemente, apenas sistemas enzimáticos que não apresentavam variações em diferentes condições ambientais e fisiológicas podiam ser utilizados como marcadores adequados, o que reduzia o

número de marcadores disponíveis, restringindo o grau de polimorfismo possível de ser detectado por análise de isoenzimas. Posteriormente, novas técnicas com base na detecção de polimorfismos nas sequências de DNA foram desenvolvidas e amplamente utilizadas com a finalidade de detecção de polimorfismos genéticos. As tecnologias de análise molecular da variabilidade do DNA permitem determinar pontos de referência nos cromossomos, tecnicamente denominados “marcadores moleculares” (BORÉM; CAIXETA, 2008).

O desenvolvimento e a utilização de marcadores moleculares têm contribuído muito para o avanço da pesquisa genética, permitindo a detecção de polimorfismo de DNA. Pelo fato de os marcadores apresentarem um padrão Mendeliano de herança, estes são passíveis de utilização em diferentes áreas da genética e do melhoramento de plantas, tais como estudos populacionais, evolução, estruturação e análise sistêmica. Essas técnicas apresentam várias vantagens em relação àquelas com base nos caracteres morfológicos e nas isoenzimas, como a exclusão de interferências ambientais, possibilidade de realização em qualquer estágio de desenvolvimento da planta, utilização de qualquer tecido ou órgão vegetal, necessidade de pequenas amostras de tecido para realização da análise, além de ser um método não destrutivo (CANÇADO et al., 2012).

Os principais marcadores podem ser classificados em grupos conforme a metodologia utilizada para identificá-los: marcadores baseados na hibridação com sondas específicas e os marcadores baseados na amplificação do DNA via reação de polimerização em cadeia “PCR” (*Polymerase Chain Reaction*) e marcadores derivados de sequências expressas. Entre os identificados por hibridação estão os marcadores RFLP (*Restriction Fragment Length Polymorphism*) e os minissatélites ou locos VNTR (*Variable Number of Tandem Repeats*). Já aqueles revelados por amplificação incluem os marcadores RAPD (*Random Amplified Polymorphic DNA*), SCAR (*Sequence Characterized Amplified Regions*), Microsatélites ou SSR (*Single Sequence Repeat*), AFLP (*Amplified Fragment Length Polymorphism*) e STS (*Sequence Tagged Sites*).

Em noqueira-pecã os estudos com marcadores moleculares ainda são escassos, porém a demanda crescente da produção no sul do país requer o aprimoramento de novas técnicas. Os marcadores moleculares são especificamente valiosos para programas de reprodução de culturas perenes, podendo auxiliar na construção de mapas de ligação, e mapeamento de traços qualitativos e quantitativos. Temos como exemplo, estudos que associaram provisoriamente marcadores moleculares ao tipo de traço fenotípico da dicogamia em pecã (protândricos ou protogíneas), onde seis marcadores AFLP foram identificados e associados à dicogamia na noz-pecã. Contudo, a associação real destes marcadores ainda precisa ser

confirmada. Se for estabelecida esta associação, estes marcadores podem ser utilizados para desenvolver marcadores SCAR (região amplificada caracterizada pela sequência). Marcadores SCAR baseados em PCR seriam facilmente gerados e muito úteis para selecionar genótipos em programas de reprodução (CERNA-CORTÉS et al. 2003). Nesse sentido, as perspectivas futuras dos avanços de estudos com marcadores moleculares em noz-pecã são tratadas no capítulo 1 dessa dissertação.

1.5 Elementos transponíveis

A mobilidade de sequências do DNA foi primeiramente sugerida no final dos anos 1940, antes da descoberta da estrutura molecular do DNA, quando Bárbara McClintock revelou os “elementos controladores”. Neste período um dos dogmas da ciência era o princípio estático do genoma e a única mudança aceitável relacionava-se àquela proporcionada pelos processos evolutivos (KIDWEL; LISCH, 2001).

A instabilidade do genoma vem sendo estudada pela comunidade científica e sabe-se que este está sujeito a vários fatores que proporcionam sua variabilidade como as mutações, inserções e deleções. Outro fator que gera instabilidade são os elementos de transposição (TEs) descobertos por McClintock em 1931, em um estudo de mutações no genoma do milho (MCCLINTOCK, 1931). Nos anos de 1945 e 1946, durante um estudo utilizando sementes de milho, esta pesquisadora verificou um padrão incomum na expressão e segregação dos determinantes da coloração da aleurona dos grãos. Tal acontecimento só poderia ser explicado por “partículas móveis” presentes neste genoma. Desde então, estes elementos móveis foram encontrados em todas as espécies que já foram pesquisados (CRAIG, 2002; CAPY, et. al.,1994).

Os elementos transponíveis (ETs) são sequências de DNA que têm a capacidade de mudar de posição no genoma, independente de homologies entre as regiões. Esta transposição pode ocorrer de forma autônoma, quando os elementos produzem suas próprias enzimas que promovem sua mobilização; ou podem ser não autônomos, os quais dependem das enzimas produzidas pelos elementos autônomos. Um dos maiores impactos no genoma hospedeiro causado pela presença e movimentação desses elementos transponíveis está relacionado com o aumento da variabilidade genética (CAPY et al., 1998; KIDWELL, LISCH, 1997).

Devido à grande diversidade de elementos de transposição existentes nos diferentes genomas foi necessária a construção de uma classificação que facilitasse os estudos destas sequências. Os TEs podem ser agrupados hierarquicamente em: classes, subclasses, ordens,

superfamílias, famílias e subfamílias, de acordo com algumas características (WICKER et al., 2007). Os TEs podem ser classificados em duas classes, dependendo do mecanismo de transposição (GOODIE; KAZAZIAN, 2008); que pode ser por meio de DNA (classe II) ou através de um intermediário de RNA (classe I). Os TEs classe II são conhecidos como transposons, e se deslocam a partir de sequências de DNA e codificam uma proteína chamada transposase (enzima que catalisa sua saída do sítio original e a sua inserção em novos sítios). Ainda na classe II, temos os helitrons, eles foram encontrados e caracterizado em eucariotos, sendo denominados de helitrons (Plantas) e heleton (Vertebrados) (POULTER et al., 2003; KAPITONOV e JURKA, 2001). Eles apresentam um mecanismo de transposição chamado ‘rolling circle’ (PILAR et. al., 2001). Estão presentes na maioria das espécies de plantas e animais e são particularmente abundantes em plantas com flores (YANG e BENNETZEN, 2009). Os TEs da classe I, denominados como retrotransposons, movem-se via RNA e codificam transcriptases reversas que utilizam sua sequência como molde para fazer uma fita de RNA que serve como molde para posteriormente dar origem a uma fita de DNA que será inserida numa outra região do genoma (RANGNER, 1996).

Há ainda uma classe intermediária onde estão agrupados os elementos que não se adequam às características dos elementos das duas classes anteriores. A princípio, esta classe é dividida em duas superfamílias: MITEs (miniatura de elementos transponíveis com repetições invertidas), que possuem curtas repetições terminais e os elementos Foldback, com longas repetições terminais. Os elementos MITEs possuem uma estrutura interna comparável com a dos elementos da classe II, mas o seu alto número de cópias no genoma e sua inserção preferencial em regiões de genes os classificam como elementos da classe I (CAPY et al., 1998).

1.6 Classificação dos elementos transponíveis

Segundo Wicker et al. (2007), podemos melhor descrever cada classe dos TEs por categorias. Os elementos desta classe são conhecidos por Retrotransposons. Eles se movimentam através de um intermediário de RNA que é codificado para DNA por uma transcriptase reversa, produzida por eles mesmos, antes da sua nova inserção. Esta ainda se subdivide em retrotransposons com LTRs (Ty1-copia e Gypsy-Ty3), que são semelhantes ao retrovírus, e os retrotransposons sem LTRs (LINEs e SINEs) do inglês *long interspersed elements* e *short interspersed elements*, respectivamente.

Retrotransposons com LTRs

São elementos estruturalmente similares aos retrovírus. Possuem longas repetições nucleotídicas nas extremidades 5' e 3'. De uma maneira geral, estas repetições terminais flanqueiam uma região central que contém três módulos abertos de leitura conhecidos por ORFs, do inglês *Open Reading Frames*. A sequência das ORFs pode variar entre os elementos deste grupo. A primeira ORF refere-se ao gene *gag* que produz uma poliproteína que é processada em três proteínas maduras: a matriz, o capsídeo e o nucleocapsídeo. A outra ORF constitui-se do gene *pol* que codifica as enzimas necessárias à transposição do elemento: protease (Pr), transcriptase reversa (TR), RNaseH e integrase (Int). A última ORF está presente em algumas famílias desta classe, podendo ou não produzir uma proteína funcional: ela corresponde ao gene *env*, que codifica a proteína do envelope viral nos retrovírus.

Retrotransposons sem LTRs

Esta subclasse é dividida em duas superfamílias, LINEs (elementos nucleares dispersos longos) do inglês *long interspersed elements* e SINEs (elementos nucleares dispersos curtos) do inglês *shorts interspersed elements*, a diferença entre eles é a presença, nos LINEs, ou ausência, nos SINEs da enzima transcriptase reversa. A superfamília LINEs compreende elementos com tamanhos que variam entre 5 e 8 kb, e possuem uma sequência rica em adeninas no término 3', muito semelhantes às caudas poli A presente em RNAs mensageiros.

Elementos de Classe II

A classe II constitui os elementos chamados de Transposons. Essas sequências se movimentam através de um intermediário de DNA por um mecanismo de excisão do sítio original e reinserção em outro sítio no genoma. Todos os elementos possuem repetições terminais invertidas (TIR) do inglês *terminal invert repeats*. Os transposons são caracterizados pela mobilização dos elementos através de um processo de excisão do sítio doador e reinserção em outro sítio em qualquer posição no genoma hospedeiro. Ainda nesta classe, temos os elementos chamados de Helitrons, fazem o processo de replicação “*rolling circle*”. Foram descobertos pela análise computacional das sequências genômicas de

Arabidopsis thaliana, *Oryza sativa* e *Caenorhabditis elegans* (KAPITONOV e JURKA, 2001). Não apresentam as estruturas típicas que são características dos TEs tradicionais de classe de DNAA. Inserem-se preferencialmente entre os nucleotídeos de adenina e timidina (KAPITONOV E JURKA, 2007). São onipresentes em todos os eucariotos, como *Arabidopsis thaliana*, *Drosophila melanogaster*, *Oryza sativa* e *Zea mays* (LANGDON et al., 2009; LAL et al., 2009; KAPITONOV e JURKA, 2003; KAPITONOV e JURKA, 2001).

Elementos de Classe III

São chamados de MITEs (Miniatura de elementos transponíveis com repetições invertidas) do inglês “*Miniature inverted-repeat transposable element*”. É um grupo considerado heterogêneo de elementos não autônomos, são pequenos, variam entre dezenas e centenas de pares de bases. Flanqueiam uma região de repetições terminais invertidas e são frequentemente achados próximos dos genes.

1.7 Retrotransposons como marcadores moleculares em plantas

Marcadores moleculares baseados em retrotransposons são utilizados para a identificação e descrição de genótipos, porque geralmente mostram dispersão cromossômica, número de cópias variável e alta distribuição aleatória no genoma (KALENDAR et al., 1999).

Esses marcadores exploram métodos diversos, podem ser facilmente desenvolvidos, e oferecem acesso à parte dinâmica e polimórfica do genoma estudado. De modo geral, marcadores moleculares são mais eficientes que os marcadores morfológicos para medir a variabilidade genética existente entre os genótipos (SCHULMAN et al., 2012).

Os marcadores moleculares baseados em retrotransposons com regiões LTRs tem se mostrado muito eficazes, e vem sendo aplicados em uma gama de cereais, gramíneas, além de várias espécies de leguminosas, fungos, pássaros e insetos. As aplicações vão desde investigações de ativação dos retrotransposons, sua mobilidade para estudos de biodiversidade e evolução do genoma, e para o mapeamento de genes e estimativa de distância genética (KALENDAR et al., 2011).

O conhecimento de várias sequências desses elementos de transposição tem possibilitado o desenvolvimento de novas técnicas para estudar a dinâmica dos genomas. A dispersão, ubiquidade e prevalência desses elementos em genomas de plantas podem ser exploradas para “DNA *fingerprinting*” (identidade molecular), servindo como uma ferramenta

para futuros estudos de biologia de plantas. (KATSIOTIS et al., 1996, FLAVELL et al., 1992; PEARCE et al., 1996, 1997).

Os retrotransposons com LTRs exploram os polimorfismos do DNA genômicos resultantes de fenômenos de inserção entre dois retrotransposons muito próximos ou entre um retrotransposon e um SSR, no caso dos marcadores IRAP (*Inter Retrotransposon Amplified Polymorphism*) e REMAP (*Retrotransposon Microsatellite Amplified Polymorphism*) respectivamente (KALENDAR et al., 1999; PROVAN et al., 1999; SCHULMAN et al., 2004). O marcador S-SAP (*Sequence-Specific Amplified Polymorphism*) é um tipo de marcador molecular baseado na detecção de variação de fragmentos de DNA que flanqueiam sítios de inserção de retrotransposons. Os fragmentos são amplificados via PCR usando um primer desenhado das regiões LTRs, e outro baseado na presença de um sítio de endonucleases de restrição próximo às LTRs (WAUGH et al, 1997).

Os *primers* utilizados são geralmente desenhados para domínios conservados nas LTRs e/ou regiões SSR conservadas. Contudo, as regiões LTR conservadas podem diferir dentro da mesma família e entre famílias. Os métodos IRAP e REMAP originam, em cada amplificação, dezenas a centenas de produtos, dependendo da prevalência da família de retrotransposons e da organização do genoma da planta em estudo (SCHULMAN, 2007).

Portanto, muitas características dos retrotransposons, como o fato de serem ubíquos do genoma, abundantes e dispersos, tornam-nos eficientes para a sua utilização como sistemas de marcadores moleculares. Simultaneamente, a sua atividade leva à diversificação do genoma e fornece meios para a sua detecção. Os retrotransposons produzem grandes alterações genéticas no local de inserção, promovendo assim sequências conservadas que podem ser utilizadas na detecção da sua própria inserção. A aplicação da técnica IRAP (*Inter Retrotransposon Amplified Polymorphism*) combinada com a técnica REMAP (*Retrotransposon Microsatellite Amplified Polymorphism*) tem se apresentado como excelente marcador molecular para a identificação de variedades.

1.8 Mecanismos de ação dos transposons

Segundo Griffiths et. al. (2008) os transposons podem ser reconhecidos por sequências de DNA curtas nas extremidades, que são iguais entre si, mas estão em orientação invertida, denominadas de repetições terminais. O mecanismo de transposição é realizado pela enzima transposase, que na maioria dos casos é codificada pelo próprio transposon.

São capazes de codificar sua própria enzima transposase, sendo capazes de iniciar seu próprio movimento de transposição, portanto são denominados autônomos. Por exemplo, o elemento de transposição *Ac*, encontrado no milho, é autônomo, enquanto o elemento *Ds*, também encontrado no milho, não é considerado autônomo. Entretanto, a transposase codificada por *Ac* pode realizar a transposição de um elemento *Ds*, bem como de seu próprio elemento. Os elementos *Ds*, neste caso, são ditos não autônomos. Como os transposons não se movem com muita frequência, a sua presença em uma determinada região do genoma pode ser inferida pela presença de repetições invertidas nas extremidades do elemento. A transposição é iniciada quando um gene que codifica uma transposase é expresso e o seu RNA mensageiro é traduzido no citoplasma. A transposase, então, entra no núcleo e liga-se às repetições terminais invertidas do elemento. De maneira geral, existem dois tipos de mecanismos de transposição. No primeiro, o transposon move-se sem que haja a replicação (transposição conservativa). No outro tipo, denominado de transposição replicativa, o transposon movimentava-se através de sua replicação e posterior inserção em outro local do genoma (SANGLARD et al., 2012).

1.9 Localização genômica de retrotransposons

Os retrotransposons são presentes no reino vegetal, detectando-se a sua presença principalmente em algas unicelulares, briófitas, gimnospermas e angiospermas (FLAVELL et al., 1992). O número de cópias existente por genoma é variável, sendo frequente a existência de um elevado número de cópias em plantas com genomas de grande dimensão. Exemplos clássicos envolvem o caso do milho, em que 50-80% do genoma nuclear é atribuído a estes elementos (SANMIGUEL e BENNETZEN, 1998) e mais de 50% do genoma da cevada, onde a principal família de retrotransposons é a família *BARE-1* que representa 5% do seu genoma (VICIENT et al. 1999). Um dos principais responsáveis por estes valores é o próprio mecanismo de transposição destes elementos, capaz de originar um elevado número de cópias em pouco tempo. Este mecanismo parece estar na base do paradoxo do valor de DNA C, isto é, a falta de correlação existente entre o tamanho do genoma e a complexidade de um organismo. Segundo Sanmiguel e Bennetzen (1998) verificaram que a diferença abrupta nos valores de C dentro da família das gramíneas era devida a uma amplificação diferencial dos retrotransposons. Por exemplo, o arroz e a cevada têm aproximadamente o mesmo número de genes, mas a diferença entre o tamanho dos seus genomas é atribuída, em parte, à amplificação destes elementos (VICIENT et al., 1999).

Nesta perspectiva, a diferença que se observa em genomas de grande dimensão, como o da fava e do milho (cerca de 14.000 Mbp e 2500 Mbp respectivamente) pode resultar de uma amplificação bem-sucedida de retrotransposons (BENNETT e LEITCH, 2005; PEARCE et al., 1996; SANMIGUEL et al., 1996) enquanto genomas de pequenas dimensões, como o de *Arabidopsis* (130 Mbp), podem resultar em pouca transposição (WRIGH et al., 1996). Paralelamente, foi demonstrada a preferência que alguns elementos exibem pela amplificação em certas espécies: elementos *copia-like* proliferaram em espécies com genomas menores enquanto elementos *gypsy-like* proliferaram em genomas de maiores dimensões, como em *Gossypium* (HAWKINS et al., 2006).

Foram feitos inúmeros estudos sobre a localização dos retrotransposons LTRs no genoma de plantas com o objetivo de conhecer a sua diversidade e assim poder clarificar a sua função na evolução destes genomas. Dados resultantes de hibridação *in situ* fluorescente (FISH) do inglês *Fluorescent in situ hybridization*, em diversas espécies vegetais permitiu concluir que, de uma forma geral, parece haver diferença no padrão de distribuição de elementos *Ty1-copia* e *Ty3-gypsy*. Elementos *Ty1-copia* parecem não ter preferência por domínios genômicos específicos uma vez que foram detectados tanto em zonas eucromáticas, isto é, zonas ricas nas bases G e C, nas quais se localiza a maioria dos genes e que estão associadas à atividade transcricional (PEARCE et al. 1996a; PEARCE et al., 1996b; PRICE et al., 2002) como também em zonas heterocromáticas, ricas em A-T e associadas á inativação transcricional, nomeadamente zonas teloméricas achados em *Allium cepa* (Pearce et al., 1996b) e paracentroméricas achados em *Arabidopsis thaliana* (BRANDES et al., 1997).

Em outras gramíneas foi detectada a presença de retrotransposons LTR em “clusters”, num tipo de arranjo designado “nested”, ou seja, elementos inseridos uns dentro dos outros. Estes podem ser da mesma família, como é o caso de *BARE-1*, em cevada (SHIRASU et al., 2000) ou de famílias diferentes, em milho (SANMIGUEL et al., 1998). Ainda em milho, 49-78% do seu genoma é composto de retrotransposons (SANMIGUEL e BENNETZEN, 1998). Em trigo, cerca de 90% do genoma consiste em sequências repetidas e, destas, 68% correspondem a elementos transponíveis (LI et al., 2004). Elementos *Ty3-gypsy* são abundantes em cereais, em regiões eucromáticas, mas também heterocromáticas, especialmente em zonas centroméricas, como se verificou em milho, centeio, trigo, arroz e cevada, mas também em *Arabidopsis thaliana* (SANMIGUEL et al., 1996).

1.10 Estudos diversos sobre retrotransposons em plantas

Estudos com retrotransposons em plantas são bem variados. O uso de técnicas de biologia molecular é bem diverso, como a transformação genética, como no caso do desenvolvimento de um sistema de vetor de transformação baseado no retrotransposon *Tnt 1* em tabaco. O elemento chamado de *Mini-Tnt 1*, também oferece um sistema facilmente manipulável para estudar mecanismos de retrotransposição em plantas (HOU et al., 2010).

Relatos da utilização de retrotransposons em plantas como *Arabidopsis thaliana*, demonstra a existência da família de retrotransposon chamada *Ta 1-10*, a qual constitui 0,1% do genoma da planta. O genoma do fumo contém o retrotransposon *Tnt 1*, o qual foi isolado após a sua transposição e inativação do gene que codifica a enzima nitrato redutase. Plantas pertencentes à família Liliaceae, possuem o retrotransposon denominado *Del*, existindo mais de 13.000 cópias no genoma de *Lilium longiflorum*. Em arroz, existem três famílias principais de retrotransposons, a do *Tos*, a do *RIRE3*, e p- *SINE 1* (KUMAR e BENNETZEN, 1999).

Estudos em *Oryza australiensis* mostraram que a família *RIRE-1* é mais abundante na zona pericentromérica que na zona distal de ambos os braços cromossômicos uma característica partilhada com outras famílias LTR no genoma de *O. sativa* (JIANG et al., 2002). Outro estudo aponta os retrotransposons LTR como responsáveis pelas modificações no genoma no gênero *Eleocharis*. Os dados sugerem fortemente que Ty1-copia /Helos1 desempenhou um papel importante na evolução, tanto no tamanho do genoma quanto no cariótipo em *Eleocharis* (ZEDEK ET al., 2010).

Em *Vitis* sp., estudos com elementos transponíveis desempenharam um papel importante na domesticação da videira e sua evolução. A publicação do genoma completo da videira abre a possibilidade para uma análise profunda do conteúdo de transposons em seu genoma para entender sua evolução (BENJAK et al., 2008).

Na ativação da transcrição do trigo, os retrotransposons alteram a expressão de genes adjacentes, e leitura através dessa transcrição é a partir dos retrotransposons que estão associados na ativação ou silenciamento desses genes (KASHKUSH et al., 2003). Assim, elementos transponíveis, incluindo retrotransposons, contribuíram mais significativamente para a evolução de genes e genomas do que se pensava anteriormente. Durante a transposição, retrotransposons são transcritos reversamente por uma transcriptase reversa, e os cDNAs resultantes são integrados no genoma. Este modo de "copiar-e-colar" contribui para a expansão do tamanho do genoma. Estudos mostram que as sequências de retrotransposon não são eliminadas facilmente do genoma e são mantidas em uma forma inativa. É reconhecido que elementos transponíveis influenciam na evolução, aumentando tamanho de genomas e na

indução de diversos tipos de alterações dos genomas eucarióticos superiores (KIDWELL e LISCH, 1997).

Em *Arabidopsis thaliana*, como se considera uma espécie modelo das dicotiledôneas e de estudos moleculares, tem contribuído muito para a compreensão dos sistemas de prevenção de invasão no genoma por elementos transponíveis, os avanços recentes sugerem que *A. thaliana* pode ser mais eficiente do que a *A. lyrata*, pelo controle da expressão e proliferação dos elementos transponíveis. A análise comparativa da transcrição de elementos transponíveis em *A. thaliana* e *A. lyrata*, diferem em 40% no tamanho do genoma, pode ajudar a compreender como os mecanismos de silenciamento contribuíram para a evolução, um importante fator de variação de tamanho nos genomas de plantas e animais (MEAUX; PECINKA, 2012). Retrotransposons LTRs também tiveram sua história evolucionária estudada recentemente, através dos domínios proteicos, nas plantas do gênero *Selaginella*, em análises comparativas e filogenéticas (NOVIKOV et al., 2012).

2 OBJETIVOS

2.1 Objetivo Geral

Caracterizar a diversidade genética de acessos selecionados de noqueira-pecã (*Carya illinoenses*) utilizando marcadores moleculares S-SAP.

2.2 Objetivos Específicos

- Utilizar retrotransposons como marcadores moleculares na caracterização genética de acessos de noqueira-pecã (*Carya illinoenses*).
- Estimar o índice de diversidade genética dos acessos de noqueira-pecã que compõem o banco de germoplasma da UNIPAMPA.
- Estudar a similaridade/divergência genética entre os acessos avaliados

3. HIPÓTESE

O uso da técnica S-SAP como marcador auxiliará na identificação da diversidade genética dos acessos de noqueira-pecã selecionados para compor um banco de germoplasma.

REFERÊNCIAS

BACKES, P.; IRGANG, B. Árvores do Sul. Guia de identificação & interesse Ecológico. As principais espécies nativas Sul-Brasileiras. Santa Cruz do Sul: Instituto Souza Cruz, 2004.

BENNETT, M.D.; LEITCH, I.J. Genome size evolution in plants. In **The Evolution of the Genome**, editado por Gregory. San Diego: Elsevier. Pages 89–162, 2005.

BENJAK, A.; FORNECK, A.; CASACUBERTA, J.M. Genome-Wide Analysis of the “Cut and-Paste” Transposons of Grapevine. **PLoS ONE** v.3, p.9, 2008.

BILHARVA, M. G.; MARTINS, C. R.; HAMANN, J. J.; FRONZA, D.; De Marco, R.; MALGARIM, M. B. Pecan: from Research to the Brazilian Reality. **American journal of experimental agriculture**, v. 23, p. 1-16, 2018.

BORÉM, A.; CAIXETA, E.T. **Marcadores moleculares**. 2.ed. Viçosa, MG: UFV, 2008. v.1, 532p.

CANÇADO, G. M. A. et al. Marcadores moleculares de DNA e suas aplicações na caracterização, identificação e melhoramento genético da oliveira. **Capítulo**, v. 8, p. 225-249, 2012.

CAPY, P.; BAZIN, C.; HIGUET, D.; LANGIN, T. Dynamics and evolution of transposable elements. Austin, Texas: **Landes Bioscience**, p. 197, 1998.

CAPY, P.; ANXOLABÉHÈRE, D.; LANGIN, T. The strange phylogenies of transposable elements: are horizontal transfers the only explanation? **Trends Genetics**, v. 10, p. 7– 12. 1994.

DUARTE, V.; ORTIZ, E. R. N. Podridão de *Phytophthora* da amêndoa e da casca da noqueira pecan. In: LUZ, E. D. M. N. et al. (Ed.). **Doenças causadas por *Phytophthora* no Brasil**. Campinas: Rural, 2001. p. 493-508.

FILIPPIN, I. L. **Viabilidade econômica do cultivo de noqueira pecã em áreas de reserva legal e de preservação permanente**. 2011. 74 f. Dissertação (Mestrado em Ciência e Tecnologia de Sementes) - Universidade Federal de Pelotas, Pelotas, 2011.

FILIPPINI ALBA, J. M.; FLORES, C. A.; WREGE, M. S. **Zoneamento edafoclimático da olivicultura para o Rio Grande do Sul**. Brasília, DF: Embrapa, 2014. 68 p.

FILIPPINI ALBA, J. M.; WREGE, M. S.; FLORES, C. A.; ALMEIDA, I. R.; HERTER, F. G. **Zoneamento edafoclimático da Pereira para Santa Catarina**. Pelotas: Embrapa, 2018a. 22 p. (Embrapa Clima Temperado. Documentos, 470).

FILIPPINI ALBA, J. M.; WREGE, M. S.; FLORES, C. A.; ALMEIDA, I. R.; HERTER, F. G. **Zoneamento edafoclimático da Pereira para o Rio Grande do Sul**. Pelotas: Embrapa, 2018b. 16 p. (Embrapa Clima Temperado. Documentos, 466).

FILIPPINI ALBA, J. M.; WREGE, M. S.; MARTINS, C. R.; ALMEIDA, I. R. **Critérios e indicadores edafoclimáticos para o cultivo da noqueira-pecã no Sul do Brasil**. Pelotas: Embrapa, 2018c. 17 p. (Embrapa Clima Temperado. Documentos, 479).

FLAVELL, A.J., E. DUNBAR, R. ANDERSON, S.R. PEARCE, R. HARTLEY, AND A. KUMAR. Ty1-copia group retrotransposons are ubiquitous and heterogeneous in higher plants. **Nucleic Acids Research**, v.20, p.3639-3644,1992.

FRONZA, D.; POLETTO, T.; HAMANN, J. J. **O cultivo de noqueira-pecã**. Santa Maria: UFSM, 2013. 301 p.

GRIFFITHS, A. J. F.; MILLER, J. H.; SUZUKI, D. T.; LEWONTIN, R. C.; GELBART, W. M. Introdução à Genética. 9. ed. Rio de Janeiro: **Guanabara Koogan**. 2008.

GATTO, D. A. et al. Características tecnológicas das madeiras de *Luehea divaricata*, *Carya illinoensis* e *Platanus x acerifolia* quando submetidas ao vergamento. **Ciência Florestal**, Santa Maria, v. 18, n. 1, p. 121-131, 2008.

HAWKINS, J.A.; KIM, H.; NASON, J.D.; WING, R.A. Differential lineage-specific amplification of transposable elements is responsible for genome size variation in *Gossypium*. **Genome Research**, v. 16, p. 1252-1261, 2006.

HOU, Y.; RAJAGOPA, J.; IRWIN, P.; VOYTAS, D. Retrotransposon vectors for gene delivery in plants, **Mobile DNA**; v.1, p. 19, 2010.

JIANG, N.; BAO, Z.; TEMNYKH, S.; CHENG, Z.; JIANG, J.; WING, R.A.; MCCOUCH, S.R.; WESSLER, S.R. Dasheng: A Recently Amplified Nonautonomous Long Terminal Repeat Element That Is a Major Component of Pericentromeric Regions in Rice. **Genetics**, v.161, p.1293–1305, 2002.

JIANG, F.; YANG, M.; GUO, W.; WANG, X.; KANG, L. Large-Scale Transcriptome Analysis of Retroelements in the Migratory Locust, *Locusta migratoria*. **PLoS ONE** v.7(7): p. 1-15, 2012.

KALENDAR, R.; GROB, T.; REGINA, M.; SUONIEMI, A.; SCHULMAN, A. H. IRAP and REMAP: Two new retrotransposon-based DNA fingerprinting techniques. **Theoretical and Applied Genetics**, v. 98, p. 704-711,1999.

KAPITONOV, V.V.; JURKA, J. Rolling-circle transposons in eukaryotes. **Proceedings of the National Academy of Sciences USA**, v. 98(15), p. 8714-8719, 2001.

KAPITONOV, V.V.; JURKA, J. Helitrons on a roll: eukaryotic rolling-circle transposons. **Trends Genetics**, v. 23(10), p.521-529, 2007.

KAPITONOV, V.V.; JURKA, J. Molecular paleontology of transposable elements in the *Drosophila melanogaster* genome. **Proceedings of the National Academy of Sciences USA**, v.100 (11), p. 6569-6574. 2003.

KASHKUSH K.; FELDMAN, M.; LEVY, A.A. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. **Nature Genetics**., v.33, p. 102-106, 2003.

KATSIOTIS, A.; SCHMIDT, T.; HESLOP-HARRISON, J.S. Chromosomal and genomic organization of Ty1-copia-like retrotransposon sequences in the genus *Avena*. **Genome**, v.39, p. 410–417, 1996.

KIDWELL, M.G; LISCH, D.R. Transposable elements as source of variation in animals and plants. **Proceedings of the National Academy of Sciences USA**, v. 94, n. 15, p. 7704-7711, 1997.

KIDWELL, M.G; LISCH, D.R. Transposable elements, parasitic DNA, and genome evolution. **Evolution**, v. 55, p. 1-24, 2001.

KUMAR, A.; BENNETZEN; J.L. Plant retrotransposons. **Annual Review of Genetics**, v.33 p. 479-532, 1999.

LAL, SK.; GIROUX, MJ.; BRENDDEL, V.; VALLEJOS, CE.; HANNAH, LC. The maize genome contains a helitron insertion. **Plant Cell**, v.15(2), p.381-391, 2003.

LANGDON, T.; THOMAS, A.; HUANG, L.; FARRA, K.; KING, J.; ARMSTEAD, I. Fragments of the key flowering gene *GIGANTEA* are associated with helitron-type sequences in the Pooideae grass *Lolium perenne*. **BMC Plant Biology**, v. 9, p. 70, 2009.

LIMA, S. S. **O cultivo de noz pecã no município de Itaqui-RS**: O estudo de caso do sítio Parintis. 2011. 49 f. Trabalho de Conclusão (Curso de Graduação Tecnológico em Planejamento para o Desenvolvimento Rural) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2011.

LI, W., ZHANG, P., FELLERS, J.P., FRIEBE, B., GILL, B.S. Sequence composition, organization, and evolution of the core Triticeae genome. **Plant Journal**, v.40, p. 500-511, 2004.

MEAUX, J. ; PECINKA, A. The *Arabidopsis* genus *An* emerging model to elucidate the molecular basis of interspecific differences in transposable element activity, **Mobile Genetic Elements**, v.2:3, p. 142-144; 2012.

NOVIKOV, A.; SMYSHLYAEV, G.; NOVIKOVA, O. Evolutionary History of LTR Retrotransposon Chromodomains in Plants, **International Journal of Plant Genomics**, v.2012, p. 17, 2012.

MARCHIORI, J. N. C. **Dendrologia das angiospermas: das magnoliáceas às flacurtiáceas**. Santa Maria: Editora da UFSM, 1997. 271 p.

MARTINS, C. R., CONTE, A.; FRONZA, D.; FILIPPINI ALBA, J. M.; HAMANN, J. J.; BILHARVA, M. G.; MALGARIM, M. B.; FARIAS, R. M.; DE MARCO, R.; REIS, T. S. **Situação e perspectiva da noqueira-pecã no Brasil**. Pelotas: Embrapa Clima Temperado, 2018. 31p. (Embrapa Clima Temperado. Documentos, 462)

MCWILLIAMS, J. **The pecan: A history of America's native nut**. Austin: Hardcover, 2013. 178 p.

ORTIZ, E. R. N.; CAMARGO, L. E. A. Doenças da noqueira pecan (*Carya illinoensis*). In: KIMATI, H. et al. (Org.). **Manual de fitopatologia**. 4. ed. São Paulo: Agronômica Ceres, 2005. v. 2. p. 501-506.

PEARCE, S.R.; HARRISON, G. D; LI, J.; KUMAR, A.; FLAVELL, A.J. The Ty1- copia group retrotransposons in Vicia species: copy number, sequence heterogeneity and chromosomal localisation. **Molecular Genomics Genetics**, v. 250, p.305-315. 1996a.

PEARCE, S.R.; PICH, U.; HARRISON, G.; FLAVELL, A.J.; HESLOP-HARRISON, J.S.; SCHUBERT, I.; KUMAR, A. The Ty1-copia group retrotransposons of *Allium cepa* are distributed throughout the chromosomes but are enriched in the terminal heterochromatin. **Chromosome Research**, v. 4, p. 357-364, 1996b.

PEAKALL, R.; SMOUSE, P.E. GENALEX 6: genetic analysis in excel - population genetic software for teaching and research. **Molecular Ecology Notes**, v.6, n.1, p.288-295, Mar. 2006.

PILAR, G.B; BERNALLES, M.; MENDIOLA, M.V; La CRUZ, F. Single-stranded DNA intermediates in IS91 rolling- circle transposit. **Molecular Microbiology**, v.39, p.494- 501, 2001.

POLETTTO, T.; MUNIZ, M. F. B.; BAGGITTO, C.; CECONI, D. E.; POLETTTO, I. Fungos associados às flores e frutos da noqueira-pecã (*Carya illinoensis*). **Revista de Ciências Ambientais**, Canoas, v. 8, n. 1, p. 5-13, 2014.

POLETTTO, T, et. al., Morphological, chemical and genetic analysis of southern Brazilian pecan (*Carya illinoensis*) accessions. **Journal Scientia Horticulturae**, 2020.

POULTER, R.M.; GOODWIN, T.J.D.; BUTTER, M.I. Vertebrate helitrons and other novel helitrons. **Gene**, v.313, p. 201-212, 2003.

PRICE, Z., F. DUMORTIER, D.W. MACDONALD, AND S. MAYES. Characterisation of copia-like retrotransposons in oil palm (*Elaeis guineensis* Jacq.). **Theoretical and Applied Genetics**, 2002.

PROVAN, J.; THOMAS, W. T. B.; FORESTER, B. P.; POWELL, W. Copia -SSR: A simple marker technique which can be used on total genomic DNA. **Genome**, v. 42, p.363-366; 1999.

RANGNER, L.P. Transposons. In ZAHA A. (ed.) **Biologia Molecular Básica**. Porto Alegre, Editora Mercado Aberto, 159p. 1996.

RASEIRA, A. **Acultura da noqueira-pecã** (*Carya illinoensis*). Pelotas: Embrapa Clima Temperado, 1990. 3 p. (Embrapa Clima Temperado. Comunicado Técnico , 63).

SANGLARD, D.A.; FERREIRA, J.V.; BRAZ. C. C. Transposons ou Elementos Genéticos Móveis: Um dos Mais Peculiares Mecanismos de Recombinação Gênica, 2012.

SANMIGUEL, P.; A. TIKHONOV, Y.K. JIN, N. MOTCHOULSKAIA, D. ZAKHAROV, A. MELAKE, Nested retrotransposons in the intergenic regions of the maize genome. **Science**, v. 274, p.765-768,1996.

SANMIGUEL, P.; BENNETZEN, J.L. Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. **Annals of Botany**, v. 82, p.37-44, 1998.

SCHULMAN, A.H.; FLAVELL, A.J; PAUX, E.; ELLIS, T.H. The application of LTR retrotransposons as molecular markers in plants. **Methods Molecular Biology**, v. 859, p.115-53; 2012.

SCHULMAN, A. H. Molecular markers to assess genetic diversity. **Euphytica**. 158: 313-321; 2007.

SCHULMAN, A.H.; FLAVELL, A. J.; ELLIS, T. H. N. The application of LTR retrotransposons as molecular markers in plants. **Methods Molecular Biology**, v. 260, p.145-173; 2004.

SHIRASU, K.; SCHULMAN, A.H.; LAHAYE, T.; SHULZE-LEFERT, P. A Contiguous 66-kb Barley DNA Sequence Provides Evidence for Reversible Genome Expansion. **Genome Research**, v. 10, p. 908-915, 2000.

SUJII, P.S., FERNANDES, E.T.M.B., AZEVEDO, V.C.R., CIAMPI, A.Y., MARTINS, K., WADT, L.H.O., 2013. Morphological and molecular characteristics do not confirm popular classification of the Brazil nut tree in Acre. Brazil. *Genet. Mol. Res.* 12 (3), 4018–4027.

VICIENT, C.M.; KALENDAR, R.; ANAMTHAWAT-JONSSON, K.; SCHULMAN, A.H. Structure, functionality, and evolution of the BARE -1 retrotransposon of barley. **Genetica**, v.107, p. 53-63, 1999.

WICKER, T.; SABOT, F.; HUA-VAN, A.; BENNETZEN, J.L.; CAPY, P.; CHALHOUB, B.; FLAVELL, A.; LEROY, P.; MORGANTE, M.; PANAUD, O.; PAUX, E.; SANMIGUEL, P.; SCHULMAN, A.H. A unified classification system for eukaryotic transposable elements. **Nature Reviews Genetic**, v.8, p.973-982. 2007.

WAUGH, Robbie et al. Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). **Molecular and General Genetics MGG**, v. 253, n. 6, p. 687-694, 1997.

WRIGHT, D.A., N. KE, J. SMALLE, B.M. HAUGE, H.M. GOODMANT, AND D.F. VOYTAS. Multiple non- LTR retrotransposons in the genome of *Arabidopsis thaliana*. **Genetics**, v. 142: 569-578, 1996.

WELLS, L. (Ed.). **Southeastern pecan growers' handbook**. Georgia: University of Georgia, 2013. 236 p.

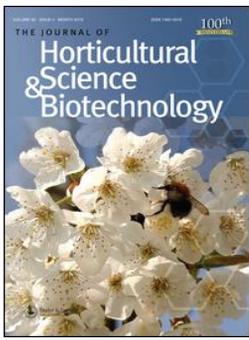
WELLS, L. **Pecan: america's native nut tree**. Tuscaloosa: The University of Alabama Press, 2017. 264 p.

YANG, L.; BENNETZEN, J. L. Structure-Based Discovery and Description of Plant and Animal "Helitrons". **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, p. 12832-12837, 2009.

ZEDEK, F.; SMARDA, J.; SMARDA, P.; BURES, P. Correlated evolution of LTR retrotransposons and genome size in the genus *Eleocharis*, **BMC Plant Biology**, v.10, p.265, 2010.

CAPÍTULO 1

MOLECULAR MARKERS IN *Carya illinoensis* (Jugladsaceae): FROM GENETIC CHARACTERIZATION TO MOLECULAR BREEDING



Molecular markers in *Carya illinoensis* (*Juglandaceae*): from genetic characterization to molecular breeding

Luana Oliveira de Oliveira, Dalvan Carlos Beise, Daniele Damian Dos Santos, Jordana Caroline Nagel, Tales Poletto, Igor Poletto & Valdir Marcos Stefenon

To cite this article: Luana Oliveira de Oliveira, Dalvan Carlos Beise, Daniele Damian Dos Santos, Jordana Caroline Nagel, Tales Poletto, Igor Poletto & Valdir Marcos Stefenon (2021): Molecular markers in *Carya illinoensis* (*Juglandaceae*): from genetic characterization to molecular breeding, The Journal of Horticultural Science and Biotechnology, DOI: [10.1080/14620316.2021.1892534](https://doi.org/10.1080/14620316.2021.1892534)

To link to this article: <https://doi.org/10.1080/14620316.2021.1892534>



Published online: 03 Mar 2021.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Molecular markers in *Carya illinoensis* (Juglandaceae): from genetic characterization to molecular breeding

Luana Oliveira de Oliveira ^a, Dalvan Carlos Beise ^a, Daniele Damian Dos Santos ^a, Jordana Caroline Nagel ^{a,b}, Tales Poletto ^c, Igor Poletto ^d and Valdir Marcos Stefenon ^{a,e}

^aPrograma de Pós-Graduação em Ciências Biológicas, Universidade Federal do Pampa, São Gabriel, Brazil; ^bDepartment of Biological Sciences, Universidade Regional Integrada do Alto Uruguai e das Missões, Santo Ângelo, Brazil; ^cPrograma de Pós-Graduação em Engenharia Florestal, Universidade Federal de Santa Maria, Santa Maria, Brazil; ^dLaboratório de Defesa Florestal, Universidade Federal do Pampa, São Gabriel, Brazil; ^eLaboratório de Fisiologia do Desenvolvimento e Genética Vegetal, Departamento de Fitotecnia, Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil

ABSTRACT

Carya illinoensis (pecan) is a fruit crop native to Mexico and the USA and cultivated across the world. Despite the advances of the molecular markers' technologies and their use in genetic studies of diversity and breeding, studies with this species only reticently accompanied this progress. Here we review the studies using molecular markers, the gaps, and the perspectives concerning these technologies in pecaniculture. The selection of plants more adapted to the particular conditions where the species is cultivated and the obligatory outcrossing mode of reproduction lead the species to generate several new genetic combinations. Molecular markers were employed to assess this diversity, using PCR-based markers and sequencing-based technologies. Genetic diversity was characterised in natural populations and collections of pecan germplasm. Markers putatively linked to traits of interest were reported and the genome sequencing technology provided insights into the evolution of the species. The outcomes of these studies still have to be integrated and further exploited to reach and impact the orchards and breeding programmes. However, the scientific advances in molecular markers research and the growth of the pecan's supply chain may drive to a scenario in which these biotechnologies will become tools regularly assisting breeders and farmers.

ARTICLE HISTORY

Accepted 15 February 2021

KEYWORDS

SSR markers; AFLPs; genome sequencing; gene mapping

Introduction

The pecan [*Carya illinoensis* (Wangenh.) K. Koch, *Juglandaceae*] is a crop fruit species native to the United States and Mexico (Grauke, Wood, & Harris, 2016). The natural occurrence range of the species is quite variable concerning climatic conditions. This fact confers high genetic variation among genotypes from different climatic regions (Sparks, 2005). This distribution over a wide climatic range gives greater adaptability to the species, enabling it to grow outside its natural occurrence area, especially in regions of temperate and subtropical climate. Given this adaptability, the pecan culture has spread across continents and is cultivated in several countries as Uruguay, Argentina, Brazil, Chile, Peru, China, India, Egypt, South Africa, and Australia.

The United States leads the global pecan production, representing about 53% of the total, with 200,300 tons produced (in-shell basis) in the 2019 harvest (INC – International Nut and Dried Fruit Council, 2019). Mexico is the second-largest producer in the world with 150,000 tons produced in 2019 (INC – International Nut and Dried Fruit Council, 2019). China, in turn, is the world's largest importer of

pecan nuts, demanding more than a third of US pecan production (Zhang, Peng, & Li, 2015).

Currently, the National Collection of Genetic Resources for Pecans and Hickories, maintained by the Agricultural Research Service of the U.S. Department of Agriculture (ARS-USDA) is the main pecans germplasm collection in the world. This collection was established in the 1930s and holds around 160 commercial cultivars of pecan (<https://cgru.usda.gov/carya/pecans/pecalph.htm>), as well as several native genotypes selected aiming at furnishing plants for use in the Pecan Breeding Program. Further pecan collections in the U.S.A. are found in the Southeastern Fruit and Tree Nut Lab, the Pecan Experimental Field, the University of Nebraska, the Louisiana State University Agricultural Center, the Alabama Pecan Collection, and the Fand Louisiana State University (<https://aggie-horticulture.tamu.edu/carya/cgc/cgc97.htm>). Germplasm banks are also found in Mexico (ca. 15 commercial cultivars from the U.S.A. and 55 selected open-pollinated plants; Eutimio Cuellar, personal communication), Argentina (tens of commercial cultivars from the U.S.A.), Brazil (ca. 45 commercial cultivars from the

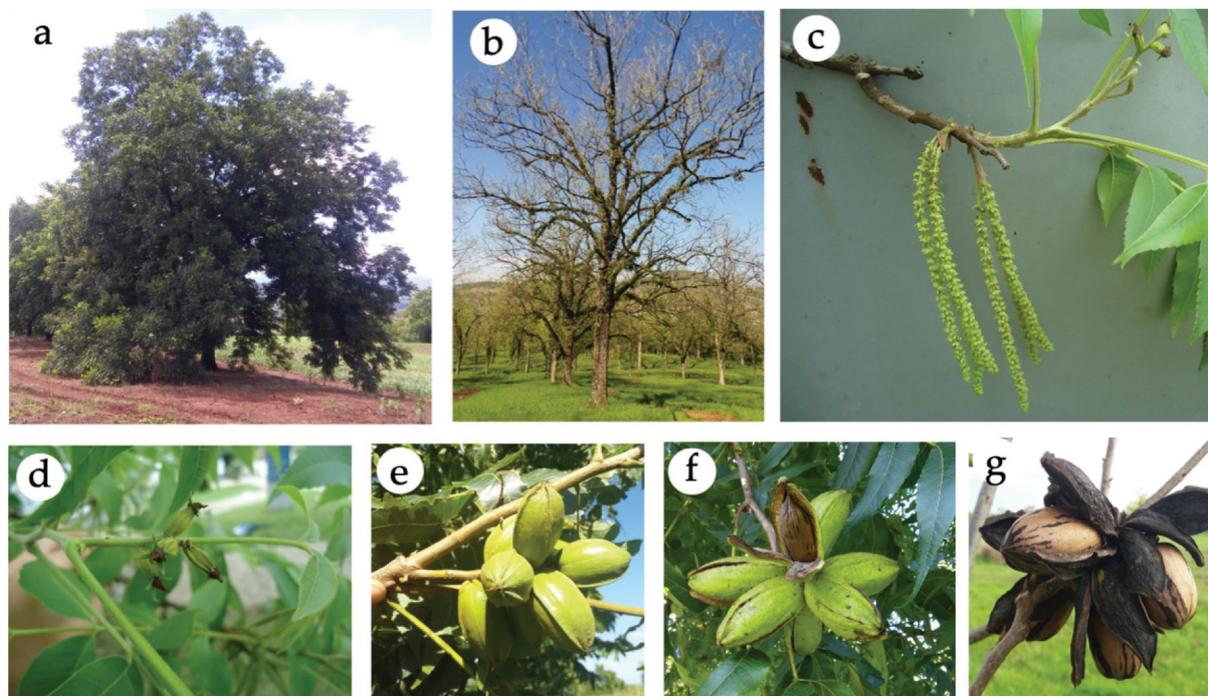


Figure 1. General characteristics of pecan. (a) Adult tree during the summer season, full of leaves; (b) adult tree during the blating season; (c) flowers of pecan (cv. Mahan); (d) cluster of pecan nuts (cv. Imperial) at the beginning of the development; (e) fruits of pecan cv. Jackson; (f) open fruit showing the nut (cv. Mahan); (g) mature fruits of pecan (cv. Jackson).

U.S.A. and 60 selected open-pollinated plants; Poletto et al., 2020), China (ca. 67 commercial cultivars from the U.S.A.; Zhang et al., 2015), and Spain (ca. 20 commercial cultivars from the U.S.A.; José M. G. Moreno, personal communication), maintained by governmental institutes and/or private orchards.

Pecaniculture is currently expanding and has been considered an attractive choice for the diversification of economic activity for rural properties (Filippin, 2011). Besides, there are favourable agronomic aspects as the possibility of inserting intercropping systems with pastures and livestock. There is also the possibility of intercropping with annual crops and other fruit species due to the wide edaphoclimatic adaptation, the longevity of production, and the easy storage of fruits after harvest (Duarte & Ortiz, 2001).

Considering this expansion, investing in pecan breeding programmes is a key factor in providing the production sector with productive and healthy cultivars. Improved pecan cultivars have traditionally been selected primarily based on nut characteristics, aiming at meeting commercial aspects, such as large nuts size, high yield, and high percent of kernel. Considering the expansion of the cultivated areas of pecan around the world, the current needs for genetic improvement of pecan should focus on aspects related to these new growth areas. The main traits required for the pecan growers are disease resistance, increased yields, the precocity of production, and reduced oscillation in the annual production.

Due to the species' long-life cycle, breeding programmes take a long-time span. Thus, the aid of

biotechnological tools like the marker-assisted selection is needed. Some worries regarding the lack of significant contributions of molecular tools for pecan breeding were shared by Conner (2015), who highlighted that this lack is not evidence that pecan breeders rejected biotechnological tools, neither that molecular markers are inefficient. Conner concluded that some barriers exist, that puzzle the application of molecular markers to aid in solving the needs of pecan breeding.

The reproduction system of pecan

Pecan is a deciduous tree with a straight trunk and an elliptical canopy that can reach a height of 50 metres (Figure 1(a–b)). As a monoecious plant, pecan has separated pistillate and staminate flowers on the same plant (Figure 1(c)). The pollination period matches the beginning of spring sprouting. The fruit is a nut enclosed in a dehiscent shell that splits along more or less winged sutures to drop the hard-shell nut (Grauke et al., 2016). Fruits group into bunches with three to seven units, consisting of an ovoid to oblong nut, surrounded by a thick green shell, which turns brown and opens when ripe (Figure 1(e–g)). The period required for fruit development varies from 150 to 180 days, depending on the cultivar (Varela, Takata, Camussi, & Zoppolo, 2015).

The development of pecan fruits comprises three main phases. The first, consists of its rapid enlargement, when the development of cotyledons occurs,

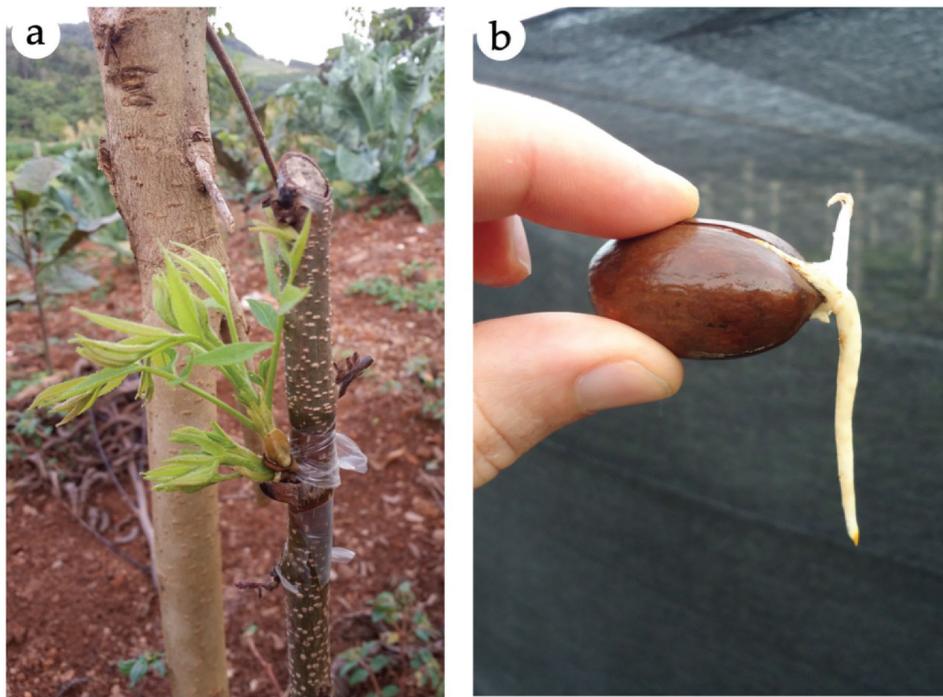


Figure 2. Propagation of pecan cultivars. (a) Grafting propagation of pecan cv. Imperial using an open-pollinated derived rootstock; (b) Germinating nut of pecan, after seed stratification treatment (Poletto et al., 2015).

continuing until the formation of the “small heart” stage and further, reaching the aqueous phase. In the second phase, the almond tissue develops inside the seed coating until it is filled (Figure 1(e–f)). In the final stage, the shell dehiscence occurs, separating from the nut and opening up, allowing it to be free of the shell (Figure 1(g)). The length of this development process differs according to the cultivar (Reid, 2015).

Pecan presents dichogamy, a reproductive system characterised by the release of pollen, and the receptivity of the pistil occurring at different times. Dichogamy is a morphological mechanism developed by the plant to prevent self-pollination, and, therefore, avoid inbreeding that can result in smaller nuts and low fruit yield (Wood, 2000).

The dichogamy system of pecan cultivars is not fixed. This reproductive system can vary according to each cultivar, the region and year of production, the age, and vigour of the plant. Even climatic conditions influence the maturation of pecan inflorescences. Periods with high relative humidity and temperatures result in the early maturation of staminate inflorescences. However, dry and cold periods anticipate the maturation of the pistil inflorescences and delay the maturation of the staminate inflorescences (Adriance, 1931).

Pecan cultivars are classified into two categories according to the dichogamy’s behaviour (Grauke & Thompson, 2012; Wood, 2000). The first is called protandrous, when the male flower ripens first, releasing pollen through the stamens. Despite being

commonly called ‘pollinators’, cultivars in this category also produce nuts with commercial value. Some examples of cultivars considered protandrous are Desirable, Western, Cheyenne, Caddo, Cape Fear, Success, Barton, and Cherokee. The second category is known as protogenic, when the female flower matures first so that the stigma becomes receptive before the stamens begin to release pollen. Protogenic cultivars are Choctaw, Wichita, Kiowa, Shawnee, Apache, Mohawk, Shoshoni, Chickasaw, Tejas, Elliott, Stuart, Comanche, and Schley.

Cross-pollination is necessary to achieve maximum pecan production. Thus, a high occurrence of cross-pollination is important, as it is one of the fundamental factors to obtain high productivity in commercial orchards (Wells, 2017). Orchards with the occurrence of inadequate pollination for a few years can experience a decrease in productivity and, consequently, in the quality of the nuts. Thus, in commercial orchards, cross-pollination should be promoted with the use of multiple cultivars. Moreover, commercial orchards should be composed of at least two pollinating cultivars, to increase the period of overlapping pollen dehiscence and the receptivity of stigma (Conner, 2015). As different trees have dissimilar bloom patterns, groups of trees are termed heterodichogamous. Heterodichogamy, in turn, promotes a significant genetic variation in pecan (Cerna-Cortés, Simpson, Martínez, & Martínez-Peniche, 2003).

Despite the evolutive importance of sexual reproduction and genetic diversity, the propagation of pecan has been carried out preferentially through



Figure 3. Morphological diversity of pecan nuts cultivated in Southern Brazil. (a) Size variability of different cultivated pecan accessions (scale in cm); (b) diversity in size and colour of the kernel of pecan accessions cultivated in Southern Brazil (scale in cm).

grafting, due to the large morphological variability existing in open-pollinated plants. Selected cultivars can be grafted (Figure 2(a)) over rootstocks produced using seeds collected from open-pollinated pecan trees (Figure 2(b)).

Molecular diversity in pecan: what we did and what we still have to do

As for other fruit crop species (e.g. Hajari, 2019), confusion concerning the naming and identification of cultivars is a challenge in pecan cultivation. Besides, open-pollination promotes high gene segregation, and the propagation through seeds may have introduced hundreds of varieties of pecan through the world. Such new varieties present great diversity in the shape of the fruits (Figure 3), tree architecture, quality of the nut, and different reproductive characteristics.

Although still incipient, the number of genetic studies on pecan using DNA-based molecular markers has increased (Table 1). Such studies, in turn, shed light on several aspects of genetic diversity and breeding of pecan both in its native range as well as in countries where it was introduced.

Genetic diversity in natural populations

Knowing the levels and patterns of genetic diversity and structure in natural populations of cultivated crop species is an important task. Besides ensuring the survival of the species, the conservation of wild populations of crop plants is important for providing genes valuable for breeding programmes of cultivated plants (Rüter, Hamrick, & Wood, 2000).

The patterns of genetic diversity accessed in four natural populations growing in the indigenous range

Table 1. Summary of molecular markers employed in pecan studies.

Marker	Cultivar	Country	Purpose	Reference
AFLP ^a	Native trees	Mexico	Genetic diversity	Cerna-Cortés et al., 2003
nSSR ^b and cpSSR ^c	Native trees	USA and Mexico	Genetic diversity	Grauke, Mendoza-Herrera, Miller, & Wood, 2011
RAPD ^d	Several	USA	Germplasm characterisation	Conner & Wood, 2001
cpSSR	Several	USA	Germplasm characterisation	Grauke, 2010
ISSR ^e and nSSR	Several	China	Germplasm characterisation	Jia, Wang, Zhai, Li, & Guo, 2011
nSSR	Several	China	Germplasm characterisation	JuanJuan, ShengYue, ShiQun, & ZhengJia, 2013
RAPD	Several	India	Germplasm characterisation	Kaur et al., 2016
RAPD	Several	Egypt	Germplasm characterisation	Elarabi & Elsoda, 2017
SNP ^f	Several	USA	Germplasm characterisation	Bentley, Grauke, & Klein, 2019
AFLP	Several	Brazil	Germplasm characterisation	Poletto et al., 2020
nDNA ^g	87MX3-2.11	USA	Nuclear genome sequencing	Jenkins, Wilson, Grimwood, Schmutz, & Grauke, 2015
nDNA ^g	Pawnee	China	Nuclear genome sequencing	Huang et al., 2019
cpDNA ^h	Imperial	Brazil	Plastid genome sequencing	Nagel et al., 2020
cpDNA ^h	Pawnee	China	Plastid genome sequencing	Mo et al., 2020
cpDNA ^h	Wichita	China	Plastid genome sequencing	Feng, Mo, & Peng, 2020
cpDNA ^h	87MX3-2.11 and Lakota	USA	Plastid genome sequencing	Wang & Shadgou, 2020

^aAmplified fragments length polymorphism; ^bnuclear single sequence repeats; ^cplastid single sequence repeats; ^drandom amplified polymorphic DNA; ^einter single sequence repeats; ^fsingle nucleotide polymorphism; ^gsequencing of the whole nuclear genome; ^hsequencing of the whole plastid genome.

of pecan's occurrence in Mexico using AFLP markers were considered relatively low. With 52.1% of poly-

plastid SSR markers of samples from 19 natural populations growing in the USA and Mexico (Grauke et al.,

Table 2. Genetic diversity and differentiation estimated for pecan cultivars using different molecular markers.

Reference	Marker	Genetic diversity estimation (H) ^a	Mean genetic distance among accessions (d) ^b
Conner & Wood, 2001	RAPD	–	0.44 (0.09–0.54)
Grauke, 2010	cpSSR ^c	0.24 (0.04–0.62)	0.508 (0.004–1.063)
Grauke, 2010	nSSR ^c	0.52 (0.47–0.56)	0.062 (0.044–0.086)
Jia et al., 2011	ISSR ^c	0.25 (0.22–0.27)	0.923 (0.828–0.974)
Jia et al., 2011	nSSR ^d	0.25 (0.24–0.26)	0.911 (0.811–0.969)
JuanJuan et al., 2013	nSSR	–	0.22 (0.05–0.39)
Elarabi & Elsoda, 2017	RAPD	–	0.83 (0.75–0.90)
Elarabi & Elsoda, 2017	ISSR	–	0.79 (0.62–0.94)
Poletto et al., 2020	AFLP ^c	0.37	0.22–0.85

^aGenetic diversity estimation for each marker and in each study was computed differently and should not be compared directly.

^bFor studies in which estimations of genetic similarity were presented, genetic distance was obtained as $d = 1 - \text{similarity}$.

^cGenetic diversity estimations obtained from cpSSR and nSSR markers range from $H = 0.0$ to $H = 1.0$, while estimations obtained from ISSR and AFLP markers (dominant markers) range from $H = 0.0$ to $H = 0.5$.

^dIn this study, nSSR markers were scored as dominant (presence/absence) and, therefore, the genetic diversity estimations range from $H = 0.0$ to $H = 0.5$.

morphic DNA fragments (196 out of 376 fragments), the pairwise genetic differentiation of trees ranging from 3% to 17% (Cerna-Cortés et al., 2003). This low differentiation among trees and overlap of individuals from different geographic origins in the UPGMA dendrogram was credited to the occurrence of gene flow through the local rivers and by anthropogenic action (Cerna-Cortés et al., 2003). Besides the characterisation of the genetic diversity in these populations, six AFLP markers were found to be associated with dichogamy. These markers were considered important tools for predicting this characteristic at the seedling stage, enabling the selection of plants in the establishment of pecan orchards, ensuring complete pollination, and, consequently, improving yield.

An equivalent pattern of genetic diversity and relatedness was reported in an analysis of nuclear and

2011). Relatively low genetic diversity ($h = 0.516$ for three cpSSR loci and $H_O = 0.461$ for 14 nSSR loci) but moderate to high number of alleles (mean total number of alleles $A = 4.33$ and mean effective number of alleles $A_e = 2.18$ for cpSSR loci, and $A = 10.21$ and $A_e = 4.37$ for nSSR loci) were reported. Whereas estimations of genetic diversity reflect the patterns of crossing and evolutive forces (as genetic drift, migration, and selection), the number of alleles is quite important for the responses of populations to biotic and abiotic stresses. Besides, a high number of alleles provide a higher opportunity for finding allelic combinations that are of interest to the genetic improvement of crop species. This analysis of cpSSR and nSSR also revealed a geographic structuring of the sampled trees in three main clusters, with slight differences concerning the structure based on cpSSR and nSSR markers (Grauke

et al., 2011). Such differences are likely related to the different heritage mode of each marker. While plastids (cpSSR loci) are inherited from only one of the genitors (usually the mother tree in angiosperms), the nuclear genome (nSSR loci) are inherited from both pollen and ovule donors. Additionally, plastid genomes virtually lack recombination and present a lower mutation rate in comparison to nuclear genomes (Stefenon, Klabunde, Lemos, Rogalski, & Nodari, 2019). Due to the higher structure revealed for the analysed populations, cpSSR markers were suggested as useful tools for initiatives towards identifying and conserving a representative sample of pecan's germplasm in *ex situ* collections (Grauke et al., 2011).

Genetic relatedness among cultivars

Provenances and cultivar collections are expected to hold lower genetic diversity than natural populations since they represent samples of the last ones. However, adaptedness and adaptability have been considered fundamental issues for forestry. Such capacity of populations to react to biotic and abiotic stresses by evolutionary adaptations deeply depends on the maintenance of high levels of genetic diversity, both in natural and planted forests (Stefenon, Gailing, & Finkeldey, 2008). Moreover, the maintenance of the evolutionary adaptive potential is an important issue if it is intended to use reproductive material from planted forests (Finkeldey & Hattemer, 2007).

Studies aiming at the characterisation of the molecular genetic diversity and relatedness among cultivars and/or accessions of germplasm collections and/or orchards have been implemented in different countries where pecan is cultivated. Overall, germplasm collections evaluated in China (Jia et al., 2011; JuanJuan et al., 2013), Egypt (Elarabi & Elsoda, 2017), India (Kaur et al., 2016), and Brazil (Poletto et al., 2020) revealed a relatively low level of genetic diversity (Table 2), likely as an effect of the initial introduction of the culture in these countries with few specimens and consequently, a narrow genetic basis.

Based on data from cpSSR markers, the *ex situ* germplasm collection of the National Clonal Germplasm Repository for Pecans and Hickories (NCGR-*Carya*) in the USA revealed a pattern of structured genetic diversity. This pattern was larger in accessions collected in Mexico, decreased to the north, and was lowest in accessions from the Northwest USA (Grauke et al., 2010). For this germplasm collection, the mean number of alleles ranged from $A = 1.33$ to $A = 3.67$, the mean effective number of alleles from $A_e = 1.04$ to $A_e = 2.71$, and genetic diversity from $h = 0.037$ to $h = 0.63$, with nine haplotypes reported in plants collected from Mexican and two haplotypes in plants sampled in the northern range of the species occurrence in the USA. These cpSSR markers were also employed to confirm the maternal inheritance of the plastid genome in *Carya* and the oldest origin of Mexican populations, preceding the Pleistocene glaciations (Grauke et al., 2010). Different from the cpSSR markers, nSSR markers revealed the highest number of alleles and the effective number of alleles ($A = 7.57$ and $A_e = 4.07$, respectively) in plants sampled from Texas (USA) and expected heterozygosity $H_e = 0.558$ in plants from the northwestern USA. The lowest values were observed in plants from the northeastern USA ($A = 5.86$ and $A_e = 3.44$) and Mexico ($H_e = 0.472$) (Grauke et al., 2010).

Ten ISSR and eight nSSR marker loci were employed to genetically characterise a pecan's germplasm collection from China, including cultivars introduced from the USA, Chinese seedling breeding cultivars, and selected, non-registered accessions (Jia et al., 2011). Despite the different nature of these two markers, the reported diversity parameters are quite similar. ISSR is biallelic dominant markers, which polymorphism is based on the presence/absence of an amplified DNA fragment. nSSRs are multiallelic codominant markers with the polymorphism based on the number of repeat motifs of the amplified allele. However, in this study, nSSR loci were scored as presence/absence of each allele and not as a codominant marker. This scoring strategy enables the comparison of nSSR with ISSR markers but triggers the loss of the high informative capacity at locus level, characteristic

Table 3. Overall characteristics of the sequenced and characterised plastomes of five pecan cultivars.

Feature	Cultivar/accession				
	Imperial ^a	Pawnee ^b	87MX3-2.11 ^c	Lakota ^c	Wichita ^d
GenBank ID	MN221384	MN977124	MH909600	MH909599	MT044463
Plastome size (bp)	160,818	160,819	160,545	160,819	160,532
SSC (bp)	18,791	18,791	18,576	18,790	18,751
LSC (bp)	90,041	90,042	89,933	90,041	89,799
IRs (bp)	25,993	25,993	26,018	25,994	25,991
GC content (%)	36.14	36.15	36.2	36.1	36.19
Number of protein-coding genes	78	86	84	84	83
Number of tRNA-coding genes	37	37	32	32	37
Number of rRNA-coding genes	8	8	8	8	8
Total number of genes	123	131	124	124	128

^aNagel et al. (2020), ^bMo et al. (2020), ^cWang and Shadgou (2020), ^dFeng et al. (2020).

of nSSR markers. The estimated number of alleles for ISSR and nSSR loci in this study was $A = 2.00$ for both markers, the effective number of alleles was $A_e = 1.46$ and $A_e = 1.42$, respectively, and the genetic diversity was $H = 0.42$ and $H = 0.41$, respectively (Jia et al., 2011). When the genetic parameters were estimated only for the North American cultivars, slightly lower values were obtained: $A = 1.73$ and $A = 1.78$, $A_e = 1.36$ and $A_e = 1.38$, and $H = 0.22$ and $H = 0.24$ for ISSR and nSSR markers, respectively (Jia et al., 2011). The pairwise genetic differentiation of the plants based on the combination of data from both ISSR and nSSR markers ranged from $d = 0.05$ to $d = 0.41$. The mean genetic differentiation among the North American cultivars was slightly lower than observed among Chinese cultivars ($d = 0.21$ and $d = 0.25$, respectively; Jia et al., 2011). The clustering of all investigated plants revealed no clear pattern of cultivars grouping considering their origin, suggesting high genetic admixture among North American, Chinese, and the non-registered accessions.

With the aid of nSSR loci scored as codominant markers, JuanJuan et al. (2013) evaluated 37 North American cultivars of pecan planted in China, reporting a mean of eight alleles per locus, ranging from three to 20 alleles. The pairwise genetic differentiation between trees ranged from $d = 0.045$ to $d = 0.393$, with an average of $d = 0.219$ (JuanJuan et al., 2013).

Seven pecan cultivars (Wichita, Grazona, Desirable, Burkett, Mahan, MoneyMaker, and Cheyenne) growing in Egypt were evaluated concerning their molecular genetic relatedness using dominant RAPD and ISSR markers (Elarabi & Elsoda, 2017). The mean pairwise genetic differentiation accessed was $d = 0.83$, ranging from $d = 0.75$ to $d = 0.90$ for RAPDs and $d = 0.79$ ranging from $d = 0.62$ to $d = 0.94$ for ISSRs. Forty-five RAPD and 14 ISSR markers were considered diagnostic for specific cultivars (Elarabi & Elsoda, 2017). However, given the use of one single individual of each cultivar in this study, these markers should not be established as markers associated with these cultivars or as markers associated with economic traits, as claimed by the authors.

Dominant AFLP markers were used to study the molecular genetic relatedness among accessions of pecan growing in Brazil (Poletto et al., 2020). The samples included commercial North American cultivars and selected accessions with potential commercial use but not registered as cultivars. The genetic diversity level observed for the AFLP markers ($H = 0.37$) was considered moderate to high. The pairwise differentiation based on 154 polymorphic AFLP markers between accessions ranged from $d = 0.22$ to $d = 0.85$ (Poletto et al., 2020). No correlation was observed among these molecular markers with morphological traits of the fruits and/or chemical composition of the nuts from each accession. However, the usefulness of

molecular markers in accessing the genetic diversity of pecan collections, aiming at planning management of orchards and breeding of the species, was validated.

Organellar genomics

Low variation is also observed in comparing the whole plastid genomes (plastomes) sequenced and characterised for five cultivars of pecan (Feng et al., 2020; Mo et al., 2020; Nagel et al., 2020; Wang & Shadgou, 2020). Features related to complete plastome size, the size of their partitions, the content of GC bases, and the number of genes (Table 3), reveals relative conservation among cultivars, although they are the result of different crossings.

The characterised plastomes presented a total length ranging from 160,818 to 160,545 bp, with the small single copy (SSC) region ranging from 18,576 to 18,791 bp and the large single copy (LSC) region from 89,799 to 90,043. The two inverted regions (IRa and IRb) presented lengths ranging from 25,001 to 26,018 bp, while the GC content oscillated around 36.1% and 36.2% (Table 3). Most importantly, the total number of genes of the plastomes ranged from 123 in cultivar Imperial to 131 in cultivar Pawnee (Table 3). The differences in the number of genes are observed in protein-coding genes (ranging from 78 to 86) and in tRNA-coding genes (ranging from 32 to 37). These results suggest the loss of genes in the plastome of some cultivars or annotation error. While gene loss or transfer of nuclear DNA to chloroplasts recurrently occurred during the evolution of the angiosperm plastome, such occurrence is not expected within the short timespans in which the commercial cultivars studied were obtained from different crossings or selections. On the other hand, annotation errors may be caused by inaccuracies in the sequencing, in the sequence assembly, or due to bias in the annotation software and are the most likely origin of such differences. However, ancient gene loss in the genitor plants that originated the studied cultivars cannot be ruled out and further studies about these differences observed in the plastomes should be performed to elucidate if they represent evolutive events or annotation troubles.

Nuclear genomics

The whole-genome sequencing of reference genomes has been proposed as an important tool towards hastening the genetic improvement of pecan (Huang et al., 2019). A first sequence draft of the nuclear genomes from the Mexican-derived accession 87MX3-2.11 and the widely planted cultivar Pawnee was generated in 2015 (Jenkins et al., 2015). The SNP calling revealed approximately 3.3 million SNPs for the 'Pawnee' draft genome, of which more than 1.3 million were homozygous and nearly 2.0 million

were heterozygous against the 87MX3-2.11 genome (Jenkins et al., 2015). The SNP rate estimated was 4.85 SNPs per kilobase for cultivar Pawnee and 1.55 SNPs per kilobase for accession 87MX3-2.11. From these results, the genome sequence of accession 87MX3-2.11 was suggested as the best choice from which to produce a reference genome for breeding proposals, given their lower variation in comparison to other pecan genotypes used in breeding programmes.

Four years later, a virtually complete reference pecan genome sequence of cultivar Pawnee was published by Huang et al. (2019), providing an assembling covering more than 97% of the estimated genome size (649.75 Mb) of the species. This sequencing effort allowed the annotation of 31,075 protein-coding genes, 378 miRNAs, 378 tRNAs, and 198 rRNAs. Around 95.7% of the predicted genes revealed homology with known functions when compared to records from different databases (Huang et al., 2019). Through the analysis of the predicted genes and metabolic pathways, the identification of genes involved in the nutritive characteristics of pecan was possible. This analysis revealed pecan as one of the oil plants with the highest number of genes related to oil synthesis. The enlarged gene copies and their high expression levels probably impact the high oil level in the nuts of pecan (Huang et al., 2019).

Accordingly, a study of 60 pecan accessions from Southern Brazil evaluated the mean content of ether extract (lipids) in pecan nuts, revealing a mean of 70.3% of the kernel mass. This level of lipids from the pecan nuts is higher than reported for other nuts such as cashew, pistachio, hazelnuts, macadamia, pine nut, peanut, walnut, and almond (Poletto et al., 2020).

Molecular markers in pecan's breeding

Molecular markers are specifically valuable for breeding programmes of perennial crops, mainly through marker-assisted selection, for the construction of linkage maps, and the mapping of qualitative and quantitative traits. Up to date, two studies tentatively associated molecular markers to the phenotypic trait type of dichogamy (protandrous or protogynous) in pecan.

Cerna-Cortés et al. (2003) identified six AFLP markers putatively associated with dichogamy in pecan. However, the real association of these markers with dichogamy still needs to be confirmed. Once this association is established, such markers can be used to develop SCAR (sequence characterised amplified region) markers. Such PCR-based SCAR markers would be easily generated and very useful for selecting protandrous or protogynous genotypes in breeding programmes.

A genetic linkage maps of two pecan cultivars (Pawnee and Elliot) was built based on 258 AFLP and 217 RAPD markers, as well as data about

dichogamy type and stigma colour (two phenotypic traits much likely under monogenic control; Beedanagari et al., 2005). The molecular markers generated maps covering around 83% of the Pawnee genome and 57% of the Elliot genome. Dichogamy type and stigma colour were closely associated (1.9 cM) in one Elliot linkage group, located between the AFLP marker ACA/CAT250 (15.4 cM from the dichogamy type locus) and the RAPD marker B29-340 (9.4 cM from the stigma colour locus) (Beedanagari et al., 2005). Mapping markers linked to dichogamy types in pecan is quite useful because the growth of cultivars of both dichogamy types is needed to ensure that orchards combine cultivars that successfully pollinate each other.

The advances in the sequencing of reference nuclear genomes of pecan, in turn, open further possibilities for implanting programmes of marker-assisted selection and for generating highly saturated maps based on SNP markers and most recently proposed mapping strategies as association mapping and genomic selection. Although the accession 87MX3-2.11 was considered a candidate to be the basis for a complete reference project for pecan given its low variation rate (Jenkins et al., 2015), the nuclear genome of cultivar Pawnee was the first completely sequenced. Sequencing this genome supported the deep evaluation of several traits of importance for breeding programmes. For instance, family genes related to response to biotic and abiotic stresses, and genes and metabolic pathways concerning the production of nutritional compounds were identified (Huang et al., 2019).

Moreover, because genetic diversity is crucial to perform genetic improvement, accessing patterns of genetic diversity in pecan collections is also important. For example, the results of an AFLP-based analysis of genetic diversity in pecan accessions growing in Brazil supported the potential of using accessions selected from pecan orchards as a source of germplasm for species breeding in the country (Poletto et al., 2020).

The future of the pecaniculture: needs and perspectives concerning molecular markers

The advances of the sequencing platforms have turned genome studies relatively easier, quicker, and cheaper. However, the demand for specialised laboratories and personal is still a bottleneck for the use of these molecular strategies in many studies. Thus, advancement in optimising the use of more classical PCR-based molecular markers is needed for the advance of the pecaniculture. Nuclear and plastid SSRs and AFLPs have shown to be quite useful markers to assess the genetic diversity and structure of natural populations and germplasm collections. Additionally, some markers may be linked to traits of interest for breeding and

management of orchards (Beedanagari et al., 2005; Cerna-Cortés et al., 2003), but no further studies were published regarding these topics.

In addition, new markers may be introduced for studying pecan. The sequenced nuclear genome of the pecan cultivar Pawnee revealed 50.43% of its content composed of repetitive elements, of which approx. 85% are transposable elements. Long terminal repeats (LTRs) make up the majority of the transposable elements in this genome, with *Gypsy*-like and *Copia*-like elements representing 15.06% and 15.41%, respectively (Huang et al., 2019). Markers based on retrotransposons can be used for the identification and description of genotypes, as they generally show chromosomal dispersion, a variable number of copies, and random distribution in the genome (Kalendar, Grob, Regina, Suoniemi, & Schulman, 1999). Thus, in addition to SSRs and AFLPs, the use of markers based on the amplification of DNA fragments using AFLP and LTR-anchored primers (S-SAP markers; Waugh et al., 1997) may be an interesting alternative to find markers linked to genes of agronomic interest, given the wide distribution of both AFLP sites and retrotransposons throughout the genome.

Taking into account the advances of the molecular genetic tools, allied to the need of improving the production of pecan nuts, the next years tend to see intense progress in molecular genetic studies of this species, towards breeding intensification. The barriers puzzling the application of molecular markers to aid in solving the needs of pecan breeding will be overleaped and these tools will become daily instruments helping breeders and farmers.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico [302502/3027-7]; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior [001].

ORCID

Luana Oliveira de Oliveira  <http://orcid.org/0000-0001-7472-4149>

Dalvan Carlos Beise  <http://orcid.org/0000-0001-8638-7004>

Daniele Damian Dos Santos  <http://orcid.org/0000-0002-3795-1796>

Jordana Caroline Nagel  <http://orcid.org/0000-0001-6876-3838>

Tales Poletto  <http://orcid.org/0000-0002-6162-4445>

Igor Poletto  <http://orcid.org/0000-0002-2441-0456>

Valdir Marcos Stefenon  <http://orcid.org/0000-0003-1091-700X>

References

- Adriance, G.W. (1931). Factors influencing fruit setting in the pecan. *Botanical Gazette*, 91, 144–166. doi:10.1086/334136
- Beedanagari, S.R., Dove, S.K., Wood, B.W., & Conner P.J. (2005) A first linkage map of pecan cultivars based on RAPD and AFLP markers. *Theor Appl Genet*, 110, 1127–1137. doi:10.1007/s00122-005-1944-5
- Bentley, N., Grauke, L.J., & Klein, P. (2019). Genotyping by sequencing (GBS) and SNP marker analysis of diverse accessions of pecan (*Carya illinoensis*). *Tree Genetics & Genomes*, 15, 8. doi:10.1007/s11295-018-1314-5
- Cerna-Cortés, J.F., Simpson, J., Martínez, O., & Martínez-Peniche, R.A. (2003). Measurement of genetic diversity of native pecan [*Carya illinoensis* (Wangenh) K. Koch] populations established in Central Mexico and correlation with dichogamous flowering using AFLP. *Food, Agriculture & Environment*, 1, 168–173.
- Conner, P.J. (2015). ROADBLOCKS AND HINDRANCES TO THE DEVELOPMENT OF MOLECULAR TOOLS IN PECAN (*Carya illinoensis*): A Breeder's PERSPECTIVE. *Acta Horticulturae*, 1070, 83–88. doi:10.17660/ActaHortic.2015.1070.9
- Conner, P.J., & Wood, B.W. (2001). Identification of pecan cultivars and their genetic relatedness as determined by randomly amplified polymorphic DNA analysis. *Journal of the American Society of Horticultural Sciences*, 126, 474–480. doi:10.21273/JASHS.126.4.474
- Duarte, V., & Ortiz, E.R.N. (2001). Podridão de Phytophthora amêndoa e da casca da noqueira pecan. In: dos Santos, A.F., Luz, E.D.M.N., & Bezerra, J.L. (Ed.), *Doenças causadas por Phytophthora no Brasil* (pp. 493–508). Campinas: Rural.
- Elarabi, N.I., & Elsoda, A.S. (2017). Molecular evaluation of genetic diversity among seven genotypes of pecan (*Carya Illinoensis*). *Journal of Agricultural Chemistry and Biotechnology*, 8, 27–33. doi:10.21608/jacb.2017.38425
- Feng, G., Mo, Z.H., & Peng, F.R. (2020). The complete chloroplast genome sequence of *Carya illinoensis* cv. wichita and its phylogenetic analysis. *Mitochondrial DNA Part B*, 5, 2235–2236. doi:10.1080/23802359.2020.1768925
- Filippin, I.L. (2011). *Viabilidade econômica do cultivo de noqueira pecã em áreas de reserva legal e de preservação permanente*. 74 f (MSc. Dissertation). Universidade Federal de Pelotas, Pelotas, Brazil.
- Finkeldey, R., & Hattemer, H.H. (2007). *Tropical forest genetics*. Berlin: Springer.
- Grauke, L.J., & Thompson, T.E. (2012). Pecan pollination. In: L.A. Stein, G.R. Mceachern, & M.L. Nesbitt (Eds.), *Texas Pecan Handbook* (pp. 199). College Station: Texas A & M AgriLife Extension Service.
- Grauke, L.J., Mendoza-Herrera, M.A., & Binzel, M.L. (2010) Plastid Microsatellite Markers in *Carya*. *Acta Hort*, 859, 237–246.
- Grauke, L.J., Mendoza-Herrera, M.A., Miller, A.J., & Wood, B.W. (2011). Geographic patterns of genetic variation in native pecans. *Tree Genetics & Genomes*, 7, 917–932. doi:10.1007/s11295-011-0384-4
- Grauke, L.J., Wood, B.W., & Harris, M.K. (2016). Crop vulnerability: *Carya*. *HortScience*, 51, 653–663. <https://doi.org/10.21273/HORTSCI.51.6.653>.

- Hajari, E. (2019). Molecular and related approaches to litchi improvement – Historical perspective and future trends. *The Journal of Horticultural Science and Biotechnology*, 94, 693–702. doi:10.1080/14620316.2019.1624202
- Huang, Y., Xiao, L., Zhang, Z., Zhang, R., Wang, Z., Huang, C., ... Huang, J. (2019). The genomes of pecan and Chinese hickory provide insights into *Carya* evolution and nut nutrition. *Gigascience*, 8, giz036. doi:10.1093/gigascience/giz036
- INC – International Nut and Dried Fruit Council. (2019). Global statistical review: Crop forecast report. *Nutfruit*, 77, 65–78.
- Jenkins, J., Wilson, B., Grimwood, J., Schmutz, J., & Grauke, L.J. (2015). Towards a reference pecan genome sequence. *Acta Horticulturae*, 1070, 101–108. doi:10.17660/ActaHortic.2015.1070.11
- Jia, X.-D., Wang, T., Zhai, M., Li, T.-R., & Guo, Z.-R. (2011). Genetic diversity and identification of Chinese-grown pecan using ISSR and SSR markers. *Molecules*, 16, 10078–10092. doi:10.3390/molecules161210078
- JuanJuan, S., ShengYue, Y., ShiQun, Y., & ZhengJia, W. (2013). SSR analysis of genetic diversity of the 37 new pecan cultivars. *Journal of Anhui Agricultural University*, 40, 42–46.
- Kalendar, R., Grob, T., Regina, M., Suoniemi, A., & Schulman, A.H. (1999). IRAP and REMAP: Two new retrotransposon-based DNA fingerprinting techniques. *Theoretical and Applied Genetics*, 98, 704–711. doi:10.1007/s001220051124
- Kaur, J., Bhardwaj, E., Shilpa, C., Kumar, K., Singh, D., & Sharma, D. (2016). Characterization of pecan (*Carya illinoensis*) genotypes using RAPD markers. *International Journal of Farm Sciences*, 6, 77–81.
- Mo, Z., Lou, W., Chen, Y., Jia, X., Zhai, M., Guo, Z., & Xuan, J. (2020). The chloroplast genome of *Carya illinoensis*: Genome structure, adaptive evolution, and phylogenetic analysis. *Forests*, 11, 207. doi:10.3390/f11020207
- Nagel, J., Machado, L.O., Lemos, R.P.M., D'Oliveira Matielo, C. B., Poletto, T., Poletto, I., & Stefenon, V.M. (2020). Structural, evolutionary and phylogenomic features of the plastid genome of *Carya illinoensis* cv. Imperial. *Annals of Forest Research*, 63, 3–18. doi:10.15287/afr.2019.1413
- Poletto, I., Muniz, M.F.B., Poletto, T., Stefenon, V.M., Baggio, C., & Ceconi, D.E. (2015). Germination and development of pecan cultivar seedlings by seed stratification. *Pesqui. Agropecu. Bras.*, 50, 1232–1235. doi:10.1590/S0100-204X2015001200014
- Poletto, T., Poletto, I., Silva, L.M.M., Muniz, M.F.B., Reiniger, L.R., Richards, N., & Stefenon, V.M. (2020). Morphological, chemical and genetic analysis of southern Brazilian pecan (*Carya illinoensis*) accessions. *Scientia Horticulturae*, 261, 108863. doi:10.1016/j.scienta.2019.108863
- Reid, W.R. (2015). The phenology of early fruit ripening in pecan. *Acta Horticulturae*, 1070, 27–32. doi:10.17660/ActaHortic.2015.1070.2
- Rüter, B., Hamrick, J.L., & Wood, B.W. (2000). Outcrossing rates and relatedness estimates in pecan (*Carya illinoensis*) populations. *The Journal of Heredity*, 9, 72–75. doi:10.1093/jhered/91.1.72
- Sparks, D. (2005). Adaptability of pecan as a species. *HortScience*, 40, 1175–1189. doi:10.21273/HORTSCI.40.5.1175
- Stefenon, V.M., Gailing, O., & Finkeldey, R. (2008). Genetic structure of plantations and the conservation of genetic resources of Brazilian pine (*Araucaria angustifolia*). *Forest Ecology and Management*, 255, 2718–2725. doi:10.1016/j.foreco.2008.01.036
- Stefenon, V.M., Klabunde, G., Lemos, R.P.M., Rogalski, M., & Nodari, R.O. (2019). Phylogeography of plastid DNA sequences suggests post-glacial southward demographic expansion and the existence of several glacial refugia for *Araucaria angustifolia*. *Scientific Reports*, 9, 2752. doi:10.1038/s41598-019-39308-w
- Varela, V., Takata, V., Camussi, G., & Zoppolo, R. (2015). Pecan: Viability of a new crop in Uruguay. *Acta Horticulturae*, 1070, 245–251. doi:10.17660/ActaHortic.2015.1070.28
- Wang, X., & Shadgou, H. (2020). Chloroplast genome sequences of *Carya illinoensis* from two distinct geographic populations. *Tree Genetics & Genomes*, 16, 48. doi:10.1007/s11295-020-01436-0
- Waugh, R., McLean, K., Flavell, A.J., Pearce, S.R., Kumar, A., Thomas, B.B., & Powell, W. (1997). Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Molecular Genetics and Genomics*, 253, 687–694. doi:10.1007/s004380050372
- Wells, L. (2017). *Southeastern pecan grower's handbook* (pp. 236). Georgia: University of Georgia Cooperative Extension Service.
- Wood, B.W. (2000). Pollination characteristics of pecan trees and orchards. *Hort Technology*, 10, 120–126. doi:10.21273/HORTTECH.10.1.120
- Zhang, R., Peng, F., & Li, Y. (2015). Pecan production in China. *Scientia Horticulturae*, 197, 719–727. doi:10.1016/j.scienta.2015.10.035

CAPÍTULO 2.

CARACTERIZAÇÃO MOLECULAR DE ACESSOS DE NOGUEIRA-PECÃ *Carya illinoensis* (WANGENH.) K. KOCH NO RIO GRANDE DO SUL USANDO MARCADORES S-SAP

1. INTRODUÇÃO

A noqueira-pecã (*Carya illinoensis* (Wangenh.) K. Koch (Jugladiaceae) é uma frutífera cultivada predominantemente nas regiões temperadas do hemisfério norte (BRISON, 1974; SPARKS, 2005; WALKER et al., 2016). Espécie nativa dos Estados Unidos e México, a sua popularização impulsionou o cultivo em vários países e diferentes continentes, como na China, África do Sul, Austrália, inclusive na América do Sul, abrangendo Uruguai, Argentina, Chile, Peru e Brasil (SPARKS, 1991; WELLS, 2017). No Brasil, a noqueira-pecã foi introduzida em 1870 no Estado de São Paulo, pelos imigrantes norte-americanos, mas foi somente por volta de 1960-1970 que a cultura passou a ser explorada comercialmente, desde o estado de Minas Gerais até o Rio Grande do Sul (RASEIRA, 1990; POLLETO et al., 2014).

Na natureza, observou-se que os frutos fazem parte da dieta de diversas aves, raposas, esquilos, gambás, porcos selvagens e guaxinins, além de muito apreciados pelos humanos, diante do seu valor nutricional. O interesse pelo cultivo dessa frutífera vem crescendo fortemente no sul do Brasil, apoiado pelo aumento do consumo e o preço de mercado. Aliado a isso, os produtores vêm identificando atrativos nesse cultivo, como o baixo custo de implantação dos pomares, comparado ao de outras frutíferas, colheita fora da época tradicional de grãos, proporcionando não só a expansão do plantio de noqueira-pecã, como a recuperação dos pomares abandonados (FRONZA; HAMANN, 2016; POLETO et al., 2014; WALKER, 2016).

Diante da apreciação do seu consumo e os benefícios da sua cultura, a noqueira-pecã vem tendo uma grande distribuição geográfica na sua produção. Com isso, estudos visando à caracterização da diversidade genética molecular e a relação entre cultivares e/ou ascendências de coleções de germoplasma e/ou pomares vem sendo implementados em diferentes países, onde a pecã é cultivada. As coleções de germoplasma avaliadas na China (JIA et al. 2011, JUANJUAN et al., 2013), Egito (ELARABI & ELSODA, 2017), Índia (KAUR et al., 2016) e Brasil (POLETO et al., 2020) revelaram um nível relativamente baixo de diversidade genética, provavelmente como efeito da introdução inicial da cultura nestes países com poucos espécimes e, conseqüentemente, uma base genética estreita.

Analisando a estrutura das populações com marcadores, os marcadores cpSSR, foram sugeridos como ferramentas úteis para identificar e conservar uma amostra representativa do germoplasma da noz-pecã em coleções *ex situ* (GRAUKE et al., 2011). Com base nos dados dos marcadores cpSSR, a coleção de germoplasma *ex situ* do Repositório Nacional de Germoplasma Clonal para Pecans e Hickories (NCGR-*Carya*) nos EUA revelou um padrão de

diversidade genética estruturada que foi maior nas adesões recolhidas no México, diminuiu para o norte, e foi menor nas adesões do Noroeste dos EUA (GRAUKE et al., 2010).

Para esta coleção de germoplasma, o número médio de alelos variou de $A = 1,33$ a $A = 3,67$, o número médio efetivo de alelos de $A_e = 1,04$ a $A_e = 2,71$, e a diversidade genética de $h = 0,037$ a $h = 0,63$, com nove haploides reportados em plantas colhidas no México e dois haploides em plantas amostradas na gama norte da ocorrência da espécie nos EUA.

Dez ISSR e oito nSSR marcadores foram utilizados para caracterizar geneticamente uma coleção de germoplasma de noz-pecã da China, incluindo cultivares introduzidas dos EUA, cultivares de reprodução de plântulas chinesas, e algumas adesões não registradas (JIA et al., 2011). Apesar da natureza diferente destes dois marcadores, os parâmetros de diversidade relatados são bastante semelhantes. A ISSR é um marcador bialélico dominante, cujo polimorfismo se baseia na presença/ausência de um fragmento de DNA amplificado.

O número estimado de alelos para ISSR e nSSR loci neste estudo foi $A = 2,00$ para ambos os marcadores, o número efetivo de alelos foi $A_e = 1,46$ e $A_e = 1,42$, respectivamente, e a diversidade genética foi $H = 0,42$ e $H = 0,41$, respectivamente. Quando os parâmetros genéticos foram estimados apenas para as cultivares norte-americanas, foram obtidos valores ligeiramente inferiores: $A = 1,73$ e $A = 1,78$, $A_e = 1,36$ e $A_e = 1,38$, e $H = 0,22$ e $H = 0,24$ para os marcadores ISSR e nSSR, respectivamente (JIA et al., 2011).

A sustentabilidade de uma espécie está condicionada a variabilidade genética desta, pois determina as possibilidades de adaptação, sobrevivência e viabilidade da espécie sob pressões ambientais (VINSON et al., 2015).

Estudos feitos com elementos transponíveis são bem variados e apresentam-se constantes, não somente em plantas, mas também em animais, como por exemplo, em *Locusta migratoria*, é um inseto com um tamanho de genoma grande, que apresentou em seu genoma muitos retroelementos, e foi feita uma análise dos dados de sequenciamento para elucidar as características de diversidade, estrutura e expressão de retroelementos (JIANG et al., 2012).

Em plantas, análises de sequências são feitas para investigar os elementos transponíveis para entender a evolução de genomas, já foi demonstrado que o uso de uma pequena sequência a partir de um genoma de planta permite a descoberta e descrição da sua dinâmica para estudos filogenéticos, estes estudos devem ser prosseguidos através de linhagens de plantas com a seleção filogenética apropriada, para ajudar entender a atividade e comportamento dos elementos transponíveis, cujos tem sido responsáveis pela variação do genoma através da evolução (ESTEP et al., 2013).

Retrotransposons também foram utilizados como marcadores moleculares em pêra japonesa (*Pyrus pyrifolia*), foram desenvolvidos através da técnica baseado em polimorfismo de inserção (RBIP) que são marcadores com base nas regiões LTRs (KIM et al., 2012). Com as inovações dos estudos evolutivos, podemos constatar que o uso de elementos transponíveis faz e continuará fazendo parte das pesquisas, que auxiliam para compreensão dos genomas e sua evolução, estudos de genótipos e sua diversidade genética.

É importante ressaltar que a demanda da sociedade por informação e orientação técnica sobre a cultura vem aumentando consideravelmente. Além das indagações por parte de produtores e órgãos de assistência técnica, algumas mobilizações evidenciam ainda mais a necessidade de pesquisas visando superar as lacunas tecnológicas do setor com a relação aos estudos genéticos.

Diante disso, o objetivo deste estudo em caracterizar a diversidade genética de acessos de noqueira-pecã (*Carya illinoenses*) no Rio Grande do Sul utilizando marcadores moleculares S-SAP, pode ser considerado um avanço para melhor elucidar as dúvidas e gerar mais informações a respeito da espécie.

2. MATERIAL E MÉTODOS

2.1 Amostragem do material vegetal

As amostras utilizadas para este estudo foram coletadas em pomares existentes no estado do Rio Grande do Sul, especificamente dos produtores da região do município de Anta Gorda, as amostras foram escolhidas com base na informação obtida dos produtores, levando em consideração características morfológicas, relacionadas principalmente ao tamanho e forma das castanhas. A maioria das árvores é derivada de enxertia utilizando as plantas locais, e outras originaram de sementes. Entretanto, todas as amostras são de árvores com mais de 50 anos, mantidas nos pomares ou nos jardins ao redor das casas dos agricultores. As amostras de cada árvore consistem em folhas jovens desprovidas de lesões e sem sinais de doença foliar, e foram acondicionadas em sacos hermeticamente fechados contendo sílica gel para retirar a umidade e conservar a planta para posterior extração de DNA.

2.2 Extração de DNA

O DNA genômico total foi isolado das amostras de folhas de cada acesso usando o protocolo CTAB (Doyle & Doyle, 1987). Foram colocados aproximadamente 200 mg de folhas picadas em tubos tipo Eppendorf de 2mL, sendo inseridas duas *beads* em seu interior. A seguir, os tubos foram colocados no TissueLyser (Qiagen®), durante dois ciclos de 30 segundos, para a maceração das folhas. Logo após, foi acrescentado 700µl do tampão de extração Brometo de Cetil-Trimetil-Amônio CTAB a 2% (NaCl a 1,4M; EDTA a 20mM; Tris-HCl (pH 8,0) a 100mM; PVP-40 a 2%), e foi feita a homogeneização com vórtex. Os tubos foram mantidos em banho-maria a 65°C por 30 min, homogeneizando por inversão a cada 10 min. Após este tempo, foram deixados na bancada até alcançar a temperatura ambiente, e então foram acrescentados 800µl de clorofórmio-álcool isoamílico-CIA (24:1; v/v) em cada tubo, e homogeneizados por inversão durante 3 min. Em seguida, centrifugou-se a 14000 RPM por 5 min.

Depois da centrifugação para a separação das partes orgânica e aquosa, foram transferidos 600µl de CIA em cada tubo, os quais foram homogeneizados durante 3 min por inversão, e centrifugados a 1400 RPM por 5 min novamente. Após esta etapa foram transferidos para um novo tubo de 1,5mL, 600µl do líquido, e adicionados 2µl de RNase (10µg mL⁻¹) e incubados a 37°C por 30min. Terminado esse tempo, foram adicionados a 800µl de álcool isopropílico a -20°C, e mantidos por 1h a -20°C. A seguir, foram centrifugados a 10000 RPM por 3min, e logo após o álcool foi descartado, e acrescentado 1 mL de etanol a 70% a -20°C, mantido por 5 min para a lavagem do precipitado de DNA, e centrifugado a 14000 RPM por 5 min. Na sequência, o etanol a 70% foi descartado, e acrescentou-se 1mL de etanol a 95% a -20°C, sendo mantido por 3 min. Após o descarte do etanol a 95%, foi mantido o precipitado em temperatura ambiente, para a secagem, por aproximadamente 30 min. Em seguida, o DNA foi diluído em 100µl de água ultra pura, e mantido por 12h em temperatura ambiente, e armazenado a -20°C até o uso.

A qualidade e quantidade de DNA foram avaliados através de eletroforese em gel de 0.8% de agarose corado com GelRed® e visualizado em transluminador sob luz UV. Um espectrofotômetro NanoVue™ (GE Healthcare) também foi utilizado para avaliar a qualidade do DNA em relação a presença de contaminantes fenólicos, polissacarídeos e proteínas, e posterior elaboração de uma solução de trabalho que foi diluída na concentração de 100 ng µL⁻¹. Para esta análise, foi usado o DNA genômico de 30 acessos de nogueira-pecã.

2.3 Análise genética e molecular

As reações AFLP (VOS et al, 1995) foram realizadas utilizando o protocolo convencional baseado na restrição do DNA total com duas endonucleases, ligação de adaptadores, amplificação via PCR e separação dos fragmentos amplificados. Para a reação de restrição/ligação cerca de 100ng de DNA genômico total foram incubados com as enzimas de restrição *PstI* e *MseI* e os adaptadores correspondentes. A restrição foi feita a 37 °C durante 8h (reação 2,5 x T4 DNA ligase buffer, 2,5 U *MseI*, 2,5 U *PstI*, 2,5 µg BSA). Seguida da ligação por mais 8h também a 37°C (reação 25 pmol de *Mse-adapter*, 25 pmol de *Pst-adapter*, 1.0 U T4 ligase de DNA). A mistura de reação foi diluída quatro vezes e utilizada como molde para a amplificação via PCR com a combinação de primers (Mse-GGA/LTR) e (Mse-GCG/LTR), em uma reação contendo 4 uL de DNA, 25 pmol de primer *Mse* 0,5 µL, 100 pmol de primer LTR 0,5µL, 1,0 U de Taq Polimerase de DNA, tampão 2x PCR, 0,25 mM de cada dNTP. Os primers para as regiões dos retrotransposons foram desenhados a partir de regiões de terminação conservadas do transposon (LTR), baseado nas sequências de DNA desses elementos, depositados no banco de dados GyDB database (LLORENS et al. 2011). Foram desenhados primers para retrotransposons de quatro diferentes elementos do cluster *Ty3/Gypsy* LTR retrotransposons: CRM, Galadriel, Reina e Del (Tabela 1).

As condições de PCR foram de um ciclo inicial de desnaturação a 95°C por 2 min, seguido de 26 ciclos a 95, 56 e 72 °C por 1 min cada e um ciclo de extensão final a 72°C por 5 min. Todas as reações de PCR foram realizadas em um termociclador BioRad C1000 Touch™. Os fragmentos de S-SAP foram separados por eletroforese em gel de agarose a 3%, sob 100volts por cerca 90 minutos, utilizando como referência um *ladder* de 100 pb (Ludwig Biotecnologia Ltda.), onde o produto da PCR foi corado com uma solução intercalante de DNA GelRed® para visualização e fotodocumentação em transiluminador de luz UV. As imagens foram analisadas pelo *software* TotalLab TL 120 1 D v. 2009 (Nonlinear Dynamics Ltda., UK). A presença/ausência de fragmentos de DNA foi codificada como presença ou ausência (1, 0) respectivamente.

Tabela 1: Primers AFLP e LTRs utilizados para a análise dos acessos de noqueira-pecã.

AFLP primers	Primer sequence (5' - 3')
MseI+AGCG	GTAGAGTCCTGAGTAAGCG
MseI+AGGA	GTAGAGTCCTGAGTAAGGA
LTR primers	
Nome	Primer sequence (5' - 3')
3'CRM	ACCCAACGGGTGCACATC
3'Galadriel	GTGCGAGAAAAATAAATCGACCCTCTTTCA
3'Reina	AACCCCAGTAGATCTACTACGTAGCTAACA
5'Del	ACACTTTGGGACTTAAAAGCGTATTTTGAT

2.4 Análise estatística

A análise dos produtos da amplificação foi feita classificando os fragmentos independentemente conforme presença (1) e ausência (0) de cada fragmento amplificado e testado em gel de agarose 3% corado com GelRed® e visualizado em transluminador sob luz UV.

A matriz binária foi utilizada para a estimativa de diversos índices de diversidade genética. A diversidade genética foi estimada assumindo ausência de endogamia, uma vez que a espécie é dioica, reproduzindo-se assim, exclusivamente por fecundação cruzada. Para marcadores cuja informação é baseada somente na presença/ausência do fragmento, a estimativa de diversidade genética (H) varia de $H = 0,0$ (todos os indivíduos possuem o mesmo “alelo”, ou seja, presença ou ausência do fragmento) a $H = 0,5$ (a mesma frequência para ambos os “alelos”).

O software GenAEx 6.5 (PEAKALL & SMOUSE, 2012) foi utilizado para o cálculo do índice de diversidade de Shannon, $I = -\sum[(p \times \ln_p) \times (q \times \ln_q)]$, onde p e q são as frequências de 1 e 0, respectivamente. Este índice varia de $I = 0,0$ (todos os indivíduos possuem o mesmo “alelo”) a $I = 1,0$ (a mesma frequência para ambos os “alelos”) e é independente do princípio de Hardy-Weinberg (POLETTTO et al. 2020).

O software Past 4.03 (HAMMER et al., 2001) foi utilizado para elaborar uma análise de componentes principais (PCA) e um dendrograma baseado no algoritmo Neighbor Joining (NJ), ambos baseados no índice de similaridade de Jaccard entre os pares de plantas. O índice

de Jaccard foi calculado como $J_{AB} = \frac{M_{11}}{M_{11} + M_{10} + M_{01}}$, onde M_{11} é o número de fragmentos presentes em A e B, M_{10} é o número de fragmentos presentes em A e ausentes em B e M_{01} é o número de fragmentos ausentes em A e presentes em B (POLETTO et al. 2020).

3. Resultados e discussões

Um total de 240 fragmentos S-SAP foram avaliados, resultando em 92 fragmentos polimórficos confiáveis e reprodutíveis. Nenhuma planta clone foi observada no total das amostras, com base nesses fragmentos analisados uma baixa diferenciação foi observada entre os indivíduos. Uma análise entre cada combinação de primer foi feita, e foi observada alta similaridade na diversidade genética (H), assim como o índice de Shannon (I) que se mostrou bem similar entre os acessos (Tabela 2).

Tabela 2. Dados das combinações de primers para diversidade genética (H) total e índice de Shannon (I).

Combinações	I^a	H^b
MGCG-CRM	0.66	0.47
MGCG-DEL	0.64	0.46
MGCG-GALADRIEL	0.62	0.44
MGCG-REINA	0.67	0.49
MGGA-CRM	0.65	0.46
MGGA-DEL	0.65	0.47
MGGA-GALADRIEL	0.64	0.45
MGGA-REINA	0.64	0.45

^a índice de Shannon total, ^b diversidade genética total

Uma média moderada da diversidade genética multilocus foi estimada para este conjunto de dados, para a diversidade genética $H = 0,46$ (S.E. = 0,05) e para o índice de Shannon $I = 0,65$ (S.E. = 0,01). O índice de similaridade de Jaccard entre pares de plantas apresentou um valor médio $J = 0,46$ (Tabela suplementar 1), o que equivale a uma distância genética média $d = 0,54$ ($d = 1 - J$). A abordagem de ordenação (Figura 1) não permite que

sejam caracterizados grupos de plantas com base na presença/ausência de fragmentos S-SAP, porém demonstra a existência de diversidade genética entre os acessos avaliados.

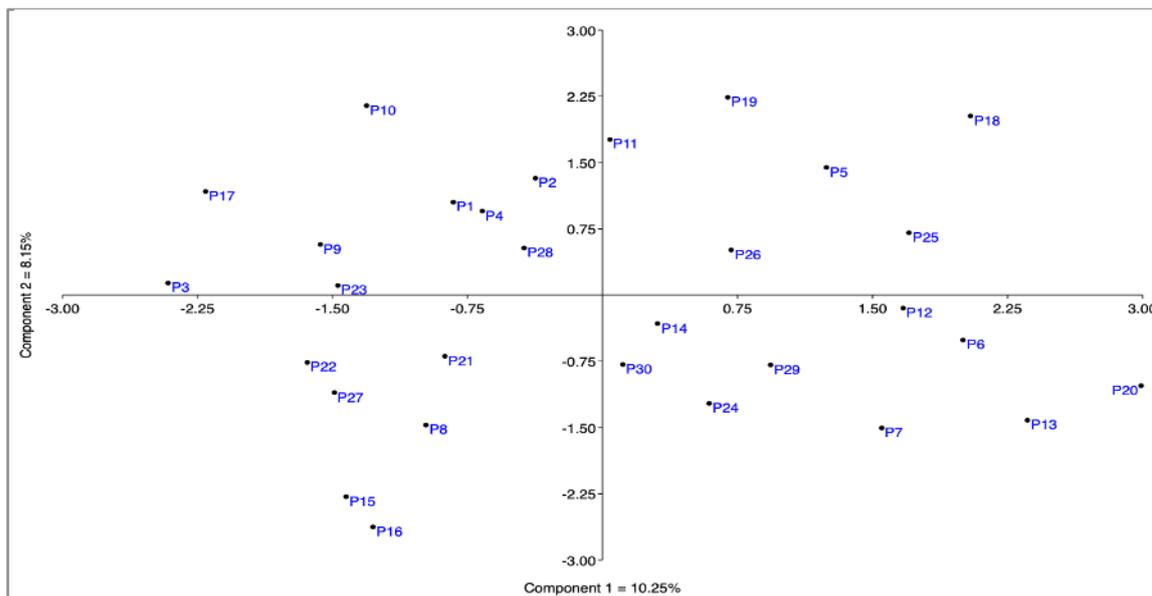


Fig. 2.1 Análise dos componentes principais da noqueira-pecã, com base nos dados genéticos moleculares

Entretanto, alguns acessos como o caso do P20 e do P3, apresentaram uma distância no eixo horizontal considerável que pode caracterizar uma variação maior entre seus padrões genéticos. O mesmo ocorre com o acesso P16 e P10. No entanto, mesmo apresentando uma dispersão na ordenação todos tendem a seguir um padrão de distanciamento, mas ainda assim, mostra a existência de diversidade genética entre os acessos. Este gráfico concentrou-se na análise dos dados genéticos polimórficos dos acessos de noqueira-pecã existentes em pomares do Rio Grande do Sul, especificamente do município de Anta Gorda, tais acessos apresentaram uma diversidade genética relativamente alta, sendo $H = 0,46$ se comparado à diversidade encontrada em estudos anteriores com outros marcadores para a espécie. As informações obtidas nesta análise podem ser utilizadas isoladamente ou combinadas com outras análises, podendo determinar padrões de diversidade, e assim posteriormente caracterizar cultivares, sendo possível fazer identificação de plantas com potencial comercial, farmacêutico ou nutricional.

Estudos de análises morfométricas, químicas e genéticas, foram feitas com noz-pecã, e os dados morfométricos apresentaram 98,88% de variação total, contrastando com os dados das análises químicas que revelaram uma homogeneidade relativa entre os acessos avaliados. Já os dados genéticos mostraram uma distribuição de grupos distintos entre os acessos, porém

não foi observado relação entre esses grupos com os dados morfológicos ou químicos, embora os marcadores AFLP podem estar ligados a genes de interesse agrônômico (Poletto et al. 2020).

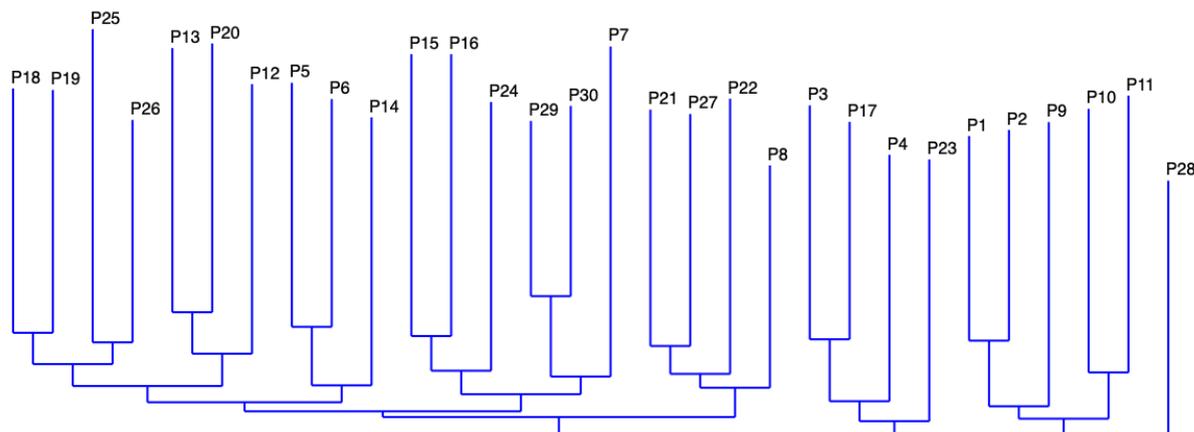


Figura. 2.2 Dendrograma de UPGMA baseada no índice de similaridade de Jaccard de 30 acessos de noqueira-pecã.

As análises genéticas deste estudo, baseadas nos marcadores S-SAP não se correlacionaram com os dados morfológicos realizados por Poletto et. al (2020), no qual o autor analisou o comprimento, diâmetro, peso da casca, peso da amêndoa e o peso total. Embora os acessos sejam provenientes do mesmo local e produtor, a morfologia apresentou-se variada, assim como a diversidade genética utilizando marcadores AFLP e também não tiveram ligação entre eles. O agrupamento do dendrograma da fig.2 nos mostra um índice de similaridade entre os pares de 0,46, onde pelo menos seis grupos se formaram o que nos indica que podem estar ligados a genes sob ação de seleção, mas não necessariamente as características morfológicas avaliadas no estudo de Poletto et. al (2020), portanto podemos dizer que se trata de marcadores neutros, que são igualmente importantes para a caracterização genética.

Sabemos que o estudo da diversidade genética é crucial para o melhoramento genético, e este resultado potencializa os avanços no estudo na seleção de pomares no Sul do Brasil como fonte de germoplasma. Essa alta diversidade genética encontrada em nosso estudo é possível que tenha se dado pelo cruzamento direto e espontâneo entre as diferentes cultivares, que resultou em novas cultivares, pois a forma de seleção das nozes durante muitos anos era baseada na escolha de indivíduos mais promissores, produtivos e resistentes/tolerantes, e isso atrapalha a garantia de uma produção confiável, pois a maioria dessas cultivares ainda não possuem registro. Além da conservação genética ser uma

necessidade para as cultivares com características superiores, a confiabilidade da produção e facilidade de seleção para o tipo de uso da pecaneira pode ser facilitado.

Os padrões de diversidade genética acessados em quatro populações naturais de noz-pecã que crescem na área indígena no México utilizando marcadores AFLP foram considerados relativamente baixos. Com 52,1% de fragmentos de DNA polimórfico (196 de 376 fragmentos), a diferenciação genética em pares de árvores variou entre 3% e 17% (CERNA-CORTÉS et al., 2003). Mesmo apresentando baixa diferenciação genética entre as árvores, e sobreposição de indivíduos de diferentes origens geográficas no dendrograma UPGMA foi associado à ocorrência de fluxo de genes através dos rios locais e por ação antropogênica (CERNA-CORTÉS et al., 2003). Além da caracterização da diversidade genética nestas populações, neste estudo seis marcadores AFLP estavam associados à dicogamia. Estes marcadores moleculares foram considerados ferramentas importantes para a previsão desta característica na fase de muda, permitindo a seleção de plantas no estabelecimento de pomares de nozes, assegurando uma polinização completa e, conseqüentemente, melhorando o rendimento.

Em outro estudo utilizando marcadores, foi relatado um padrão equivalente de diversidade genética numa análise de marcadores nucleares e plastídeos SSR, amostras de 19 populações naturais de noz-pecã em crescimento nos EUA e México (GRAUKE et al., 2011). Também foi encontrado uma diversidade genética relativamente alta ($h = 0,516$ para três loci cpSSR e $H_O = 0,461$ para 14 loci nSSR) mas um número moderado a elevado de alelos (número médio total de alelos $A = 4,33$ e número médio efetivo de alelos $A_e = 2,18$ para loci cpSSR, e $A = 10,21$ e $A_e = 4,37$ para loci nSSR).

Enquanto as estimativas da diversidade genética refletem os padrões de cruzamento e as forças evolutivas (como deriva genética, migração e seleção), o número de alelos é importante para as respostas das populações às tensões bióticas e abióticas.

Além disso, um número elevado de alelos proporciona uma maior oportunidade para encontrar combinações alélicas que são de interesse para o melhoramento genético das espécies de culturas. Esta análise de cpSSR e nSSR revelou também uma estruturação geográfica das árvores amostradas em três grupos principais, com pequenas diferenças relativamente à estrutura baseada em marcadores cpSSR e nSSR (GRAUKE et al., 2011). Tais diferenças estão provavelmente relacionadas com os diferentes modos de herança de cada marcador.

Tanto os dados genéticos obtidos neste estudo, como os dados morfométricos, químicos e genéticos de outras análises, podem ser considerados isoladamente ou em

combinações para determinar padrões para caracterização de cultivares. Temos como exemplo um estudo feito com o fruto *Bertholletia excelsa Bonpl.* (castanha do Brasil), onde foi observada significativa diferença na largura da semente, no comprimento dos frutos e em sua massa seca. Por outro lado, os marcadores moleculares não mostraram uma diferenciação genética entre as cultivares analisadas, e mesmo assim o resultado melhorou a forma de avaliar o rendimento da produção. (SUJII et al, 2013).

As análises genético-moleculares com S-SAP utilizadas nesse estudo, demonstraram uma diversidade genética considerada moderada a alta nos acessos estudados, e essa diversidade é crucial para realizar o melhoramento genético. Uma vez que as árvores utilizadas para este estudo são árvores antigas, todas com mais de 50 anos, pode-se concluir que as cultivares possuem um padrão genético desejável para fins comerciais. Embora possa ser realizados mais estudos a cerca da espécie, estes resultados podem ser utilizados para justificar a preservação dos recursos genéticos destes genótipos superiores.

REFERÊNCIAS

BAKER, J. B.; BROADFOOT, W. M. **A practical field method of site evaluation for eight important southern hardwoods.** New Orleans: USDA, 1979. 31 p.

BECKER J, Vos P, Kuiper M, Salamini F, Heun M (1995) Combined mapping of AFLP and RFLP markers in barley. *Mol Gen Genet* 249:65–73.

BENJAK, A.; FORNECK , A.; CASACUBERTA, J.M. Genome-Wide Analysis of the “Cut and-Paste” Transposons of Grapevine. **PLoS ONE** v.3, p.9, 2008.

BRANDES, A.; HESLOP-HARRISON, J.S.; KAMM, A.; S. KUBIS, R.L.; SCHMIDT, T. Comparative analysis of the chromosomal and genomic organization of Ty1-copialike retrotransposons in pteridophytes, gymnosperms and angiosperms. **Plant Molecular Biology**, v.33, p. 11-21, 1997.

BRISON, F. R. **Pecan Culture.** Austin, Texas: Capital Printing, 1974. 300 p.

CASTELO BRANCO, J. S.; VIEIRA, E.A.; MALONE, G.; MAURICIO, M.K.; MALONE, E.; BERNARDES, A.; MISTURA, C.C.; CARVALHO, F.I.F.; OLIVEIRA, A.C. IRAP and

REMAP assessments of genetic similarity in rice. **Journal of Applied Genetics**, v.48, n.2, p.107–113, 2007.

CERNA-CORTÉS, J.F., SIMPSON, J., MARTÍNEZ, O., & MARTÍNEZ- PENICHE, R.A. (2003). Measurement of genetic diversity of native pecan *Carya illinoensis* (Wangenh) K. Koch populations established in Central Mexico and correlation with dichogamous flowering using AFLP. **Food, Agriculture & Environment**, 1, 168–173.

DOYLE, J. J.; DOYLE, J.L.; HORTORIUM, L.H.B. Isolation of plant DNA from fresh tissue. *Focus*, v.12, n.1, p. 13-15, 1987.

ELARABI, N.I., & ELSODA, A.S. (2017). Molecular evaluation of genetic diversity among seven genotypes of pecan (*Carya Illinoensis*). *Journal of Agricultural Chemistry and Biotechnology*, 8, 27–33. doi:10.21608/jacb.2017.38425

EMATER/RS-ASCAR. **Cultura da noz pecã em 2017 no Rio Grande do Sul**: áreatotal, produção e número de produtores. Tabela. Porto Alegre: EMATER/RS-ASCAR, 2017.

ESTEP, M.C.; DEBARRY, J.D.; BENNETZEN, J.L. The dynamics of LTR retrotransposon accumulation across 25 million years of panicoid grass evolution. **Heredity**, v.110, p.194 - 204, 2013.

FRONZA, D.; HAMANN, J. J. **Técnicas para o cultivo da noqueira-pecã**. Santa Maria: Universidade Federal de Santa Maria: Colégio Politécnico da UFSM, 2016. 424 p.

INTERNATIONAL NUT AND DRIED FRUIT COUNCIL FOUNDATION. **Global Statistical Review 2008-2013**. Reus: INC, 2013. Disponível em: <http://www.nutfruit.org/global-statistical-review-2008-2013_85959.pdf>. Acesso em: 05 de junho. 2019.

GRAUKE, L. J. and THOMPSON, T. E. Pecan pollination. IN: Stein, L. A., Mceachern, G. R. & Nesbitt, M. L. Texas Pecan Handbook, Texas A & M AgriLife Extension Service, College Station. p. 199. (eds.). 2012.

HAMMER, Oyvind et al. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia electronica*, v. 4, n. 1, p. 9, 2001.

JIA, X.-D., WANG, T., ZHAI, M., LI, T.-R., & Guo, Z.-R. (2011). Genetic diversity and identification of Chinese-grown pecan using ISSR and SSR markers. *Molecules*, 16, 10078–10092. doi:10.3390/molecules161210078.

JUANJUAN, S., SHENGYUE, Y., SHIQUN, Y., & ZHENGJIA, W. (2013). SSR analysis of genetic diversity of the 37 new pecan cultivars. *Journal of Anhui Agricultural University*, 40, 42–46.

JIANG, N.; BAO, Z.; TEMNYKH, S.; CHENG, Z.; JIANG, J.; WING, R.A.; MCCOUCH, S.R.; WESSLER, S.R. Dasheng: A Recently Amplified Nonautonomous Long Terminal Repeat Element That Is a Major Component of Pericentromeric Regions in Rice. *Genetics*, v.161, p.1293–1305, 2002.

JIANG, F.; YANG, M.; GUO, W.; WANG, X.; KANG, L. Large-Scale Transcriptome Analysis of Retroelements in the Migratory Locust, *Locusta migratoria*. *PLoS ONE* v.7(7): p. 1-15, 2012.

Kaur, J., Bhardwaj, E., Shilpa, C., Kumar, K., Singh, D., & Sharma, D. (2016). Characterization of pecan (*Carya illinoensis*) genotypes using RAPD markers. *International Journal of Farm Sciences*, 6, 77–81.

KIM, H.; TERAOKAMI, S.; NISHITANI, C.; KURITA, K.; KANAMORI, H.; KATAYOSE; Y. SAWAMURA, Y.; SAITO, T.; YAMAMOTO, T. Development of cultivar-specific DNA markers based on retrotransposon-based insertional polymorphism in Japanese pear, *Breeding Science* v. 62, p. 53–62, 2012.

PETERSON, J. K. *Carya illinoensis* (Wangenh.) K. Koch pecan. In: BURNS, R. M.; HONKALA, B. H. (Org.). *Silvics of North America*. Hardwoods. Washington: U.S. Department of Agriculture, Forest Service, 1990. v. 2. p. 205-210.

POLETTTO, T.; MUNIZ, M. F. B.; BAGGITTO, C.; CECONI, D. E.; POLETTTO, I. Fungos associados às flores e frutos da noqueira-pecã (*Carya illinoensis*). **Revista de Ciências Ambientais**, Canoas, v. 8, n. 1, p. 5-13, 2014.

POLETTTO, T, et. al., Morphological, chemical and genetic analysis of southern Brazilian pecan (*Carya illinoensis*) accessions. **Journal Scientia Horticulturae**, 2020.

RASEIRA, A. **Acultura da noqueira-pecã (*Carya illinoensis*)**. Pelotas: Embrapa Clima Temperado, 1990. 3 p. (Embrapa Clima Temperado. Comunicado Técnico , 63).

SPARKS, D. Adaptability of pecan as a species. **HortScience**, v. 40, n. 5, p. 1175–1189, 2005.

UNITED STATES DEPARTMENT OF AGRICULTURE. October 2015 USDA Pecan Crop Estimate (Thousand Pounds – Inshell Basis). **Pecan South**, v. 48, n. 9, p. 1. 2015.

VOS P, Hogers R, Bleeker M, Reijans M, Lee T van de, Hornes M, Fritjers A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new concept for DNA fingerprinting. *Nucleic Acids Res* 23:4407–4414.

WALKER, C. **Identificação de espécies de *Cladosporium* e a reação de cultivares de noqueira-pecã**. 2016. 57 f. Tese (Doutorado em Engenharia Florestal) - Programa de Pós-Graduação em Engenharia Florestal, Concentração em Silvicultura, Universidade Federal de Santa Maria, Santa Maria, 2016.

WALKER, C.; MUNIZ, M. F. B.; MARTINS, R. R. O.; MEZZOMO, R.; ROLIM, J. M.; BLUME, E. First report of species in the *Cladosporium cladosporioides* complex causing pecan leaf spot in Brazil. **Journal of Plant Pathology**, v. 98, n. 2, p. 369-377, 2016.

WELLS, L. (Ed.). **Southeastern pecan growers' handbook**. Georgia: University of Georgia, 2013. 236 p.

WELLS, L. **Pecan**: america's native nut tree. Tuscaloosa: The University of Alabama Press, 2017. 264 p.

Tabela suplementar 1. Matriz de similaridade genética de *Jaccard* entre os pares dos 30 acessos selecionados de noqueira-pecã.
 Similaridade genética média $J = 0,46$.

0,64
0,50 0,47
0,41 0,40 0,52
0,39 0,42 0,37 0,56
0,41 0,44 0,33 0,45 0,59
0,43 0,45 0,34 0,35 0,36 0,57
0,51 0,47 0,53 0,53 0,46 0,46 0,53
0,55 0,49 0,45 0,48 0,38 0,36 0,39 0,51
0,47 0,53 0,42 0,47 0,45 0,44 0,30 0,43 0,51
0,46 0,44 0,42 0,52 0,48 0,47 0,43 0,39 0,45 0,54
0,44 0,43 0,44 0,49 0,47 0,52 0,43 0,49 0,38 0,36 0,45
0,37 0,40 0,36 0,47 0,38 0,46 0,49 0,42 0,38 0,33 0,38 0,51
0,44 0,53 0,44 0,53 0,51 0,52 0,40 0,58 0,45 0,45 0,37 0,42 0,49
0,47 0,45 0,43 0,42 0,36 0,46 0,40 0,51 0,43 0,37 0,34 0,34 0,33 0,51
0,42 0,36 0,52 0,42 0,33 0,43 0,43 0,55 0,45 0,36 0,35 0,38 0,41 0,38 0,52
0,46 0,49 0,61 0,56 0,44 0,39 0,41 0,53 0,51 0,51 0,48 0,38 0,35 0,42 0,39 0,50
0,57 0,50 0,40 0,47 0,55 0,49 0,44 0,46 0,44 0,43 0,49 0,48 0,48 0,50 0,32 0,38 0,45
0,53 0,51 0,51 0,55 0,44 0,44 0,43 0,49 0,43 0,45 0,45 0,43 0,41 0,53 0,34 0,33 0,49 0,58
0,41 0,39 0,33 0,38 0,43 0,45 0,40 0,43 0,39 0,35 0,41 0,50 0,54 0,44 0,35 0,39 0,26 0,47 0,44
0,48 0,44 0,52 0,54 0,48 0,49 0,35 0,59 0,42 0,45 0,44 0,46 0,38 0,55 0,47 0,40 0,42 0,38 0,42 0,49
0,46 0,45 0,42 0,43 0,40 0,41 0,38 0,57 0,47 0,54 0,46 0,44 0,36 0,46 0,45 0,46 0,44 0,38 0,37 0,39 0,54
0,51 0,49 0,51 0,53 0,42 0,44 0,45 0,51 0,53 0,49 0,44 0,51 0,45 0,47 0,47 0,43 0,53 0,43 0,49 0,31 0,42 0,53
0,46 0,44 0,41 0,48 0,42 0,47 0,47 0,51 0,44 0,45 0,44 0,46 0,45 0,44 0,47 0,52 0,42 0,45 0,44 0,51 0,39 0,42 0,53
0,43 0,44 0,32 0,41 0,43 0,40 0,42 0,43 0,46 0,40 0,32 0,44 0,40 0,41 0,38 0,35 0,41 0,51 0,46 0,47 0,38 0,34 0,35 0,43
0,46 0,51 0,49 0,49 0,46 0,46 0,47 0,60 0,44 0,51 0,49 0,49 0,49 0,49 0,40 0,44 0,49 0,50 0,58 0,54 0,51 0,49 0,44 0,46 0,54

0,47 0,43 0,51 0,55 0,44 0,39 0,39 0,53 0,49 0,43 0,45 0,43 0,47 0,51 0,53 0,49 0,47 0,39 0,41 0,41 0,60 0,53 0,45 0,42 0,39 0,53
0,51 0,50 0,40 0,58 0,55 0,51 0,39 0,54 0,52 0,56 0,45 0,56 0,43 0,54 0,46 0,40 0,45 0,48 0,45 0,39 0,53 0,51 0,56 0,49 0,40 0,47 0,54
0,44 0,51 0,43 0,48 0,46 0,53 0,49 0,55 0,46 0,44 0,43 0,49 0,49 0,55 0,43 0,41 0,43 0,51 0,48 0,45 0,48 0,46 0,48 0,50 0,40 0,48 0,46 0,54
0,47 0,51 0,44 0,48 0,38 0,44 0,49 0,51 0,51 0,43 0,44 0,43 0,49 0,51 0,43 0,43 0,53 0,44 0,45 0,44 0,42 0,44 0,51 0,53 0,39 0,44 0,41 0,50 0,69

CONSIDERAÇÕES FINAIS

Com base nos dados obtidos nessa dissertação, pode-se inferir que configuram avanços para o conhecimento científico da *Carya illinoensis*, uma vez que foi possível caracterizar a diversidade genética de acessos da espécie utilizando marcadores moleculares usando a técnica S-SAP. As informações obtidas até o momento nos permitem auxiliar no aprofundamento de análises da diversidade genética, pois sabemos por meio de estudos recentes que a polinização aberta promove alta segregação genética, e a propagação através de sementes pode ter introduzido centenas de variedades de noz-pecã pelo mundo. Essa diversidade na forma dos frutos, das árvores, a qualidade das nozes e suas diferentes formas de se reproduzir, ainda causa uma grande confusão quanto à nomenclatura e identificação.

Os resultados obtidos neste estudo evidenciaram alta diversidade genética ($H= 0,46$), tendo em vista que para marcadores dominantes, como é o caso do S-SAP, esta estimativa de diversidade genética é bem significativa, e este resultado pode ser utilizados para determinar o padrão genético desses acessos analisados. Tais resultados mostram a possibilidade do uso de marcador molecular aliado à técnica S-SAP, e posteriormente a caracterização de cultivares, podendo ainda identificar plantas com potencial comercial, farmacêutico ou nutricional. As perspectivas dos estudos com noz-pecã são favoráveis, embora laboratórios especializados e mão de obra ainda sejam escassos, a possibilidade de novos estudos é promissora e se fazem necessárias, principalmente no Rio Grande do Sul onde a ascendência da produção é esperada. Esse resultado corrobora com a importância do avanço em estudos genéticos para a espécie, contribuindo dessa forma para a seleção de cultivares superiores e assim gerar novos bancos de germoplasma ligados a produtividade no sul do Brasil. Além da possibilidade de garantir a sobrevivência das espécies, a conservação das populações selvagens de plantas cultivadas.