



**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO DEPARTAMENTO
DE AGRONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM MELHORAMENTO GENÉTICO DE
PLANTAS**

GISELE EVANGELISTA DOS SANTOS

**FILOGENIA MOLECULAR DE ESPÉCIES DA FAMÍLIA MYRTACEAE E ANÁLISES
in silico DA PROTEÍNA 1-AMINOCICLOPROPANO-1-CARBOXILATO SINTASE EM
REPRESENTANTES DE DICOTILEDÔNEAS**

RECIFE – PE

2020

GISELE EVANGELISTA DOS SANTOS

**FILOGENIA MOLECULAR DE ESPÉCIES DA FAMÍLIA MYRTACEAE E ANÁLISES
in silico DA PROTEÍNA 1-AMINOCICLOPROPANO-1-CARBOXILATO SINTASE EM
REPRESENTANTES DE DICOTILEDÔNEAS**

Dissertação apresentada à Universidade Federal Rural de Pernambuco, como parte das exigências do Programa de Pós-Graduação em Agronomia “Melhoramento Genético de Plantas”, para obtenção do título de Mestre.

Orientadora: Prof^a. Dr^a. Luiza Suely Semen Martins.

Coorientador: Dr. Rômulo Maciel de Moraes Filho.

RECIFE – PE

2020

Dados Internacionais de Catalogação na Publicação
Universidade Federal Rural de Pernambuco
Sistema Integrado de Bibliotecas
Gerada automaticamente, mediante os dados fornecidos pelo(a) autor(a)

S237f

Evangelista dos Santos, Gisele

FILOGENIA MOLECULAR DE ESPÉCIES DA FAMÍLIA MYRTACEAE E ANÁLISES *in silico* DA
PROTEÍNA 1-AMINOCICLOPROPANO-1-CARBOXILATO SINTASE EM REPRESENTANTES DE
DICOTILEDÔNEAS / Gisele Evangelista dos Santos. - 2020.

81 f.

Orientador: Luiza Suely Semen Martins.

Coorientador: Romulo Maciel de Moraes Filho.

Inclui referências.

Dissertação (Mestrado) - Universidade Federal Rural de Pernambuco, Programa de Pós-Graduação em
Agronomia - Melhoramento Genético de Plantas, Recife, 2020.

1. Goiabeira. 2. Bioinformática. 3. Modelagem por Homologia. 4. Filogenia. I. Martins, Luiza Suely
Semen, orient. II. Filho, Romulo Maciel de Moraes, coorient. III. Título

CDD 581.15

GISELE EVANGELISTA DOS SANTOS

**FILOGENIA MOLECULAR DE ESPÉCIES DA FAMÍLIA MYRTACEAE E ANÁLISES
in silico DA PROTEÍNA 1-AMINOCICLOPROPANO-1-CARBOXILATO SINTASE EM
REPRESENTANTES DE DICOTILEDÔNEAS**

Dissertação defendida e aprovada pela banca examinadora em:

ORIENTADORA:

Dra. Luiza Suely Semen Martins
Área de Genética / Departamento de Biologia - UFRPE

EXAMINADORES:

Dr. Rafael Trindade Maia (Membro Titular Externo)
Universidade Federal de Campina Grande/UFCG)

Dra. Nara Suzy Aguiar de Freitas (Membro Titular Externo)
Área de Genética / Departamento de Biologia - UFRPE

Dra. Gheysa Coelho Silva (Membro Titular Interno)
Área de Fitotecnia/Departamento de Agronomia - UFRPE

SUPLENTE:

Dra. Angélica Virginia Valois Montarroio (Membro Suplente Interno)
Área de Fitotecnia/Departamento de Agronomia - UFRPE

RECIFE – PE

2020

Dedico à minha família: José Evangelista dos Santos, Valter Evangelista dos Santos, Valdir Evangelista dos Santos e Maria Magdala da Silva Santos por estarem comigo sempre. Vocês são tudo o que tenho e tudo que sou.

AGRADECIMENTOS

Agradeço a Deus por iluminar o meu caminho nessa jornada, por ter estado comigo nos momentos em que estive mais sozinha, diante de tantos problemas e por ter me dado força pra chegar até aqui.

A Universidade Federal Rural de Pernambuco (UFRPE), pela oportunidade. Gratidão aos professores que tanto ensinaram, à direção e a administração do programa e dessa instituição de ensino.

Ao Laboratório de Biotecnologia da UFRPE, lugar de aprendizado e de descontração com os membros da equipe de pesquisa.

Aos meus orientadores Prof^a. Dr^a. Luiza Suely Semen Martins e Dr. Rômulo Maciel de Moraes Filho por terem me acolhido nesse desafio e me ajudado a concluir esta fase.

Aos colegas de Laboratório Allan, Horace e Fabian pelas risadas, conversas, ajudas na pesquisa e por terem sido de grande apoio e conforto em dias que eu mal tinha forças para chegar na universidade.

A Prof^a Dr^a Nara Freias por todo apoio emocional e acadêmico. A universidade precisa de mais profissionais como ela.

Aos amigos que me acompanham desde a graduação Felipe Pessoa, Débora Eveny, Tainá Lacerda, Mariana Pontes, Lorena Aguiar e Laura Miranda por estarem comigo nos momentos bons e ruins.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudos.

A Lia, *in memoriam*, minha segunda mãe que sempre se preocupou comigo mesmo quando eu estava distante.

A minha mãe, Terezinha que embora não esteja presente em corpo físico, seus ensinamentos continuam me dando discernimento para fazer minhas escolhas.

Ao meu pai, simplesmente por ser quem ele é, aos meus irmãos Valdir e Valter que me suportam desde sempre e estão comigo para tudo.

Aos meus familiares Tia Ivone e Karina por todo carinho, amor, força e ajuda. Vocês me motivaram mesmo sem ter nem ideia do que eu fazia.

As amizades construídas durante esses dois anos de mestrado: Ruana Cabral, Fernanda Lima, Roberta Paes, Agnes Guedes, Eduardo, Clint Wayne, Juliana,

Joseane, Islan, Alane. Todos vocês são importantes para mim. Grata por estarem comigo no desespero e nas alegrias.

Aos amigos de infância Priscila Roberta, Rodrigo Oliveira, Rodrigo Vinícius, Bruno Victor, Ítalo Fontes, Genivaldo Trindade, Adriana Santana e demais que foram de grande apoio quando eu não tinha mais energia para seguir, ouviam minhas lamúrias e me faziam sorrir.

A Marcelo e Amanda, por tomar chopp comigo nas sextas possíveis.

Aos amigos do IFPE Michael, Raphael e Henrique que me dão apoio desde a adolescência e me deram uma força sem igual para conclusão dessa jornada.

A todos os professores da área da genética do departamento de biologia que estiveram presentes em vários momentos dessa jornada.

Obrigada a todos que me entenderam, a todos que chegaram com palavras de conforto e apoio, a todos que se preocuparam comigo mais do que qualquer um que estava próximo a mim, pois viver não cabe no lattes.

Muito obrigada!

LISTA DE FIGURAS

CAPÍTULO II

ANÁLISE MOLECULAR *in silico* DE ESPÉCIES DA FAMÍLIA MYRTACEAE

- Figura 1.** Alinhamento de sequências completas de plastomas de 18 espécies da família Myrtaceae. Regiões em azul identificam regiões codificantes segundo anotação para *P. guajava*, obtida pelo servidor DOGMA52
- Figura 2.** Representação linear do Plastoma completo das espécies *P. guajava* e *P. aureana* obtida pelo servidor OGDRAW (Lohse et al., 2013).....53
- Figura 3.** Filogenia da Família Myrtaceae com genomas completos54
- Figura 4.** Filogenia da Família Myrtaceae baseada em quatro genes plastidiais.....55
- Figura 5.** Filogenia da ordem Myrtales baseada em quatro genes plastidiais..... 58
- Figura 6.** Filogenia e Árvore do tempo de 108 espécies da família Myrtaceae, baseada em quatro genes plastidiais completos..... 62

CAPÍTULO III

- Figura 1.** Alinhamento de sequências de ACC sintase avaliadas. As sequências foram alinhadas por ClustalW, onde os resíduos idênticos e similares são exibidos na mesma cor.76
- Figura 2.** Previsão dos efeitos funcionais das mutações em aminoácidos pelo servidor SNAP2 e Domínios funcionais da 1-aminociclopropano-1-carboxilato sintase observados pelo servidor ProDom..... 77
- Figura 3.** Árvore filogenética das espécies de dicotiledôneas, geradas com base nas sequências de ACC sintase pelo método máxima verossimilhança (MV). 79
- Figura 4.** Estrutura tridimensional predita pelo servidor Phyre2 para ACC sintase da espécie *Glycine max* e *Psidium guajava* 81
- Figura 5.** Gráfico de Ramachandran gerado pelo servidor MolProbity, a partir do modelo tridimensional construído pelo servidor Phyre2 para *Glycine max* e *Psidium guajava*.

LISTA DE TABELAS

CAPÍTULO II

ANÁLISE MOLECULAR *in silico* DE ESPÉCIES DA FAMÍLIA MYRTACEAE

Tabela 1. Número de pares de bases (PB), conteúdo de guanina e citosina (GC), genes que codificam proteínas (Protein), RNA ribossômico (rRNA), RNA transportador (tRNA), número de genes (Gene) Espaço gênico de Espécies da família Myrtaceae.
.....49

Tabela 2 - Estatística dos quatro genes utilizados neste estudo.....57

CAPÍTULO III

Tabela 1. Análise da estrutura primária e localização subcelular das seqüências de ACC sintase de espécies de Dicotiledôneas 74

SUMÁRIO

1 INTRODUÇÃO	13
2 REVISÃO DE LITERATURA	17
2.1 CARACTERÍSTICAS DA FAMÍLIA MYRTACEAE	17
2.2. IMPORTÂNCIA ECONÔMICA.....	19
2.3 MELHORAMENTO GENÉTICO DA FAMÍLIA MYRTACEAE	22
2.4 FILOGENIA MOLECULAR	23
2.5 BIOINFORMÁTICA.....	25
2.6 FILOTAXONOMIA	27
2.7 GENOMA PLASTIDIAL.....	29
2.8 SÍNTESE DE ETILENO EM PLANTAS	30
REFERÊNCIAS	33
CAPÍTULO II.....	38
Genômica Comparativa de Plastomas de Espécies da Família Myrtaceae	39
1. INTRODUÇÃO	40
2. MATERIAL E MÉTODOS	42
3. Resultados e Discussão	44
4. CONCLUSÃO.....	57
5. AGRADECIMENTOS	57
CAPÍTULO III.....	63
Caracterização <i>in silico</i> e análise filogenética de uma 1-aminociclopropano-1-carboxilato sintase em representantes de Dicotiledôneas	
INTRODUÇÃO	66
MATERIAL E MÉTODOS	67
RESULTADOS E DISCUSSÃO.....	67
AGRADECIMENTOS	77
REFERÊNCIAS	78

CAPÍTULO I

(Revisão de Literatura)

FILOGENIA MOLECULAR DE ESPÉCIES DA FAMÍLIA MYRTACEAE E ANÁLISES *in silico* DA PROTEÍNA 1-AMINOCICLOPROPANO-1-CARBOXILATO SINTASE EM REPRESENTANTES DE DICOTILEDÔNEAS

RESUMO: A família Myrtaceae é de grande importância econômica e ecológica no Brasil. Diversos representantes da família, com destaque para os eucaliptos, tem grande importância na produção de madeira para uso industrial, e frutos de espécies da tribo Myrteae como a goiaba, pitanga, araçá e jaboticaba, são utilizados para produção de alimentos sejam eles *in natura* ou industrializados. Com a evolução das ferramentas de sequenciamento de DNA, diversas formas de estabelecer correlações evolutivas entre as espécies desta família têm sido utilizadas tanto em conjunto como para confirmar correlações estabelecidas com dados morfológicos, como os marcadores moleculares. Devido à sua característica de baixa taxa mutacional, o genoma de cloroplasto possui marcadores que podem ser utilizados para inferências filogenéticas. Transcritos de proteínas importantes também podem ser ferramentas de estabelecer correlações entre espécies já que partem do mesmo princípio de baixa taxa de mutação para que não se perca sua função nos organismos. Com os objetivos de realizar uma análise comparativa entre os plastomas de espécies de Myrtaceae, realizar uma reconstrução filogenética de espécies da família Myrtaceae, e realizar análise filogenética com base em quatro genes plastidiais para ordem Myrtales, foram obtidas sequências completas dos plastomas de espécies da família Myrtaceae e dos genes isolados em formato FASTA do banco de dados GenBank. A reconstrução filogenética e alinhamento de sequências foram realizadas com uso do software MEGA7. Para reconstrução das hipóteses filogenéticas foi utilizado o método de máxima verossimilhança com bootstrap de 1000 para todas as análises. As correlações filogenéticas para família Myrtaceae resultaram em uma árvore monofilética com clados fortemente suportados, com baixos valores de bootstrap para as relações entre espécies da tribo Eucalipteae. A filogenia da ordem Myrtales resultou em parafiletismo entre algumas famílias também com clados fortemente suportados. Conclui-se que estas correlações obtidas tanto para a família Myrtaceae como para a ordem Myrtales são semelhantes às correlações realizadas com marcadores morfológicos, confirmando a confiabilidade de genes plastidiais para estudos filogenéticos. As sequências das proteínas de 1-aminociclopropane-1- carboxylate synthase foram recuperadas do banco de dados GenBank através da ferramenta BlastP. Os parâmetros físico-químicos foram avaliados através da ferramenta ProtParam. A identificação de domínios funcionais foi avaliada por meio do servidor ProDom e a estimativa de efeitos mutacionais foi feita por meio do servidor SNAP2. O alinhamento de sequência das proteínas foi realizado por meio do algoritmo ClustalW e as árvores filogenéticas produzidas com o software MEGA 7 utilizando o método de máxima verossimilhança. Para a predição de estrutura das proteínas em modelos 3D foi utilizado o servidor Phyre2. As árvores filogenéticas das espécies de dicotiledôneas construídas por meio da sequência da proteína da ACC sintase apresentam consistência com proposta de classificação apresentada na literatura. No entanto os métodos utilizados se mostraram suficientes para a separação até o nível de tribo dentro das famílias Malvaceae, Rutaceae, Euphorbiaceae, Vitaceae, Moraceae, Cannabaceae, Solanaceae, Compositae, Myrtaceae, Juglandaceae, Fagaceae, Rosaceae e Leguminosae. Os modelos 3D gerados demonstraram excelente qualidade estereoquímica no gráfico de Ramachandran e boa qualidade estrutural global e local, possibilitando a análise de suas estruturas terciárias e funções moleculares.

Palavras Chaves: Goiabeira, Bioinformática, Modelagem por Homologia, Filogenia

FILOGENIA MOLECULAR DE ESPÉCIES DA FAMÍLIA MYRTACEAE E ANÁLISES *in silico* DA PROTEÍNA 1-AMINOCICLOPROPANO-1-CARBOXILATO SINTASE EM REPRESENTANTES DE DICOTILEDÔNEAS

ABSTRACT:

The Myrtaceae family is of great economic and ecological importance in Brazil. Several representatives of the family, especially the *Eucalyptus* genus, are of great importance in the production of wood for industrial use, and fruits of species from the Myrteae tribe such as guava, pitanga, araçá and jaborcaba, are used for food production, whether *in natura* or industrialized. With the evolution of DNA sequencing tools, several methodologies to establish evolutionary correlations between the species of this family, such as molecular markers, have been used together with morphological data. Due to its characteristic of low mutation rate, the chloroplast genome has markers that can be used for phylogenetic inferences. Important protein transcripts can also be tools for establishing correlations between species since they start from the same principle of low mutation rate so that their function in organisms is not lost. In order to perform a comparative analysis between plasmas of species of Myrtaceae, perform a phylogenetic reconstruction of species of the Myrtaceae family, and perform phylogenetic analysis based on four plastidial genes for the order Myrtales, complete plastome and gene sequences of species of the Myrtaceae family were obtained in FASTA format from the GenBank database. Phylogenetic reconstruction and sequence alignment were performed using the MEGA7 software. To reconstruct the phylogenetic hypotheses, the maximum likelihood method with a bootstrap of 1000 was used for all analyzes. Phylogenetic correlations for the Myrtaceae family resulted in a monophyletic tree with strongly supported clades, with low bootstrap values for the relationships between species of the Eucalipteae tribe. The phylogeny of the Myrtales order resulted in paraphyletism among some families also with strong supported clades. It is concluded that these correlations obtained for both the Myrtaceae family and for the order Myrtales are similar to the correlations performed with morphological markers, confirming the reliability of plastidial genes for phylogenetic studies. The sequences of the 1-aminocyclopropane-1-carboxylate synthase proteins were retrieved from the GenBank database using the BlastP tool. The physical-chemical parameters were evaluated using the ProtParam tool. The identification of functional domains was evaluated using the Prodom server and the mutational effects were estimated using the SNAP2 server. The protein sequence alignment was performed using the ClustalW algorithm and the phylogenetic trees produced with the MEGA 7 software using the maximum likelihood method. To predict the structure of proteins in 3D models, the Phyre2 server was used. Phylogenetic trees of dicotyledonous species built using the ACC synthase protein sequence are consistent with the classification proposal presented in the literature. However, the methods used proved to be sufficient for separation up to the tribe level within the Malvaceae, Rutaceae, Euphorbiaceae, Vitaceae, Moraceae, Cannabaceae, Solanaceae, Compositae, Myrtaceae, Juglandaceae, Fagaceae, Rosaceae and Leguminosae families. The use of *in silico* tools proved to be feasible for building a 3D model of proteins through homology modeling. The generated 3D models demonstrated stability in terms of energy level, enabling the analysis of their tertiary structures and molecular functions.

Key words: Guava, Bioinformatics, Homology Modeling, Phylogeny

1 INTRODUÇÃO

A família Myrtaceae é considerada uma das maiores famílias da flora brasileira, sendo a maioria dos seus representantes distribuídos no bioma da Floresta Atlântica (MORAIS et al., 2014). A tribo Eucalypteae, na qual estão inclusas as espécies conhecidas como Eucalipto (Gênero *Eucalyptus*), é representada por diversas espécies, em quase totalidade de origem australiana, mas que se adaptaram bem ao clima brasileiro. Estas possuem grande potencial econômico e são utilizadas na produção de derivados da celulose e de madeira (VIA CAMPESINA, 2006; EMBRAPA, 2014).

Além dos eucaliptos, as espécies da tribo Myrteae são utilizadas principalmente para produção de frutos, com destaque para a goiabeira (*Psidium guajava*) que é utilizada para consumo *in natura* e para produção de doces, compotas, geleias e sucos; Jabuticabeira (*Plinia cauliflora*), utilizada como ornamental e consumo *in natura* de seus frutos; e pitangueira (*Eugenia uniflora*) para produção de polpa e consumo *in natura* de frutos, e também na indústria de cosméticos (MORAIS et al., 2014). As espécies *E. uniflora* e *Plinia trunciflora* (O. Berg) Kausel também são utilizadas para produção de óleos essenciais antibacterianos, antifúngicos, antioxidantes (LAGO et al., 2011). Além de seu uso na medicina tradicional principalmente para a produção de chás com atuações diversas (CRUZ e CAPLAN, 2012).

Para a realização de inferências evolutivas das espécies do grupo, durante muito tempo, marcadores morfológicos foram as únicas ferramentas disponíveis para os pesquisadores. Com a descoberta do DNA e com a evolução das metodologias de sequenciamento, foram desenvolvidos métodos mais detalhados e precisos para resolver problemas como identificação equivocada de espécies e para o entendimento de suas relações filogenéticas. Uma das formas de complementar as lacunas

deixadas pelos estudos baseados apenas em marcadores morfológicos, vem do uso de marcadores moleculares. Estes marcadores possuem algumas características consideradas vantajosas, quando comparado com os marcadores morfológicos como a não influência de fatores ambientais e também por não haver a necessidade de esperar o desenvolvimento vegetativo, floração e frutificação para identificar estas características. (ALEIXO et al., 2006; GOMPERT et al., 2006; MAIRA et al., 2013).

O DNA de cloroplasto tem sido utilizado para estudos de filogenia, pois possui alto grau de conservação tanto em nível de sequência quanto de arranjo estrutural de seus genes. Ao comparar espécies também é possível perceber que os genes plastidiais possuem tamanho similar quando comparado com espécies próximas (JUDD et al., 2002).

Bancos de dados genômicos reúnem e disponibilizam sequências que podem ser de interesse global e podem ser utilizadas para estudos filogenéticos e de evolução molecular. Para recriar estas relações com dados moleculares, podem ser usadas sequências genômicas derivadas do DNA presente na célula, seja ele nuclear, mitocondrial ou plastidial (VIEIRA, 2017; MACHADO et al, 2015).

O DNA plastidial é importante para a fotossíntese, pois está presente na organela responsável pelo processo, o cloroplasto (BENDICH, 2004); também é muito utilizado para estabelecer correlações filogenéticas em plantas por ter origem nas eubactérias primitivas, sofrer poucas mutações, existir em grande quantidade nas células, e também por seu alto grau de conservação (PALMER e STEIN, 1986; TAKEUCHI, 2007). Por possuírem estas características, as sequências de DNA plastidial se destacam nos estudos de filogenia e possuem herança uniparental, permitindo a criação de correlações a nível intraespecífico (TANGPHATSORNRUANG et al., 2010; EDWARDS; BATLEY, 2004).

Estudos evolutivos também podem ser realizados utilizando dados de sequências de proteínas. Assim como os genomas plastidiais, existem proteínas que são de extrema importância para manutenção das funções celulares e de sistemas, desta forma, tais proteínas possuem alto grau de conservação, pois alterações em sua estrutura podem resultar em perda de função, sendo assim, estas proteínas sofrem poucas modificações evolutivas (WATSON, 2005). Estas modificações sofridas pelas proteínas indicam e auxiliam a resolver questões evolutivas entre espécies. Com o avanço das ferramentas de sequenciamento e análise de proteínas e suas funções, os bancos de dados possuem diversas sequências de proteínas informativas que podem ser utilizadas para comparações, correlações filogenéticas e estruturais (KASAP e BENKRID, 2010).

Uma proteína de grande importância para espécies frutíferas é enzima carboxilase sintase, responsável pela produção do 1-aminociclopropano-1-carboxilase um composto essencial para síntese de etileno. Este hormônio é o responsável pelo amadurecimento de frutos e envelhecimento de folhas, atuando diretamente no ciclo de vida de determinadas células (RAVEN et al., 2001; MATTIUZ; DURIGAN, 2001; DURIGAN et al., 2005; DE MARTINO et al., 2006).

O etileno é um hormônio importante para a agricultura, pois ao acelerar o amadurecimento de frutos, permite um melhor controle da colheita e armazenamento pós-colheita. A forma de sua produção separa os frutos climatéricos de não-climatéricos, assim também como é utilizado na agricultura, adicionando-o a frutos quando colhidos ainda verdes para chegarem aos locais de venda maduros para o cliente (RAVEN et al., 2001; KERBAUY, 2008).

Diante do exposto, a presente pesquisa tem como objetivos realizar análise comparativa dos plastomas de espécies da família Myrtaceae; realizar uma reconstrução filogenética com base nas sequências de genoma plastidial; efetuar análise filogenética utilizando quatro genes plastidiais da ordem Myrtales; construir um modelo tridimensional da proteína 1-aminociclopropano-1-carboxilato sintase, e por fim, verificar seu uso como marcador para análise filogenética.

2 REVISÃO DE LITERATURA

2.1 CARACTERÍSTICAS DA FAMÍLIA MYRTACEAE

A família Myrtaceae abrange mais de 3500 espécies distribuídas em cerca de 140 gêneros presentes em todo planeta – com exceção da Antártida - (MORAIS et al., 2014), e é uma das maiores famílias da flora brasileira, sendo no país composta por cerca de 1000 espécies distribuídas em 26 gêneros (LORENZI; SOUZA, 2008), encontradas em sua maioria na Mata Atlântica (MORAIS et al., 2014).

A importância econômica da família se caracteriza pela exploração da madeira das espécies de eucalipto, devido ao seu crescimento acelerado, capacidade de adaptação e potencial econômico, com uso diversificado de sua madeira como para a produção de móveis, ou obtenção de celulose para a produção de papel (Via Campesina, 2006; EMBRAPA, 2014).

No geral, as espécies de eucalipto têm seu centro de origem na Austrália, lugar em que compõem grande parte da vegetação. Costumam se desenvolver em qualquer tipo de clima, mas são considerados ausentes em florestas tropicais. Mais de 700 espécies de eucalipto podem ser encontrados em território australiano, apenas duas espécies de eucalipto não podem ser encontrados na Austrália, uma dessas espécies é *Eucalyptus deglupta*, encontrado no hemisfério norte, na Nova Guiné e em partes da Indonésia (EMBRAPA, 2019; ANPS, 2019)

No grupo de espécies considerados popularmente como eucalipto, classificadas como Tribo Eucalyptae (WILSON et al., 2005), estão as espécies dos gêneros *Eucalyptus* com mais de 600 espécies; *Corymbia* com mais de 100 espécies; *Angophora* com 13 espécies; *Arillastrum* com uma espécie, *Alossyncarpia* com uma

espécie; *Eucalyptopsis* com duas espécies; e por fim *Stockwellia* com uma espécie (UDOVICIC; LADIGES, 2000).

Embora sejam caracterizados como grandes árvores, os eucaliptos são diversificados de forma que é possível classificá-los morfológicamente. As flores de eucalipto possuem grande quantidade de estames que envolvem o estigma de onde surge o ovário da flor. As espécies do gênero *Eucalyptus* e *Corymbia* possuem opérculos que cobrem os botões florais formados pela união das sépalas e pétalas unidas, gerando um opérculo externo e um interno. Quando a flor se abre, a tampa do opérculo é derramada. A tampa do opérculo não está presente no gênero *Angophora*, para este, as flores desenvolvidas possuem pequenas pétalas (ANPS, 2019).

A identificação dessas plantas na natureza se caracteriza pelos seus frutos, estes são tipo vagem, duros e lenhosos. A grande variação na forma e no tamanho da vagem ocorre de espécie para espécie, sendo útil em casos de árvores muito altas em que há dificuldade de identificar a folhagem, pois pode ser realizada a coleta da vagem que cai ao redor da planta. As sementes das vagens podem levar até 12 meses para amadurecer após o florescimento, a germinação vai depender do ambiente, ou do agente externo como animais dispersores (EMBRAPA, 2019; ANPS, 2019).

As espécies dessa família que não pertencem à tribo Eucalypteae tem a madeira destinada ao uso em produção de lenha e carvão (GRESSLER et al., 2006). Em contrapartida, as espécies da tribo Myrteae (WILSON et al, 2005) são excelentes provedoras de frutas tropicais como a goiabeira (*P. guajava*), a jabuticabeira (*P. cauliflora*) e a pitangueira (*E. uniflora*), utilizadas na alimentação de diversas formas (MORAIS et al., 2014).

Nativa da América do Sul, a espécie *P. guajava* foi e levada pelos colonizadores para as colônias africanas e indianas. A espécie *E. uniflora*, está presente em todo

território brasileiro, com sua produção concentrada nos Estados de Pernambuco e Bahia. A espécie *P. cauliflora* é nativa da Mata Atlântica brasileira, sua beleza fez com que ela fosse cultivada em outros países da América do Sul (FLORA DO BRASIL, 2020).

A tribo Myrteae dispõe, além do que já foi comentado, de espécies provedoras de óleos essenciais com propriedades antibacteriana, antifúngica, antimicrobiana, antioxidante, etc (LAGO et al., 2011). Também é utilizada na medicina tradicional, com a produção de chás com propriedades anti-reumática, anti-inflamatória, antipirética, hipoglicemiante, diurética e reguladora do aparelho digestivo (CRUZ; KAPLAN, 2004; LAGO et al., 2011).

2.2. IMPORTÂNCIA ECONÔMICA

Dados do Censo do Instituto Brasileiro de Geografia e Estatística (IBGE) de 2018, no levantamento Produção da Extração Vegetal e Silvicultura (Pevs) de 2017, verificou que o país possui 9,85 milhões de hectares de florestas plantadas, sendo estas divididas entre eucalipto com cerca de 75% e pinus com 20,6%, com o Paraná se destacando como o maior produtor. A produção nacional de eucalipto tem origem principal do Sudeste, com 53% da área plantada, seguida pelo Nordeste com 15,8%, Centro-Oeste com 13,4%, Sul com 11,5% e Norte com 6,2%. Existe dificuldade de encontrar dados de produtividade no estado de Pernambuco, entretanto tem registro de produção de eucalipto no semiárido do Estado com 166,14m³ de volume de madeira (LUNDGREN et al, 2018), em Araripina (GADELHA et al, 2018), e no Araripe (SILVA et al, 2018).

No Brasil, as espécies utilizadas para plantio são *Eucalyptus grandis*, *E. saligna*, *E. urophylla*, *E. vitaminalis*, *E. dunnii*, e híbridos de *E. grandis* com *E.*

urophylla. Na Região Sul do Brasil, como é comum atingir baixas temperaturas, chegando até a pequenas geadas, a espécie *E. benthamii* é utilizada por sua tolerância a essas condições (SANTAROSA et al, 2014).

As florestas plantadas geram empregos e renda fora das áreas urbanas, reduzindo o deslocamento de famílias (SANTAROSA et al, 2014). Segundo dados do anuário de 2013, o valor estimado é de 4,36 milhões de empregos divididos em direto, indireto, setor primário, de processamento industrial e empregos resultantes do efeito-renda.

O Eucalipto tem sido plantado em função de suas características que lhe permitem um vasto uso como obtenção de madeira serrada para produção de móveis, lâminas e compensados, chapas de fibra, produção de energia com carvão feito a partir de galhos menores das árvores, celulose e papel, para esta última utilização, as empresas no geral possuem parâmetros para produção como idade da árvore e diâmetro do tronco (SANTAROSA et al, 2014).

A goiabeira (*P. guajava*) pertencente à tribo Myrteae, é a espécie mais economicamente importante do gênero *Psidium* que possui cerca de 100 espécies (LANDRUM; KAWASAKI, 1997), *P. guajava* é a única espécie que é cultivada comercialmente, com sua área de colheita correspondente a 21.579 hectares (IBGE, 2018). Existem outras espécies do gênero que são frutíferas, como os araçazeiros, mas que também não se encontram nos registros de produção por ser comercializada em pequenas regiões; as demais espécies do gênero são importantes para compor bancos de germoplasma, constituindo de uma grande fonte de material genético (SILVA, 2010).

A goiabeira (*P. guajava*) é uma árvore de pequeno porte, podendo atingir de três a seis metros de altura; possui folhas opostas de formato elíptico-oblongo; e apresenta raiz pivotante quando propagada por semente, mas as mudas propagadas por estacas

possuem apenas raízes secundárias. Suas flores são brancas, hermafroditas, que surgem em botões isolados ou em grupos de dois ou três botões na axila das folhas; apresenta fecundação cruzada com taxa entre 25,7% a 41,3%, evidenciando a autofecundação como a forma mais comum de propagação da goiabeira (SILVA, 2010).

A jabuticabeira (*P. cauliflora*) é encontrada em diversas regiões do Brasil, e é utilizada na produção de móveis e de lenha, os frutos são comestíveis e também são utilizados para produzir doces, geléias, vinho, vinagre, entre outros produtos. Através de pesquisas, foi verificado que a jabuticaba dispõe de diversos nutrientes importantes, na polpa são encontrados zinco, manganês, potássio e magnésio, a semente e a casca são ricas em fibras, além de altos teores de manganês na casca e de compostos antioxidantes (FERNANDES, 2018).

Esta planta pode chegar até 15m de altura com a copa densa. O tronco é liso e as flores brancas se dispõem no caule. Os frutos têm cor negra ou roxo escuro e são brilhantes. Possuem polpa de cor branca e de sabor adocicado. O florescimento da jabuticabeira ocorre no geral, duas vezes por ano, variando nos intervalos de julho a agosto e novembro a dezembro, a maturação dos frutos ocorre um mês após esse período (FERNANDES, 2018).

A pitangueira (*E. uniflora*), possui um comércio que em maioria é realizado por extrativismo. Existem registros de produção de pitanga em diversos lugares do mundo, mas é estimado que o Brasil é o maior produtor, com produção focada no Nordeste, principalmente nos estados de Pernambuco e Bahia. Dentre suas características marcantes está o sabor, o aroma e alto valor nutritivo, que atrai diversos tipos de consumidores (LIRA JUNIOR, 2007).

A árvore é do tipo arbusto podendo crescer até 4m de altura com ramificações de copa com até 6m de diâmetro e de formato arredondado. Suas folhas são opostas, simples pecioladas de cor verde escura quando adultas. Em cada ramo de flor tem de 4

a 8 flores brancas monóicas, perfumadas produtoras de pouco néctar, possui diversos estames, as anteras possuem cor amarelada e grande quantidade de pólen, os estiletos são brancos. Morton (1987) descreve o fruto como baga globosa, achatada nos pólos, com 7 a 10 sulcos longitudinais, sua cor varia de verde quando jovem, até vermelho escuro quando maduros e pode frutificar diversas vezes ao ano.

2.3 MELHORAMENTO GENÉTICO DA FAMÍLIA MYRTACEAE

As espécies frutíferas da família Myrtaceae estão inclusas nas fruteiras tropicais de potencial econômico, conseqüentemente, existem instituições que atuam na pesquisa com o intuito de caracterizar, selecionar, multiplicar genomas e preservar a diversidade de espécies, incluindo-as nos bancos de germoplasmas. No Brasil, além dos banco de germoplasma das universidades envolvidas em pesquisas com espécies da família, as instituições Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) e Empresa Pernambucana de Pesquisa Agropecuária (IPA), são destaque em bancos de germoplasma (SAMPALIO et al, 2017; ALMEIDA, 2019).

Espécies do gênero *Eucalyptus* têm passado por processos de melhoramento desde os anos 40, no Instituto Agrônomo de Campinas (IAC), posteriormente foi transferido para o Instituto de Pesquisa em Assuntos Florestais (IPAF), na Escola Superior de Agricultura Luiz de Queiroz (ESALQ). As espécies de eucalipto passaram a ser avaliadas como produto promissor em plantios, seleções e estabelecimentos de espécies produtivas no Brasil (CASTRO, 2016). Atualmente, a Embrapa dispõe das maiores informações sobre o cultivo de tais plantas e de bancos de sementes (EMBRAPA, 2019).

A goiaba é uma fruta de grande importância econômica e nutricional, pois possui altos teores das vitaminas A, C e de algumas do complexo B; além de ser consumida em diversos lugares do mundo. No Brasil o melhoramento da goiaba se iniciou ainda

nos anos 50 em São Paulo na ESALQ. A seleção de plantas de goiaba têm sido realizada por outros institutos de pesquisa como Embrapa, IPA e IAC, mas ainda assim, grande parte das plantas selecionadas foi realizada no Japão (LIMA, MOREIRA, 2010)

As espécies de goiabeira plantadas no Brasil variam de acordo com as características desejadas. A variedade Paluma é a mais utilizada na agricultura no Nordeste, pois pode ser utilizada tanto na indústria como no consumo *in natura*, dependendo das características de interesse do produtor. No Sudeste também são produzidas as variedades Tailandesa, Kumagai, Pedro Sato, Cortibel 'CLG', Cortibel 'RM' e Cortibel 'RG' (FLORI, 2016).

Entre outras espécies, uma das que têm chamado atenção graças às características organolépticas de seus frutos é a *Acca sellowiana*, conhecida popularmente como feijoa ou goiabeira serrana. Assim, diversos estudos foram ou estão sendo realizados buscando entendimento de como se realizam seus cruzamentos, formação de híbridos e domesticação (MORETTO, 2014; CARDOSO, 2009)

Trabalhos também têm sido realizados com o objetivo de entender melhor como funciona o cultivo da pitangueira (*E. uniflora*), desta forma, já foram realizados estudos de comportamento com diferentes quantidades de irrigação, testes de germinação de pólen e de variabilidade genética de populações conservadas e em regeneração, validação de marcadores microssatélites, diversidade química e genética, assim como também propagação vegetativa (SILVA, 2016; SANTOS, 2018; SLAVIERO, 2009; FRANZON, 2010; FRANZON, 2007; BEZERRA, 2004)

2.4 FILOGENIA MOLECULAR

A filogenia molecular é um conjunto de métodos, teorias e técnicas, utilizados para a elucidação das correlações evolutivas entre espécies, onde o diferencial é o uso

de marcadores moleculares, que são sequências de DNA, RNA e aminoácidos que podem diferenciar indivíduos e espécies (GRAUR e LI, 2000). Estes marcadores têm vantagem sobre os marcadores morfológicos, pois possuem neutralidade fenotípica, tem poucas interações epistáticas e pleiotrópicas e podem ser identificados independentemente do estágio de desenvolvimento do organismo. Desta forma, os dados moleculares têm sido usados cada vez mais de forma a complementar e confirmar as correlações evolutivas realizadas com dados morfológicos (ALEIXO et al., 2006). Para classificação biológica e filogenia são utilizados os conceitos de homologia, quando há a descendência de um ancestral comum; similaridade que mede a semelhança ou a diferença de forma independente de sua origem; agrupamento que une os itens que são similares; e por fim, o agrupamento hierárquico que forma os grupos de filogenia (LESK, 2008).

Nos métodos de agrupamento, também chamados de relações fenéticas, não são levadas em consideração as questões históricas, diferentemente do que ocorre nos métodos cladísticos, onde eles atuam nos padrões de ancestralidade que levariam a determinada relação entre as espécies. Os métodos cladísticos que são mais utilizados são os de máxima parcimônia e máxima verossimilhança (LESK, 2008).

A metodologia de Máxima Parcimônia tem como princípio, gerar reconstruções filogenéticas com base na escolha da hipótese que implica no menor número de etapas para se atingir o resultado final. Ao se praticar esse método, somente parte da informação armazenada nas sequências pode ser recuperada, visto que nem toda variabilidade será empregada como sítios informativos (LESK, 2008; NEI; KUMAR 2000).

No método de máxima verossimilhança, além de contar os eventos mutacionais, também são sugeridas probabilidades quantitativas, ela também reconstrói a ancestralidade e atribui comprimentos aos ramos com base nos valores de probabilidade (LESK, 2008), neste método, são levados em consideração todos os

sítios (SCHNEIDER, 2007; LEMEY et al., 2009), e apenas a filogenia com maior probabilidade é apontada como correta (SCHNEIDER, 2007; LEMEY et al., 2009).

A inferência Bayesiana fundamenta-se em um método estatístico relacionado com a máxima verossimilhança. A hipótese mais adequada é a que mostra a máxima probabilidade posterior, que é estimada a partir de um modelo evolutivo previamente selecionado. A probabilidade posterior para uma hipótese é dada pela verossimilhança multiplicada pela probabilidade a priori desta hipótese. Este valor pode ser analisado como sendo a probabilidade da árvore representar o cenário mais real (HUELSENSBECK et al. 2001).

O neighbor-joining é um método de distância e processa de forma mais rápida a análise de dados, que se baseia na evolução mínima. Ele é muito utilizado quando há um grande número de sequências para análise, poucos dados e ainda assim, não se torna menos preciso que os demais métodos para resolução de questões de forma rápida (SAITOU e NEI, 1987). O método NJ produz árvores agrupando sequências vizinhas paulatinamente. Nas fases do agrupamento de sequências, ele minimiza a soma dos comprimentos de ramificação e, portanto, verifica muitas topologias. Para enormes conjuntos de dados, o NJ analisa somente uma fração minúscula do número total de topologias possíveis, sendo portanto mais rápido do que metodologias como a máxima parcimônia ou máxima verossimilhança (Tamura et al. 2004).

2.5 BIOINFORMÁTICA

A bioinformática se consolidou nos anos 80 em conjunto com os avanços de tecnologia para sequenciamento de genomas e se tornou uma ferramenta essencial para resolução dos mais diversos problemas na biologia molecular. A bioinformática é uma união das ciências computacionais, estatística e biologia, e tem por objetivo organizar, processar e analisar a grande quantidade de informações geradas pela biologia molecular. Com essa finalidade, vários recursos computacionais foram desenvolvidos visando processar a grande quantidade de informação dos bancos de

dados moleculares (EDWARDS e BATLEY, 2004).

Os bancos de dados genômicos tiveram um crescimento acelerado no volume de dados armazenados quando as ferramentas de sequenciamento se tornaram mais baratas, precisas e ágeis, resultando em sequências de nucleotídeos, aminoácidos e proteínas para estudo disponíveis para uso em pesquisas (MORAIS, 2003; BAYAT, 2002).

O *GenBank* é o maior banco de dados genômicos, que contém todas as sequências obtidas por pesquisadores de todo mundo, disponível no banco de dados de informações National Center of Biotechnological Information (NCBI). A base do GenBank é colaborativa com outros bancos de dados como o banco de dados de DNA do Japão (Bank of Japan - DDBJ), instituto de Bioinformática Europeu (EMBL) e pesquisadores das mais diversas instituições de ensino pelo mundo (BENSON et al., 2000; TATENO et al., 2002).

No NCBI é possível encontrar ferramentas complementares para ajudar na manipulação dos dados encontrados no próprio servidor como o *Basic Alignment Search Tool* (BLAST), onde alguma sequência de interesse pode ser inserida para posterior busca de regiões semelhantes em todo banco de dados do *GenBank*, fundamenta-se na busca de homologia entre os genomas (GIBAS E JAMBECK, 2001)

Com o BLASTn é possível buscar nucleotídeos em bancos de nucleotídeos; como BLASTp é feita uma busca por proteínas em bancos de dados de proteínas; com o BLASTx é possível realizar proteínas usando nucleotídeos traduzidos; com o tBLASTn é pesquisado banco de dados de nucleotídeos com consulta de proteína; e por fim tBLASTx, que busca bancos de dados de nucleotídeos traduzidos com uma consulta de nucleotídeos traduzidos (NCBI, 2019). Outra ferramenta útil que é encontrada no NCBI é o *Open Reading Frame Finder* (ORF-Finder) que traduz as sequências de nucleotídeo - seja DNA ou RNA - em seis tipos de quadros de leitura diferentes, podendo identificar os possíveis genes naquela região retirando as regiões

que não fazem parte do gene (NCBI, 2019; LESK, 2008).

As sequências podem ser obtidas pelo NCBI no formato *GenBank* ou no formato FASTA. O formato fasta é o que é mais utilizado nos programas de processamento de dados de sequência de moléculas. O *Clustal* é uma ferramenta computacional utilizada para realizar alinhamentos múltiplos de sequências, sendo o *ClustalW* a versão mais utilizada, a partir dele é possível importar alinhamentos no formato NEXUS e no formato FASTA, sendo o último o formato mais comum para uso em outros programas. A partir do alinhamento, seja de nucleotídeos ou aminoácidos, é possível fazer correlações filogenéticas e gerar árvores evolutivas (MORAIS, 2003; CHENNA et al., 2003).

Os programas que são utilizados para comparar as sequências como o BLAST (NCBI, 2020), levam em consideração os princípios da genética evolutiva e molecular, para que assim possa estabelecer uma hipótese filogenética mais próxima possível da realidade (KUMAR et al., 2004). Um desses programas é o Molecular Evolutionary Genetics Analysis (MEGA), que possui os modelos estatísticos: UPGMA (Unweighted Pair Group Method with Arithmetic Means) (Sneath e Sokal 1973), NJ (Neighbor-Joining) (Nei, 1973), Máxima Parcimônia (Fitch 1971) e Máxima Verossimilhança (Kumar et al., 2016) possibilitando a realização de inferências filogenéticas.

2.6 FILOTAXONOMIA

Dos estudos que permitem a construção de hipóteses filogenéticas como os morfológicos e paleontológicos, existe a sistemática, onde é verificada a diversidade dos organismos; ela descreve e interpreta a diversidade biológica, a qual tem como responsabilidade de descobrir os ramos da árvore evolutiva da vida, documentando as diferenças entre os ramos e descrever espécies o máximo possível, incluindo nestas análises a separação de populações que resultou em outras linhagens e modificações evolutivas que aconteceram (JUDD et al., 2009).

A sistemática tem o papel de reunir as espécies dentro de um sistema e leva em consideração suas características morfológicas, relações genéticas e afinidades. A taxonomia, muitas vezes tratada como sinônimo de sistemática atua na elaboração das leis da classificação, enquanto a sistemática classifica os seres vivos, baseando-se na morfologia (BARROSO, 2004).

Existe uma conexão entre os estudos de morfologia e evolução, onde os estudos de evolução se utilizam de dados morfológicos - seja perda, ganho ou modificações de uma característica - que foram se modificando com a evolução da vida no planeta. Filogenia é definida como as sequências de eventos obtidas que possibilitam a reconstrução de uma história evolutiva. Uma árvore filogenética é um esquema que resume as relações entre ancestrais e descendentes (JUDD et al, 2009). Em árvores filogenéticas, um táxon é frequentemente um grupo monofilético, que é caracterizado por possuir um ancestral para todos seus descendentes (SANTOS, 2008). Os grupos monofiléticos recriados em uma filogenia são chamados de clados. Para os estudos com DNA são avaliadas as sinapomorfias, que são características comuns entre espécies do mesmo grupo que foram originadas de um ancestral comum (LESK, 2008).

O termo sistemática molecular define o método de realizar análises filogenéticas com dados de genoma, seja em nucleotídeos ou aminoácidos. Esses dados facilitam as análises pois possuem um limite de possibilidades, como por exemplo a identificação das bases Adenina, Guanina, Citosina, Timina e Uracila, ao contrário de dados morfológicos como por exemplo uma folha composta que possui diferentes fenótipos a depender do ambiente. O caso é que não existe plasticidade para bases nitrogenadas. Em células vegetais podem ser encontrados genomas em três locais: no núcleo (herança biparental, genoma de maior tamanho), nas mitocôndrias (herança materna, tem entre 200 - 2.500pb) e nos plastídeos (135- 160pb) (JUDD et al, 2009).

2.7 GENOMA PLASTIDIAL

Sequências de DNA podem ser encontradas além do núcleo celular, como por exemplo o DNA plastidial, encontrado nos cloroplastos (cpDNA) e o DNA mitocondrial, encontrado nas mitocôndrias (mtDNA). Ambos têm origem nas eubactérias primitivas, sendo o cpDNA proveniente das cianobactérias (Tiimmis et al. 2004), e possui várias funções importantes, como destaque a responsabilidade pela fotossíntese (BENDICH, 2004; JUDD et al, 2009).

O cpDNA possui um tamanho que varia de 120 a 180 kb em espécies fotossintetizantes (MALIGNA, 2004) e têm características marcantes de repetições em grandes quantidades e alto grau de conservação mesmo que observado em espécies distantes (PALMER e STEIN, 1986; TAKEUCHI, 2007; JUDD et al, 2009). Mas ainda assim, podem ocorrer alterações no cpDNA de acordo com a espécie, como o reposicionamento de alguns genes (FREITAS e BERED, 2003).

O alto grau de conservação faz com que seja um meio vantajoso para realizar estudos em filogenia, em conjunto com dados da sistemática, demonstrando provas das correlações evolutivas (DHINGRA e FOLTA, 2005), além da herança uniparental, que permite diferenciação de correlações genéticas que podem ser realizadas a nível intraespecífico (TANGPHATSORNRUANG et al. 2010).

O cpDNA possui uma região de cópia única pequena e uma região de cópia única longa, chamadas de SSC e LSC respectivamente, que forma um arranjo quadripartido em conjunto com as duas regiões de repetição invertida que são chamadas de IRA e IRB. Essas regiões IRs são iguais em questões de disposição de genes, mas todos esses genes possuem duas cópias por genoma, mas no sentido invertido do quadro de leitura (GREEN, 2011).

Dentre os genes de cloroplasto utilizados para análises filogenéticas destaca-se o gene *rbcL*, codificador da enzima Rubisco (ribulose-1,5-bifosfato

carboxilase/oxigenase), a enzima mais abundante em plantas, que é considerada também a proteína mais abundante do planeta, esta atua na captação de CO₂ do ar.

O gene tem evolução lenta e pouca variação de família para família (Salamin et al, 2005) e graças às suas características marcantes como a presença do íntron trnL e os espaçadores intergênicos trnL-trnF em regiões importantes, são úteis para encontrar as variações interespecíficas e intraespecíficas em diversas espécies (TSAI et al, 2006).

O gene da maturase K (matK) dispõe de cerca de 1570 pb e possui alta taxa evolutiva, sendo recomendado para uso de construção de relações filogenéticas em níveis taxonômicos como ordem ou família (LAHAYE et al., 2008) e também em níveis interespecíficos e de gênero (WILSON et al., 2005), ou seja, possui uma ampla utilização em estudos filogenéticos e até em estudos de DNA *Barcode* de plantas (GAO et al., 2008).

Os genes plastidiais se localizam geralmente, de forma padronizada, seguindo uma ordem entre eles, entretanto, este padrão tem sido modificado com a evolução, e comumente ocorrem com uma série de inversões à região longa de cópia única que são mais frequentes nas IRs (KIM e LEE, 2004).

No NCBI, através do banco do GenBank podem ser encontrados mais de 1300 genomas de cloroplastos disponíveis para uso em pesquisas. Dentre estes, 43 pertencem a espécies da família Myrtaceae (MACHADO, 2017), onde a primeira delas a ser sequenciada foi a espécie *Eucalyptus globulus* (STEANE, 2005), e com estes genomas é possível realizar análises comparativas, e alinhamentos para posterior análises filogenéticas (LIN et al, 2010).

2.8 SÍNTESE DE ETILENO EM PLANTAS

O etileno é um hormônio vegetal que possui papel importante para o desenvolvimento das plantas. Possui a fórmula C₂H₄ e contém como precursor a metionina, atua principalmente nos processos de amadurecimento de frutos e envelhecimento de folhas. O efeito desse hormônio foi descoberto quando as

lâmpadas utilizadas nos postes queimavam gás, pois os vazamentos destes gases faziam com que as folhas das árvores caíssem, em 1901 foi descoberto que o etileno estava presente nos gases usados para iluminação. Anteriormente a este fato, a ação do etileno no amadurecimento do abacaxi foi verificada e em uma casa de vegetação (RAVEN et al., 2001)

Poluição do ar e danos que ocorrem por impacto, também leva a ter aumento na produção de etileno, como, por exemplo, em maçãs, onde no local exposto pelos danos oxida, escurece, aumenta a respiração, e conseqüentemente, a produção de etileno, desta forma, danos mecânicos podem acelerar o processo de amadurecimento de frutos, diminuindo sua durabilidade (MATTIUZ; DURIGAN, 2001; DURIGAN et al., 2005; DE MARTINO et al., 2006).

Para a produção do hormônio etileno, a Metionina reage com o ATP e forma o composto S-adenosilmetionina (SAM), durante o processo ele é quebrado formando entre os compostos o ácido 1-aminociclopropano-1-carboxílico (ACC). O ACC é convertido em Etileno + CO₂ + Amônio. Nestes processos, o que é afetado por agentes externos a ponto de estimular a produção de etileno é o ACC (RAVEN et al., 2001; KERBAUY, 2008).

Trabalhos realizados indicam que o 1-metilciclopropeno atua na redução da produção de etileno (BRACKMANN et al., 2004; DAL CIN et al., 2006; DE MARTINO et al., 2006). A presença do 1-metilciclopropeno em frutas resultou em menor produção de enzimas essenciais para a síntese do etileno como a 1-aminociclopropano carboxilase sintase e 1-aminociclopropano carboxilase oxidase, assim como também atuou nos receptores de etileno ETR1 e ERS1 (DAL CIN et al, 2006) . O 1-metilciclopropano também reduziu a ação das enzimas (WATKINS, 2006).

É um hormônio vegetal eliminado na forma de vapor e pode ser produzido em

qualquer parte de plantas superiores, a depender do tipo de tecido e da fase do ciclo de vida. Como por exemplo, o etileno atua no envelhecimento do talo de frutos causando sua queda ao amadurecer, assim como também pode fazer o mesmo com as folhas, levando-as ao envelhecimento e posterior queda. A depender da quantidade de etileno produzido para o amadurecimento é que ocorre a divisão entre climatéricos (maçã, banana) e não-climatéricos (laranja, limão), possuindo alta produção de etileno e baixa produção de etileno, respectivamente (KERBAUY, 2008).

É utilizado na agricultura sendo adicionado a frutos colhidos verdes antes de irem para as vendas. Um dos exemplos mais comuns é o do tomate, pois ele acelera o processo de amadurecimento dos frutos, garantindo que vai estar próprio para vendas ao ser entregue ao cliente. É uma técnica utilizada desde o surgimento da civilização egípcia, onde eram feitos furos no figo para estimular o amadurecimento (KERBAUY, 200; RAVEN et al, 2001).

REFERÊNCIAS

- ALEIXO, A.; BURLAMAQUI, T.; GONCALVES, E.; SCHNEIDER, P. (2006) Molecular systematics of the Ocellated Woodcreeper complex (Dendrocolaptidae) in tropical South America: Implacations for taxonomy, conservation, and historical biogeograph. **Journal of Ornithology** **147**:125-126, 2006.
- BAYAT, A. Science, medicine, and the future Bioinformatics. **The BMJ** **324**:1018-1022, 2002.
- BENDICH, A. J. Circular chloroplast chromosomes: thegrandillusion. *PlantCell* **16**:1661–1666, 2004.
- BENSON, D.A.; KARSCH-MIZRACHI, I.; LIPMAND, J.; OSTELL, J.; RAPP, B.A.E.; WHEELER. D. L. GenBank. **Nucleic Acids Research** **28**: 15-18, 2000.
- BEZERRA, João Emmanoel Fernandes et al. Comportamento da pitangueira (*Eugenia uniflora* L.) sob irrigação na região do vale do rio Moxotó, Pernambuco. **Revista Brasileira de Fruticultura, Jaboticabal**, v. 26, n. 1, p. 177-179, 2004.
- BRACKMANN, A.; SESTARI, I.; ANDRÉ, C. S.; HETTWER, R. F. G. Qualidade da maçã cv. Gala tratada com 1-metilciclopropeno. **Ciência Rural**, Santa Maria, v.34, n.5, p.1415-1420, 2004.
- CARDOSO, J. H. Cultivo e conservação da feijoa: Uma homenagem a um agricultor guardião. **Embrapa Clima Temperado-Documentos (INFOTECA-E)**, 2009.
- CHENNA, R.; SUGAWARA, H.; KOIKE, T.; LOPEZ, R.; GIBSON, T. J.; HIGGINS, D.; THOMPSON, J. D. Multiple sequence alignment with the Clustal series of programs. **Nucleic Acids Research** **31**: 3497–350, 2003.
- DAL CIN, V.; RIZZINI, F. M.; BOTTON, A.; TONUTTI, P. The ethylene biosynthetic and signal transduction pathways are differently affected by 1-MCP in apple and peach fruit. **Postharvest Biology and Technology**, Amsterdam, n.42, p.125-133, 2006.
- DE MARTINO, G.; VIZOVITIS, K.; BOTONDI, R.; BELLINCONTRO, A.; MENCARELLI, F. 1-MCP controls ripening induced by impact injury on aprictos by affecting SOD and POX activities. **Postharvest Biology and Technology**, Amsterdam, n.39, p.38-47, 2006.
- Edwards, D.; Batley, J. (2004) Plant bioinformatics: from genome to phenome. **Trends in Biotechnology** **22**: 232-237.
- FERNANDES, L. L. Alimento funcional: propriedades da jaboticaba (*Myrciaria cauliflora*). **Revista FAROL**, v. 6, n. 6, p. 49-60, 2018.
- FLORI, J. E. Principais variedades de goiaba. *Revista Campo & Negócios HF*, v. 8, n. 129, p. 72-73, fev. 2016.
- FLORY, W.S.; SMITH, G. L. (1980) The chromosomes of *Habranthus martinezii* and *robustus* and their hybrid. **Plant Life** **36**: 54-62.
- FRANZON, R. C.; GONÇALVES, D. D. S.; ANTUNES, L. E. C.; RASEIRA, B. M. C. Propagação vegetativa de genótipos de pitangueira (*Eugenia uniflora* L.) do sul do Brasil por enxertia de garfagem. **Revista Brasileira de Fruticultura**, v. 32, n. 1, p. 262-267, 2010.
- FRANZON, R. C.; RASEIRA, M. C. B.; JÚNIOR, A. W. Testes de germinação in vitro e armazenamento de pólen de pitangueira (*Eugenia uniflora* L.). **Acta Scientiarum. Agronomy**, v. 29, n. 2, p. 251-255, 2007.
- GAO, X.; ZHU, Y. P.; WU, B. C.; ZHAO, Y. M.; CHEN, J. Q.; HANG, Y. Y. (2008) Phylogeny of *Dioscorea* sect. *Stenophora* based on chloroplast matK, rbcL and trnL-F sequences. **Journal of Systematics and Evolution** **46**: 315–321.

- GIBAS C.; JAMBECK, P. (2001) **Desenvolvendo Bioinformática-Ferramentas de software para aplicações em Biologia**. Editora campus, Rio de Janeiro, 440p.
- Graur, D.; Li W. (2000) **Fundamentals of Molecular Evolution**. 2a edição, Sinauer Press. Sunderland, 481p.
- GREEN, B. R. (2011) Chloroplast genomes of photosynthetic eukaryotes. **The Plant Journal** **66**:34-44.
- JUDD, W. S.; CAMPBELL, C. S.; KELLOGG, E. A.; STEVENS, P. F. E DONOGHUE, M. J. (2002) **Plant systematics: a phylogenetic approach**. 2ª ed, Sinauer Associates, Sunderland, 576p.
- KASAP, S.; BENKRID, K. High performance phylogenetic analysis with maximum parsimony on reconfigurable hardware. **IEEE Transactions on Very Large Scale Integration (VLSI) Systems**, v. 19, n. 5, p. 796-808, 2010.
- KERBAUY, G. B. Fisiologia Vegetal. 2ª edição—Rio de Janeiro: Ed. 2008.
- KUMAR, S.; TAMURA, K.; NEI, M. (2004) MEGA3: an integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. **Briefings In Bioinformatics** **5**:150-163
- HUESELBECK J.P.; RONQUIST, F.; NIELSEN, R.E.; BOLLECK, J.P. Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology. **Science** v. 294: p. 2310-2314, 2001.
- LADIGES, P. Y.; UDOVICIC, F. Comment on a new classification of the eucalypts. **Australian Systematic Botany**, v. 13, n. 1, p. 149-152, 2000.
- LADIGES, P. Y.; UDOVICIC, F.; NELSON, G. Australian biogeographical connections and the phylogeny of large genera in the plant family Myrtaceae. **Journal of Biogeography**, v. 30, n. 7, p. 989-998, 2003.
- LEMEY, P.; SALEMI, M.; V, D. A-M. (2009) **The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing**. 2ª ed. Cambridge University Press, Cambridge, p.750.
- LESK, A.M. **Introduction to Bioinformatics**. 3ª ed. Oxford University Press, 2008, p. 432.
- LI QIN-QIN; SONG-DONG ZHOU; XING-JIN HE; YAN YU; YU-CHENG ZHANG; XIAN-QIN WEI (2010) Phylogeny and biogeography of *Allium* (Amaryllidaceae: Allieae) based on nuclear ribosomal internal transcribed spacer and chloroplast *rps16* sequences, focusing on the inclusion of species endemic to China. **Annals of Botany**. 106: 709–733.
- LIRA JÚNIOR, J. S.; BEZERRA, J. E. F.; LEDERMAN, I. E.; SILVA JUNIOR, J.F. Pitangueira. Empresa Pernambucana de Pesquisa Agropecuária-IPA. 2007.
- MACHADO, L. D. O. Análise comparativa do genoma plastidial de Myrtaceae: (Mart.) O. Berg, *Plinia cauliflora* (Mart.) Kausel, *Plinia aureana* (Mattos) Mattos e *Acca sellowiana* (O. Berg) Burret, *Eugenia uniflora* L., *Campomanesia xanthocarpa*. *Plinia* sp. 2017.
- MATTIUZ, BEN-HUR; DURIGAN, J. F. Efeito De Injúrias Mecânicas No Processo Respiratório e nos parâmetros químicos de goiabas 'paluma' e 'Pedro Sato'. **Revista Brasileira de Fruticultura**, p. 282-287, 2001.
- MORAIS, D. A. L. Análise bioinformática de genes de resistência à patógenos no genoma expresso da cana-de-açúcar. Tese de Mestrado. UFPE. Recife, 2003.

MORAIS, L. M. F.; CONCEIÇÃO, G. M.; NASCIMENTO, J. M. Família Myrtaceae: análise morfológica e distribuição geográfica de uma coleção botânica. *AgrarianAcademy*, Centro Científico Conhecer - Goiânia, v. 1, p. 317- 346, 2014.

MORETTO, Samira Peruchi; NODARI, Eunice Sueli; NODARI, Rubens Onofre. A Introdução e os Usos da Feijoa ou Goiabeira Serrana (*Acca sellowiana*): A perspectiva da história ambiental. **Fronteiras: Journal of Social, Technological and Environmental Science**, v. 3, n. 2, p. 67-79, 2014.

NCBI-National Center for Biotechnology Information (2019) Disponível em: <https://www.ncbi.nlm.nih.gov/> Acesso em: 01 nov 2019.

NEI, M.; KUMAR, S. **Molecular evolution and phylogenetics**. Oxford University Press, New York, 333p, 2000.

PALMER, J. D.; STEINS, D. B. Conservation of chloroplast genome structure among vascular plants. **Current Genetics** **10**:823-833, 1986.

RAVEN, P. H.; EVERT, R. F.; EICHHORN, S. E. *Biologia vegetal*. 6ª edição. **Guanabara Koogan SA**, 2001.

SAITOU, N.; NEI, M. (1987) O método de união de vizinhos: um novo método para reconstruir árvores filogenéticas. **Molecular Biology and Evolution** **4**: 406-25.

SANTOS, K. L.; LENZI, M.; CAPRESTANO, C. A.; DANTAS, A. C. D. M.; DUCROQUET, J. P. H. J.; NODARI, R. O., ... GUERRA, M. Evidência da atuação do sistema de auto-incompatibilidade tardia em *Acca sellowiana* (Berg) Burret.(Myrtaceae). **Rev. Bras. Frutic.**, v. 29, n. 1, 2007.

SANTOS, L. B. Estudo da diversidade química e genética de *Eugenia* spp. no Estado de Sergipe. 2018.

Schneider, H. (2007) **Métodos de Análise Filogenética: Um Guia Prático**. 3ª ed, **Holos Editora Sociedade Brasileira de Genética**, Ribeirão Preto, 200p.

SILVA, C. Validação in silico de marcadores ssr de *eugenia* uniflora e transferibilidade para sete espécies de *eucalyptus*. 2016.

SLAVIERO, L. B. Variabilidade genética de *eugenia* uniflora I. Em áreas conservadas e em regeneração. 2009.

Souza, V. C.; Lorenzi, H. *Botânica sistemática: Guia ilustrado para identificação das famílias de Fanerógamas nativas e exóticas no Brasil, baseado em APGIII*. 3ª ed. **Editora Plantarum**, São Paulo, 768 p, 2012.

STEFFENS, C. A.; ESPÍNDOLA, B. P.; DO AMARANTE, C. V. T.; SILVEIRA, J. P. G.; CHECHI, R.; BRACKMANN, A. Respiração, produção de etileno e qualidade de maçãs "Gala" em função do dano mecânico por impacto e da aplicação de 1- metilciclopropeno. **Ciência Rural**, v. 38, n. 7, p. 1864-1870, 2008.

TAKEUCHI, R.; KIMURA, S.; SAOTOME, A.; SAKAGUCHI, K. Biochemical properties of a plastidial DNA polymerase of rice. **Plant molecular biology**, v. 64, n. 5, p. 601- 611, 2007.

TAMURA, K.; NEI, M.; KUMAR, S. Perspectivas para inferir filogenias muito grandes usando o método neighbor-joining. **Proceedings da Academia Nacional de Ciências dos Estados Unidos da América**. **101**: 11030-5, 2004.

TANGPHATSORNRUANG, S.; SANGSRAKRU, D.; CHANPRASERT, J.; UTHAIPAIWANWONG, P.; YOOCHA, T.; JOMCHAI, N.; TRAGOONRUNG, S. The chloroplast genome sequence of mung bean (*Vigna radiata*) determined by high-through put pyro sequencing: structural organization and phylogenetic relationships. **DNA Res** v. 17, p. 11–22, 2010.

TSAI, L. C.; YU, Y.C.; HSIEH, H. M.; WANG, J. C.; LINACRE, A. Species identification using sequences of the trnL íntron and the trnL-trnF IGS of chloroplast genome among popular plants in Taiwan. **Forensic Science International** **164**:193-200, 2006.

UDOVICIC, F.; LADIGES, P. Y. Informative ness of nuclear and chloroplast DNA regions and the phylogeny of the eucalypts and related genera (Myrtaceae). **Kew Bulletin**, p. 633-645, 2000.

WATKINS, C.B. The use of 1-methylcyclopropene (1-MCP) on fruits and vegetables. **Biotechnology Advances**, Kidlington, n.24, p.389-409, 2006.

WATSON, J. D.; LASKOWSKI, R. A.; THORNTON, J. M. Predicting protein function from sequence and structural data. **Current opinion in structural biology**, v. 15, n. 3, p. 275-284, 2005.

WENDLING, I. W.; DUTRA, L. F. **Produção de mudas de eucalipto**. Colombo: Embrapa Florestas, 2010.

WILSON, P. G., O'BRIEN, M. M., HESLEWOOD, M. M., & QUINN, C. J. Relationships within Myrtaceae sensu lato based on a matK phylogeny. **Plant Systematics and Evolution**, v. 251, n. 1, p. 3-19, 2005.

CRUZ, A. V. M.; KAPLAN, M. A. C. Uso medicinal de espécies das famílias Myrtaceae e Melastomataceae no Brasil. **Floresta e ambiente**, v. 11, n. 1, p. 47-52, 2012.

MACHADO, L.; CARVALHO, E. L.; RANGEL, D. L.; DA SILVA, L. F. D.; PINTO, P. M.; DA LUZ WALLAU, G. Caracterização do genoma mitocondrial e plastidial da alga antártica prasiola crispa. **Anais do Salão Internacional de Ensino, Pesquisa e Extensão**, v. 7, n. 2, 2015.

VIEIRA, L. N. Estrutura e evolução do genoma plastidial em Araucariaceae Henkel & W. Hochst. e Podocarpaceae Endl. 2017.

Psidium in Flora do Brasil 2020 em construção. Jardim Botânico do Rio de Janeiro. Disponível em: <<http://floradobrasil.jbrj.gov.br/reflora/floradobrasil/FB24034>>. Acesso em: 04 ago. 2020

Eugenia in Flora do Brasil 2020 em construção. Jardim Botânico do Rio de Janeiro. Disponível em: <<http://floradobrasil.jbrj.gov.br/reflora/floradobrasil/FB10560>>. Acesso em: 04 ago. 2020

Plinia in Flora do Brasil 2020 em construção. Jardim Botânico do Rio de Janeiro. Disponível em: <<http://floradobrasil.jbrj.gov.br/reflora/floradobrasil/FB10828>>. Acesso em: 04 ago. 2020

MOREIRA, F. R. B.; LIMA, M. F. A cultura da goiaba. Área de Informação da Sede- Col Criar Plantar ABC 500P/500R Saber (INFOTECA-E), 2010.

SANTAROSA, E.; PENTEADO JR, J. F.; GOULART, I. C. G. R. Cultivo de eucalipto em propriedades rurais: diversificação da produção e renda. Colombo: **Embrapa Florestas**, 2014.

SAMPAIO, A. A.; DA SILVA FONTES, L.; BARBOSA, D. R. S.; DA COSTA LIMA, V. M. G. Bioatividade dos extratos aquosos de folhas de pimenta malagueta e eucalipto sobre *Zabrotes subfasciatus* em grãos de fava. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, v. 12, n. 2, p. 198-203, 2017.

ALMEIDA, R. S. Adaptação e seleção de matrizes em um teste de espécies e procedências de eucalipto. 2019.

LIMA GADELHA, F. H.; SILVA, J. A. A.; FERREIRA, R. L. C.. Viabilidade econômica do cultivo de clones de *Eucalyptus* em função do espaçamento e do sistema silvicultural, em Araripina-PE. **Anais da Academia Pernambucana de Ciência Agronômica**, v. 15, n. 2, p. 105-123, 2018.

SILVA, J. A. A. ; ROCHA, K. D.; FERREIRA, R. L. C.; TAVARES, J. A. Modelagem do crescimento volumétrico de clones de eucalipto (*Eucalyptus urophylla*) no Polo Gesseiro do Araripe-PE. **Anais da Academia Pernambucana de Ciência Agronômica**, v. 13, p. 173-190, 2018.

LUNDGREN, W. J. C.; SILVA, J. A. A.; FERREIRA, R. L. C. Estimação do Volume de Eucaliptos por Krigagem e Cokrigagem no Semiárido Pernambucano. **Floresta e Ambiente**, v. 24, p. 0-0, 2018.

CAPÍTULO II

Este artigo será submetido para revista Scientia Plena

Genômica Comparativa de Plastomas de Espécies da Família Myrtaceae

Comparative Genomics of Plastomas of Species of the Myrtaceae Family

G. E. dos Santos¹; R. M. Moraes Filho¹; A. D. F. da Silva¹; L. S. S. Martins¹

¹Departamento de Agronomia, Universidade Federal Rural de Pernambuco, 52171-900, Recife-PE, Brasil

*gisaesantos@gmail.com

(Recebido em dia de mes de ano; aceito em dia de mes de ano)

A família Myrtaceae é de grande importância econômica, sendo utilizada de diversas formas no Brasil, desde a produção de madeira para móveis e indústria de forma geral, até produção de alimentos como doces, polpas e consumo in natura. Desta forma, sequências de plastomas de espécies da família Myrtaceae e da ordem Myrtales foram obtidas com o objetivo de avaliar genomas de bancos de dados públicos; realizar reconstrução filogenética com os genes plastidiais matK, rbcL, ndhF e atpB para Myrtaceae; e realizar reconstrução filogenética com representantes da ordem Myrtales também com genes atpB, rbcL, ndhF e matK. Sequências de 51 espécies da família Myrtaceae foram obtidas em formato fasta. Os alinhamentos foram feitos no software MEGA7, mesmo programa no qual foi realizada a reconstrução filogenética com o método de máxima verossimilhança com bootstrap de 1000, utilizando quatro genes. O mesmo método foi aplicado para as 108 espécies da ordem Myrtales. Foi verificada uma inversão de sequência de genes da espécie *Plinia aureana*. As correlações filogenéticas para família Myrtaceae resultaram em uma árvore monofilética com clados fortemente suportados, com baixos valores de bootstrap para as relações entre espécies da tribo Eucalypteae. A filogenia da ordem Myrtales resultou em parafiletismo entre algumas famílias também com clados fortemente suportados. Conclui-se que estas correlações obtidas tanto para a família Myrtaceae como para a ordem Myrtales são semelhantes as correlações realizadas com marcadores morfológicos, confirmando a confiabilidade de genes plastidiais para estudos filogenéticos.

Palavras-chave: Myrtales, Mirteae, Eucalypteae, Bioinformática

The Myrtaceae family is of great economic importance, being used in several ways in Brazil, from the production of wood for furniture and industry in general, to the production of foods such as sweets, pulps and fresh consumption. In this way, sequences of plastomas of species of the family Myrtaceae and of the order Myrtales were obtained with the objective of evaluating genomes of public databases; perform phylogenetic reconstruction with the plastidial genes matK, rbcL, ndhF and atpB for Myrtaceae; and perform phylogenetic reconstruction with representatives of the Myrtales order also with atpB, rbcL, ndhF and matK genes. Sequences of 51 species of the Myrtaceae family were obtained in fasta format. The alignments were made using the MEGA7 software, the same program in which phylogenetic reconstruction was performed using the maximum likelihood method with a 1000 bootstrap, using four genes. The same method was applied to the 108 species of the order Myrtales. An inversion of the sequence of genes of the species *Plinia aureana* was verified. Phylogenetic correlations for the Myrtaceae family resulted in a monophyletic tree with strongly supported clades, with low bootstrap values for the relationships between species of the Eucalypteae tribe. The phylogeny of the Myrtales order resulted in paraphyletism among some families also with strong supported clades. It is concluded that these correlations obtained for both the Myrtaceae family and for the order Myrtales are similar to the correlations performed with morphological markers, confirming the reliability of plastidial genes for phylogenetic studies.

Keywords: Myrtales, Mirteae, Eucalypteae, Bioinformatics

1. INTRODUÇÃO

A família Myrtaceae pertencente à ordem Myrtales, é composta por mais de 3500 espécies distribuídas em cerca de 140 gêneros por todo planeta, com exceção da Antártida (Morais et al., 2014), e grande parte da riqueza desta família estão inclusos na tribo Myrteae. No Brasil é constituída por cerca de 1000 espécies distribuídas em 26 gêneros (Lorenzi; Souza, 2008), encontradas em sua maioria na Mata Atlântica (Morais et al., 2014).

Devido ao fato de não produzirem madeiras valiosas, exceto pelas espécies do gênero *Eucalyptus*, as plantas dessa família são bastante utilizadas para produção de lenha (Gressler et al., 2006), por outro lado, são excelentes provedoras de frutas tropicais como a goiabeira (*Psidium guajava* L.) – que pode ser consumida *in natura* ou industrializada nos mais diversos tipos de produto -, a jabuticabeira (*Plinia cauliflora*) e a pitangueira (*Eugenia uniflora* L.), que também são utilizadas na alimentação do ser humano e de pequenos animais herbívoros (Morais et al., 2014).

Em plantas, encontram-se três tipos de DNA: nuclear, mitocondrial e plastidial. O DNA plastidial (cpDNA) é encontrado nos plastídios, organelas essenciais para que o metabolismo das plantas, pois são responsáveis pela fotossíntese e outras funções (Bendich, 2004). O cpDNA geralmente é constituído por duas repetições invertidas idênticas que separam uma grande cópia única e uma pequena região de cópia única (Chumley et al., 2006). O cpDNA geralmente apresenta entre 108 e 218 kb e se configura por apresentar um alto grau de conservação em questões de tamanho, estrutura, ordem dos genes e conteúdo genético (Provan et al., 2001; Chumley et al., 2006; Shaw et al., 2007). Suas características de natureza não-recombinante, herança uniparental e abundância nas células fazem com que sejam muito utilizados em estudos de filogenia e sistemática e relações evolutivas (Provan et al., 2001; Nishikawa et al., 2002; Provan, 2008). Desta forma, podem ser realizadas diferenciações de linhagens e correlações genéticas a nível intraespecífico graças à herança uniparental (Tangphatsornruang et al. 2010).

A utilização de sequências completas de plastomas podem fornecer informações importantes relacionadas principalmente a historia evolutiva assim como servir de base para o desenvolvimento de marcadores genéticos moleculares que além de favorecer estudos ecológicos, são uteis na identificação das espécies (Hand et al., 2013). Diversos estudos dessa natureza vêm sendo realizados em genômica comparativa (Wilson et al, 2005; Thornhill et al, 2015; Gu et al, 2018) e análises filogenéticas (Machado, 2017; Thornhill, 2015; Gu, 2018).

Análises filogenéticas e de taxonomia de gêneros da família Myrtaceae, ainda possuem dificuldades, para serem classificadas corretamente, pois possui diversidade de gêneros e de espécies de importância econômica como *Eucalyptus*, *Psidium*, *Mirtus* e *Eugenia*. Nos trabalhos realizados, a filogenia e taxonomia desses gêneros são feitas de forma isolada, dentro de cada gênero, sendo difícil encontrar e estabelecer uma relação entre estes gêneros resultando em uma análise completa das relações filogenéticas (Wilson et al, 2005; Thornhill et al, 2015; Machado et al, 2017; Gu et al, 2018).

Espécies da família myrtaceae tem como características possuir folhas simples com glândulas, flores polistêmones e ovario ínfero (Conti et al. 1997).

A família Myrtaceae já foi dividida em duas subfamílias, Mirtoideae e Leptospermoideae (Mcvaugh, 1968). Esta divisão foi contestada e foi criada uma nova classificação proposta por Wilson et al (2005), onde a família foi dividida em duas subfamílias: Myrtoideae, com duas tribos e Psiloxylloideae com 15 tribos.

Com o avanço das tecnologias de sequenciamento e ferramentas computacionais, as avaliações filogenéticas passaram a ser realizadas com marcadores genéticos, como por exemplo, os genes plastidiais. As avaliações com marcadores morfológicos têm se mostrado insuficientes para resolver as questões da família, levando à realização das análises morfológicas e filogenéticas em conjunto para se ter uma quantidade de informações razoáveis para realizar a análise (Santos, 2014; Faria Júnior, 2014).

Filogenias que utilizam como ferramenta genes plastidiais têm se mostrando de grande importância revelando informações para melhor entendimento da origem e evolução da família Myrtaceae. Com os avanços das tecnologias de bioinformática, análises filogenéticas passaram a ser realizadas com metodologias *in silico*. O uso do gene *matK* é feito em diversas análises filogenéticas de plantas das mais diversas famílias, sendo estes estudos realizados com o gene isolado ou com uma seleção de genes plastidiais (Wilson et al, 2005; Machado, 2012; Cruz-Maia, 2018)

No banco de dados *GenBank*, no NCBI podem ser encontrados cerca de 1300 genomas de cloroplasto já sequenciados. Dentre estes, 51 pertencem à família Myrtaceae (Machado, 2017) e com estes genomas, é possível realizar análises comparativas *in silico* relacionadas à filogenia e evolução (Lin et al., 2010). Neste mesmo banco de dados, podem ser encontradas 81 espécies de Myrtaceae com o gene *matK* já sequenciado (Wilson et al, 2005)

Diante isto, sequências de plastomas de 51 espécies da família Myrtaceae, disponíveis em bancos de dados públicos, foram avaliadas com base em ferramentas da bioinformática com os objetivos: 1) efetuar uma análise comparativa entre os plastomas de espécies de Myrtaceae, com base no levantamento de sequências completas de genoma plastidial no banco de dados do GenBank; 2) realizar uma reconstrução filogenética de espécies da família Myrtaceae, com base em sequências de genoma plastidial para compreender as relações entre as espécies do grupo; 3) realizar análise filogenética com base em quatro genes plastidiais (matK, rbcL, ndhF e atpB) para espécies da família Myrtaceae e compreender suas relações, e (4) realizar análise filogenética com base em quatro genes plastidiais (matK, rbcL, ndhF e atpB) para espécies representantes da ordem Myrtales e entender suas relações.

2. MATERIAL E MÉTODOS

21 Obtenção e caracterização de sequências de genomas plastidiais de espécies da família Myrtaceae

As sequências de genoma plastidial completo foram obtidos do banco de dados NCBI (National Center of Biotechnology Information) realizando *download* em formato FASTA, totalizando 51 espécies pertencente à família Myrtaceae e uma espécie *Punica granatum* pertencente a família Lythraceae como grupo externo (*outgroup*).

Destes genomas completos, 30 foram utilizados para alinhamento com o servidor mVista. As demais foram utilizadas na avaliação filogenética com quatro genes selecionados.

22 Obtenção de sequências de genomas plastidiais de espécies da Ordem Myrtales

Foram encontrados 108 genomas completos de plastomas da ordem Myrtales no banco de dados NCBI e posteriormente foi feito o download em formato FASTA dos genes ndhF, matK, atpB e rbcL. Foram selecionadas também duas espécies da ordem Geraniales para utilização como *outgroup*.

23 Alinhamento e Reconstrução filogenética

O alinhamento dos genomas plastidiais completos foi realizado no servidor mVISTA (Brudno et al., 2003) utilizando o algoritmo mLAGAN. Os parâmetros padrões foram aplicados, e a estrutura de anotação do genoma do cloroplasto de *Psidium guajava* foi utilizado como referência. O gráfico VISTA (Frazer et al., 2004) permitiu a verificação percentual de identidade de cada plastoma em relação à *P. guajava*.

Para análise filogenética com genoma completo de espécies da família Myrtaceae foram recuperados e alinhados genomas de 16 espécies, sendo destes 15 da família Myrtaceae e 1 da família Lythraceae (*P. granatum*) para ser utilizada como grupo externo. A árvore filogenética foi criada com o software MEGA7 com o método de máxima verossimilhança e com suporte estatístico calculado utilizando 1000 replicatas.

O alinhamento realizado foi importado para o software MEGA7 (Molecular Evolutionary Genetics Analysis), versão 7.0 (Kumar et al., 2016) para a realização da análise filogenética, empregando o método de Máxima verossimilhança. O suporte estatístico foi calculado pelo método de bootstrapping com 1000 replicatas (Felsenstein, 1985). Os bootstrap de 90-100 foram considerados como fortemente suportados, os de 80-89 moderadamente suportados e 50-79 fracamente suportados. Assim, a análise filogenética foi estimada pela máxima verossimilhança com 31 espécies.

A escala evolutiva do tempo entre as espécies da família Myrtaceae foi verificada através do método Reltime (Tamura et al., 2012) e o modelo de Tamura- Nei (Tamura et al., 1993). A árvore foi analisada utilizando três pontos de calibração oriundos pelas escalas do site *Timetree* (<http://www.timetree.org/>).

A análise filogenética da família Myrtaceae, foram utilizados genes de cloroplasto de 52 espécies, sendo 51 espécies da família Myrtaceae e uma da família Lythraceae. O alinhamento das sequências foi realizado com o uso do algoritmo *ClustalW* e a árvore filogenética foi criada com o software MEGA7 com o método de máxima verossimilhança e com suporte estatístico calculado utilizando 1000 replicatas.

Para análise filogenética da ordem Myrtales, foram utilizados genes de cloroplasto de 108 espécies divididas entre as famílias Myrtaceae, Onagraceae, Lythraceae, Melastomataceae, Combretaceae e Vochysiaceae. O alinhamento das sequências foi realizado com o uso do algoritmo *ClustalW* e a árvore filogenética foi criada com o software MEGA7 com o método de máxima verossimilhança e com suporte estatístico calculado utilizando 1000 replicatas.

3. Resultados e Discussão

Cinquenta e duas espécies da família Myrtaceae das quais o genoma plastidial foi recuperado do banco de dados do NCBI, são apresentadas na Tabela 1. Quanto ao conteúdo GC (guanina e citosina), houve a variação de 36,8% a 39,3%; nos genes que codificam proteínas foram observados os valores com variação de 84 e 85; o espaço gênico variou de 91.648 pb a 96.246 pb; o número de pares de bases variou de 158.445pb (*Eugenia uniflora*) a 161.071pb (*Eucalyptus spathulata*); e a quantidade de genes variou de 118 (*Pisidium guajava*) a 138 (*Corymbia torelliana*).

Tabela 1. Número de pares de bases (PB), conteúdo de guanina e citosina (GC), genes que codificam proteínas (Proteínas), RNA ribossômico (rRNA), RNA transportador (tRNA), número de genes (Gene) Espaço gênico de Espécies da família Myrtaceae.

Espécie	Identificação	Tamanho(MB)	GC%	Proteínas	rRNA	tRNA	Nº de Genes	Espaço Gênico
<i>Acca sellowiana</i>	KX289887.1	159370	37	84	8	34	136	94155
<i>Allosyncarpia ternata</i>	NC_022413.1	159593	36,8	85	8	37	135	94725
<i>Angophora costata</i>	NC_022412.1	160326	36,8	85	8	37	135	94612
<i>Angophora floribunda</i>	NC_022411.1	160245	36,8	85	8	37	135	94596
<i>Corymbia citriodora</i>	KM885985.1	160146	36,8	85	8	37	135	94639
<i>Corymbia eximia</i>	NC_022409.1	160012	36,8	85	8	37	135	94675
<i>Corymbia gummifera</i>	NC_022407.1	160713	36,8	85	8	37	135	96246
<i>Corymbia henryi</i>	NC_028409.1	160095	36,8	85	8	36	129	94684
<i>Corymbia maculata</i>	NC_022408.1	160045	36,8	85	8	37	135	94675
<i>Corymbia tessellaris</i>	NC_022410.1	160127	36,8	85	8	37	135	94662
<i>Corymbia torelliana</i>	NC_028410.1	159994	36,8	84	8	36	138	94654
<i>Eucalyptus aromaphloia</i>	NC_022396.1	160149	36,9	85	8	37	135	94635
<i>Eucalyptus baxteri</i>	NC_022382.1	160032	36,8	85	8	37	135	94761
<i>Eucalyptus camaldulensis</i>	KC180791.1	160164	36,9	85	8	37	135	94735
<i>Eucalyptus cladocalyx</i>	NC_022394.1	160213	36,9	85	8	37	135	94668
<i>Eucalyptus cloeziana</i>	NC_022388.1	160015	36,8	85	8	37	135	94602
<i>Eucalyptus curtisii</i>	NC_022391.1	160038	37	85	8	37	135	94717
<i>Eucalyptus deglupta</i>	NC_022399.1	160177	36,9	85	8	37	135	94729
<i>Eucalyptus delegatensis</i>	NC_022380.1	159724	36,9	85	8	37	135	94731
<i>Eucalyptus diversicolor</i>	NC_022402.1	160214	36,8	85	8	37	135	94636
<i>Eucalyptus diversifolia</i>	NC_022383.1	159954	36,8	85	8	37	135	94686
<i>Eucalyptus elata</i>	NC_022385.1	159899	36,8	85	8	37	135	94713
<i>Eucalyptus erythrocorys</i>	NC_022406.1	159742	36,9	85	8	37	135	94726
<i>Eucalyptus globulus</i>	NC_008115.1	160286	36,9	85	8	37	135	94108
<i>Eucalyptus grandis</i>	HM347959.1	160137	36,9	85	8	37	135	91648
<i>Eucalyptus guilfoylei</i>	NC_022405.1	160520	36,9	85	8	37	135	94942
<i>Eucalyptus marginata</i>	NC_022390.1	160076	36,8	85	8	37	135	94714
<i>Eucalyptus melliodora</i>	NC_022392.1	160386	36,8	85	8	37	135	94591
<i>Eucalyptus microcorys</i>	NC_022404.1	160225	36,8	85	8	37	135	94666
<i>Eucalyptus nitens</i>	NC_022395.1	160271	36,9	85	8	37	135	94719
<i>Eucalyptus obliqua</i>	NC_022378.1	159527	36,9	85	8	37	135	94656

<i>Eucalyptus patens</i>	NC_022389.1	160187	36,8	85	8	37	135	94717
<i>Eucalyptus pauciflora</i>	MG921592.1	159942	36,8	85	8	37	135	94756
<i>Eucalyptus polybractea</i>	NC_022393.1	160268	36,9	85	8	37	135	94624
<i>Eucalyptus radiata</i>	NC_022379.1	159529	36,9	85	8	37	135	94731
<i>Eucalyptus regnans</i>	NC_022386.1	160031	36,8	85	8	37	135	94686
<i>Eucalyptus saligna</i>	NC_022397.1	160015	36,8	85	8	37	135	94411
<i>Eucalyptus salmonophloia</i>	NC_022403.1	160413	36,8	85	8	37	135	94696
<i>Eucalyptus sieberi</i>	NC_022384.1	159985	36,8	85	8	37	135	94737
<i>Eucalyptus spathulata</i>	NC_022400.1	161071	36,8	85	8	37	135	96016
<i>Eucalyptus torquata</i>	NC_022401.1	160223	36,9	85	8	37	135	94543
<i>Eucalyptus umbra</i>	NC_022387.1	159576	36,8	85	8	37	135	94365
<i>Eucalyptus verrucata</i>	NC_022381.1	160109	36,8	85	8	37	135	94680
<i>Eugenia uniflora</i>	NC_027744.1	158445	37	85	8	37	133	94220
<i>Pimenta díóica</i>	NC_034684.1	158984	37	85	8	37	134	94501
<i>Plinia aureana</i>	NC_039557.1	158918	37	84	4	30	119	93723
<i>Plinia cauliflora</i>	NC_039395.1	159095	37	84	4	30	119	94227
<i>Plinia trunciflora</i>	NC_034801.1	159512	37	84	8	36	132	94428
<i>Psidium guajava</i>	NC_033355	158841	39,3	84	4	30	118	94129
<i>Punica granatum</i>	NC_035240.1	158663	37	85	4	30	119	94679
<i>Stockwellia quadrifida</i>	NC_022414.1	159561	36,9	85	8	37	135	94351
<i>Syzygium cumini</i>	GQ870669.3	160373	37	85	8	37	135	94132

O tamanho do genoma variou de 158.445 kb (*Eugenia uniflora*) a 161.071 kb (*Eucalyptus spathulata*), também encontrados por Eguiluz et al. (2017) para *E. uniflora*, e dentro da faixa de 107 kb (*Cathaya argyrophylla*) a 218 kb (*Pelargorium*) visto por Machado et al. (2016) em seus trabalhos para as mesmas espécies. Valores no mesmo intervalo comparando genoma de *Heimia myrtifolia* com outras espécies de Myrtaceae mostraram o resultado no mesmo intervalo com 159.219 kb (Gu et al, 2018). Alves (2005) encontrou o valor de 160kb para *E. grandis*.

O conteúdo GC% encontrado de 37 % para *E. uniflora* é o mesmo encontrado por Machado (2017) e a média geral de 36,9 é semelhante ao encontrado por Alves (2005) e Gu et al. (2018). O número de genes variou entre 118 (*P. guajava*) e 138 genes (*C. torelliana*), valores semelhantes ao encontrado por Machado (2017) em seu trabalho com Myrtaceae. Esses genes são divididos em codificadores de proteína, rRNA e tRNA. Eles encontraram valores de quantidade de gene em angiospermas de 120 a 130, que atuam principalmente na fotossíntese, transcrição e tradução (Bendich, 2004).

3.1 Comparação genômica das sequências do plastoma de espécies da família Myrtaceae

Como é possível ver na Figura 1, ao verificar o gráfico obtido pelo mVista comparando o genoma plastidial das espécies observou-se um alto nível de conservação. É possível constatar que as regiões em azul são as regiões codificantes de gene e as regiões rosa indicam as áreas não codificantes, ao comparar com a anotação da espécie *P. guajava*, obtida através do servidor DOGMA.

Embora observado o alto grau de conservação, destaca-se a espécie *P. aureana* com uma parte de seu genoma comportando-se de forma diferente a das outras espécies com uma região sem notação. Ao realizar o comparativo entre espécies do mesmo gênero (*P. trunciflora* e *P. cauliflora*), utilizando a mesma notação de *P.*

guajava obtida do servidor DOGMA, observa-se que a espécie *P. aureana* apresentou diferença na ordenação dos genes, provavelmente originário de um evento de inversão de uma região gênica do plastoma.

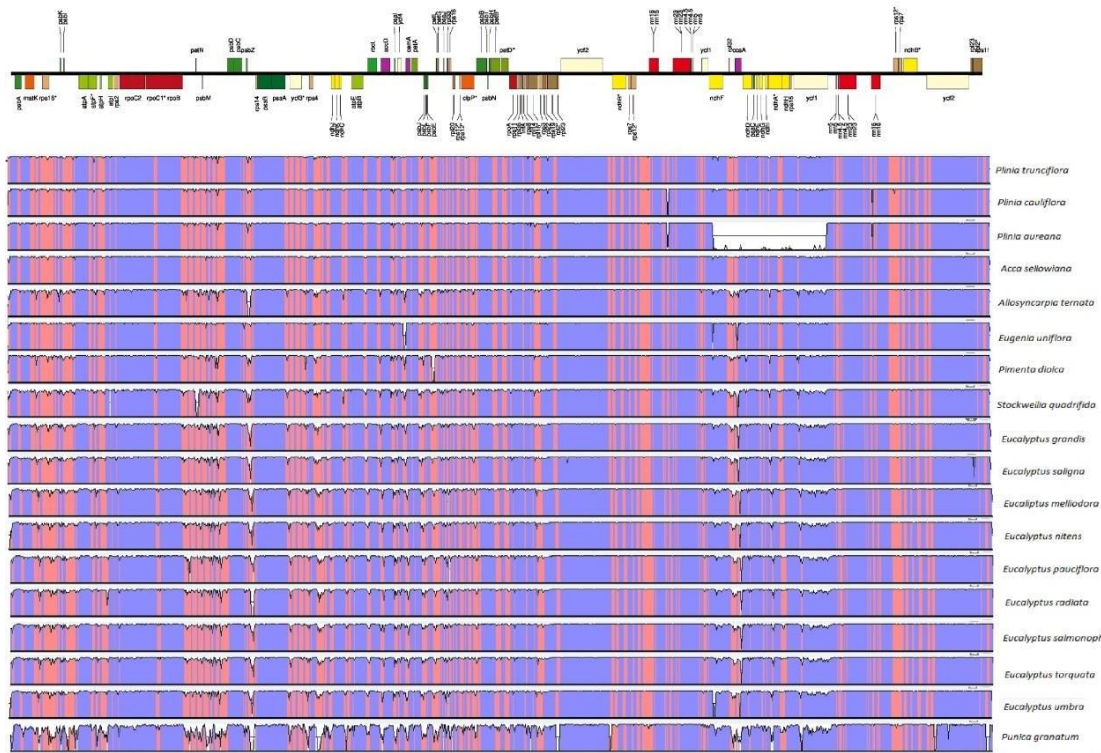


Figura 1. Alinhamento de seqüências completas de plastomas de 18 espécies da família Myrtaceae. Regiões em azul identificam regiões codificantes segundo anotação para *P. guajava*, obtida pelo servidor DOGMA. A representação linear do Plastoma de *P. guajava* foi obtida pelo servidor OGDRAW (Lohse et al., 2013). Regiões em rosa representam regiões não-codificantes conservadas. A espécie *P. granatum* foi utilizada como grupo externo, representando a família Lythraceae. SV = Sítios variáveis.

O mesmo resultado foi obtido na representação linear dos plastomas. Para confirmar esta análise, ao verificar a notação no OGDRAW para *P. guajava*, *P. cauliflora* e *P. aureana* com o intuito de verificar melhor as diferenças entre os genes presentes nos cloroplastos das espécies, é possível observar que na região em que existe um destaque para *P. aureana*, existe uma inversão na seqüência de genes.

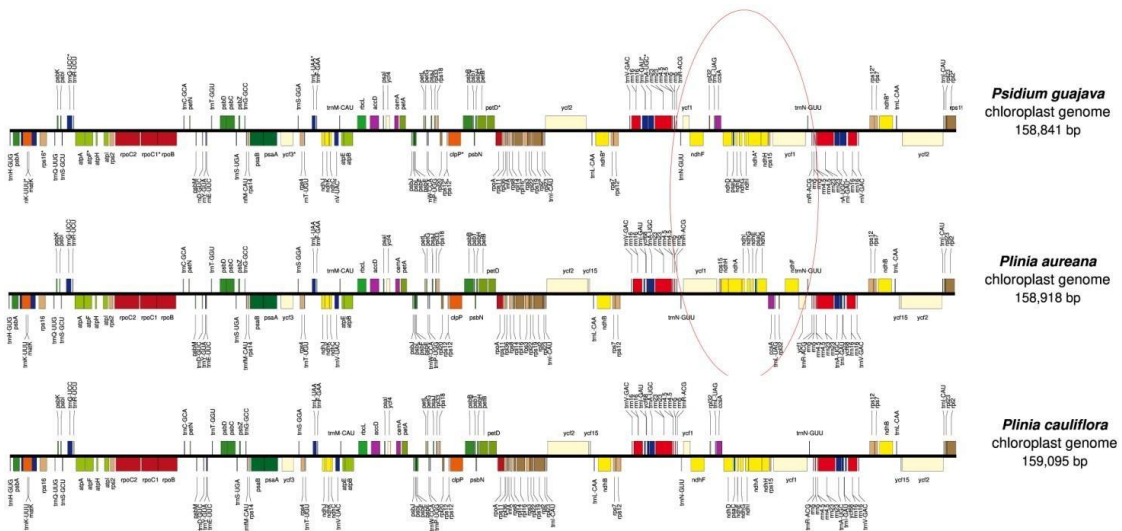


Figura 2. Representação linear do Plastoma completo das espécies *P.guajava* e *P. aureana* obtida pelo servidor OGDRAW (Lohse et al., 2013). A região evidenciada representa a inversão presente no genoma da espécie *P. aureana*.

32 Análises filogenéticas

3.2.1 Família Myrtaceae com genomas plastidiais completos

A figura 3 exibe a correlação filogenética entre espécies da família Myrtaceae. A árvore filogenética resultante dos 16 plastomas completos é bem suportada com todos os *bootstraps* com valores acima de 90%. As tribos Myrteae e Eucalypteae formaram dois grandes grupos com as espécies correspondentes a classificações anteriores (APG IV, 2016). Correlação semelhante foi obtida por Machado (2017) ao realizar análise com plastomas completos da família Myrtaceae.

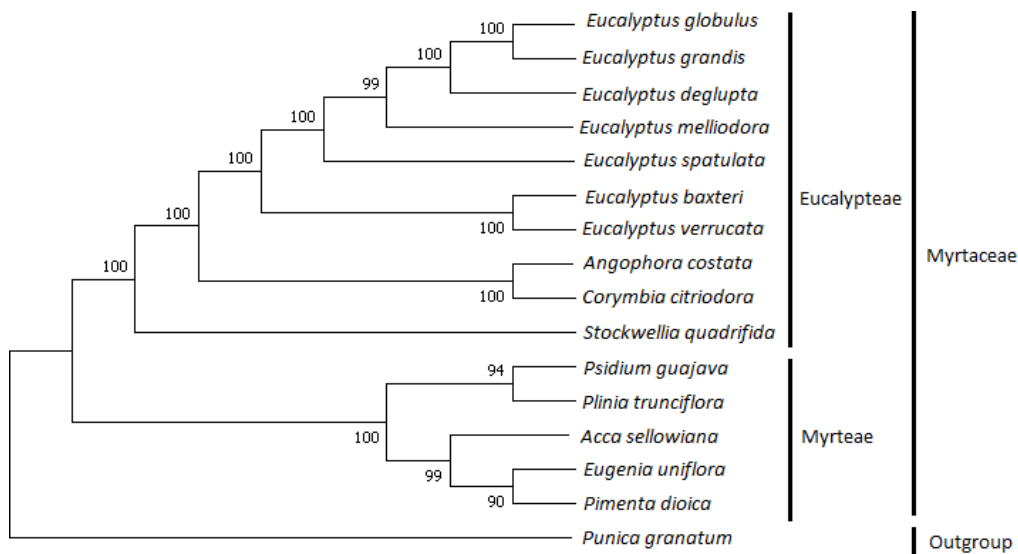


Figura 3: Árvore filogenética com 15 espécies da família Myrtaceae, feita com base no alinhamento do genoma plastidial completo. O plastoma da espécie *Punica granatum* da família Lythraceae, está incluída como grupo externo. A árvore filogenética foi gerada pelo software MEGA7, usando o método de máxima verossimilhança e o suporte de bootstrap foi calculado usando 1000 replicatas. Valores de suporte de bootstrap (%) são mostrados acima da interseção de cada ramo.

3.2.2 Família Myrtaceae com quatro genes plastidiais

Para análises filogenéticas são utilizados vários genes de cloroplasto, como por exemplo, *accD*, *matK*, *ycf1*, *ycf2*, *rpoC2*, *ndhF* são mais utilizados para relacionar espécies pertencentes a mesma família ou gênero; já com os genes *psbC*, *psbB*, *rbcL*, *ndhB*, *psbA*, *psbD*, *psaB* e *psaA*, é possível realizar relações entre os angiospermas e gimnospermas (Kim et al, 2009). Levando esses fatores em consideração para este estudo, foram utilizados dois genes de ambos os grupos visando obter uma relação mais clara entre as espécies avaliadas. A árvore filogenética resultante pode ser observada na figura 4.

A filogenia das espécies da família Myrtaceae utilizando quatro genes plastidiais gerou uma árvore monofilética com um suporte forte de bootstrap de 100%. Entretanto, parte da relação entre as espécies desses clados ainda é difícil de ser estabelecida, principalmente entre as espécies da tribo Eucalypteae, neste grupo estão inclusos as espécies do gênero *Eucalyptus*, *Corymbia* e *Angophora* (APG IV, 2016); essa dificuldade de separação é confirmada com o valor de bootstrap baixo para várias relações entre *Eucalyptus*.

As tribos Myrteae e Eucalypteae foram bem separadas com bootstrap de 99% ao mesmo tempo em que indicam parafiletismo. O clado da tribo Myrteae é composto pelas espécies *P. cauliflora*, *P. trunciflora*, *P. aureana*, *P. guajava*, *Eugenia uniflora*, *Acca sellowiana* e *P. dioica*. A tribo Eucalypteae foi representada por quatro clados, sendo um destes clados composto pelas espécies *Angophora costata*, *Angophora floribunda*, *C. gummifera*, *C. eximia*, *C. maculata*, *C. torelliana*, *C. citriodora* e *C. tessellaris*; os demais clados da tribo Eucalypteae são compostos por espécies do gênero *Eucalyptus*.

A localização da espécie *P. granatum* (romã), utilizada como grupo externo, é fortemente suportada com bootstrap de 100%, formando um clado isolado.

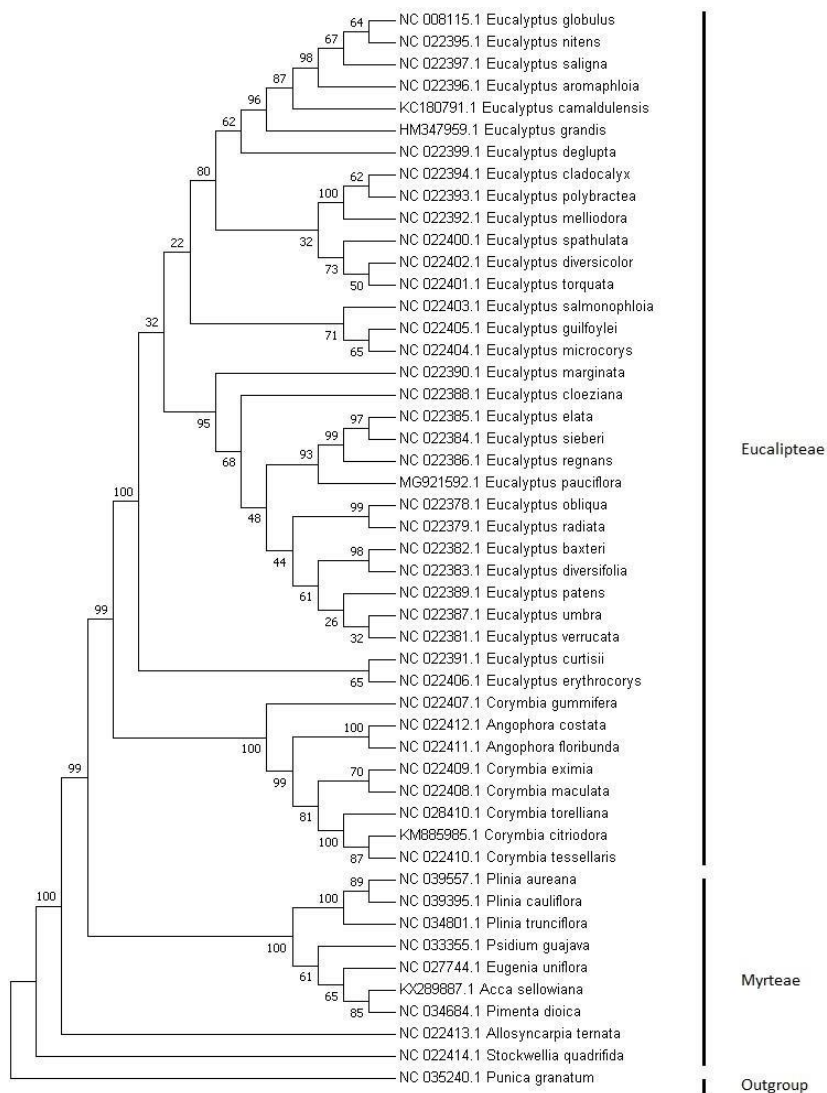


Figura 4. Filogenia da Família Myrtaceae baseada em quatro genes plastidiais: *matK*, *rbcL*, *ndhF* e *atpB*. Foi incluída como grupo externo a espécie *Punica granatum* (NC_0352401). A árvore filogenética foi gerada pelo software MEGA7, usando o método de máxima verossimilhança, e o suporte de bootstrap foi calculado usando 1000 replicatas. Valores de suporte de bootstrap (%) são mostrados na interseção de cada ramo. Ordens e famílias indicadas seguem a classificação de APG-IV (2016).

Os eucaliptos (tribo Eucalypteae) formaram um grande grupo, no qual é possível observar valores de bootstrap baixos indicando que a relação naquele ponto ainda não é bem definida, assim como já foi visto por Wilson et al. (2005), Udovicic e Ladiges (2000), e Steane et al. (2002). Com estudos de biogeografia, morfologia, RFLPs e dados de regiões de sequência de DNA plastidial, é possível entender a história geológica e evolução dos eucaliptos e como houve a separação dos gêneros, afirmando que não são sinônimos (Udovicic e Ladiges, 2000). Neste grupo estão as espécies de gênero *Angophora*, *Corymbia* e *Eucalyptus*, todas as plantas deste grupo são utilizadas na produção de celulose e de móveis, pois possuem madeira de boa qualidade (Via Campesina, 2006).

A filogenia obtida neste trabalho para tribo Myrteae, é semelhante à encontrada por (Machado, 2017), onde a separação dos grupos mostrou um suporte de bootstrap de 100%. Neste grupo estão inclusas espécies de importância econômica como a goiabeira (*P. guajava*), utilizada para produção de diversos alimentos e consumo in natura; pitanga (*Eugenia uniflora*), uso principal para produção de polpas; e jabuticaba (*Plinia trunciflora*), geralmente consumida in natura.

3.2.3 Filogenia Ordem Myrtales

As sequências das espécies da Família Myrtales alinhadas resultou em 6.426 bases sendo estas 1.288 para *atpB*, 1515 para *matK*, 2206 para *ndhF* e 1418 para *rbcL*; com 3.809 sítios semelhantes entre si e 2.862 sítios variáveis com os valores de 763 para *atpB*, 863 para *matK*, 1004 para *ndhF* e 298 para *rbcL*. O conteúdo GC(%) variou de 42,5, 33,9, 31,8, 43,2 para *atpB*, *matK*, *ndhF* e *rbcL*, respectivamente e de 36,8 para o valor combinado. A análise realizada com a união destes genes não foi feita anteriormente por outros autores.

Tabela 2: Estatística dos quatro genes utilizados neste estudo.

Característica	atpB	matK	ndhF	rbcL	Combinado
Alinhado	1288.4	1515.7	2206,6	1418	6426,7
Incluso	918	1432	1305	1005	3809
Sítios variáveis	763	863	1004	298	2862
Sítios Parsimônia	699	660	762	210	2333
informativos					
G+C(%)	42,5	33,9	31,8	43,2	36,8

A análise filogenética da ordem Myrtales gerou um parafiletismo no qual em um grupo engloba a família Melastomataceae e no outro as demais famílias da ordem utilizadas na análise. Que estão inclusas: Myrtaceae, Vochysiaceae, Combretaceae, Onagraceae e Lythraceae. Essas famílias, em maioria, estão separadas por relações fortemente suportadas, a exceção é apenas considerada fracamente suportada (bootstrap 62%) para o parafiletismo que divide a família Melastomataceae das famílias Myrtaceae, Vochysiaceae, Combretaceae, Onagraceae e Lythraceae. Esta classificação da ordem Myrtales demonstra correlação entre as espécies e mantém a classificação anterior realizada com dados morfológicos (APG IV).

Nesta árvore (Figura 5) para Myrtaceae, as espécies que na árvore anterior não haviam sido incluídas na tribo Eucalypteae agora estão, que são *Alossyncarpia ternata* e *Stockwellia quadrifida* no mesmo clado em que estão as espécies do gênero *Angophora* e *Corymbia*. Há um parafiletismo com as espécies da tribo Myrteae e Eucalypteae, indicando a possibilidade de haver uma ancestralidade comum entre os dois grupos. Destaque para outra espécie da família Myrtaceae, o *Syzygium forrestii*, espécie da tribo Syzigeae, formou uma linhagem a parte fortemente suportado com Bootstrap de 100%. Análises deste tipo não foram realizadas anteriormente em outros trabalhos para esta família.

O grupo externo formado por espécies da ordem Geraniales gerou um grupo que teve origem em um ancestral de todas as espécies da ordem Myrtales com bootstrap fortemente suportado de 99%.

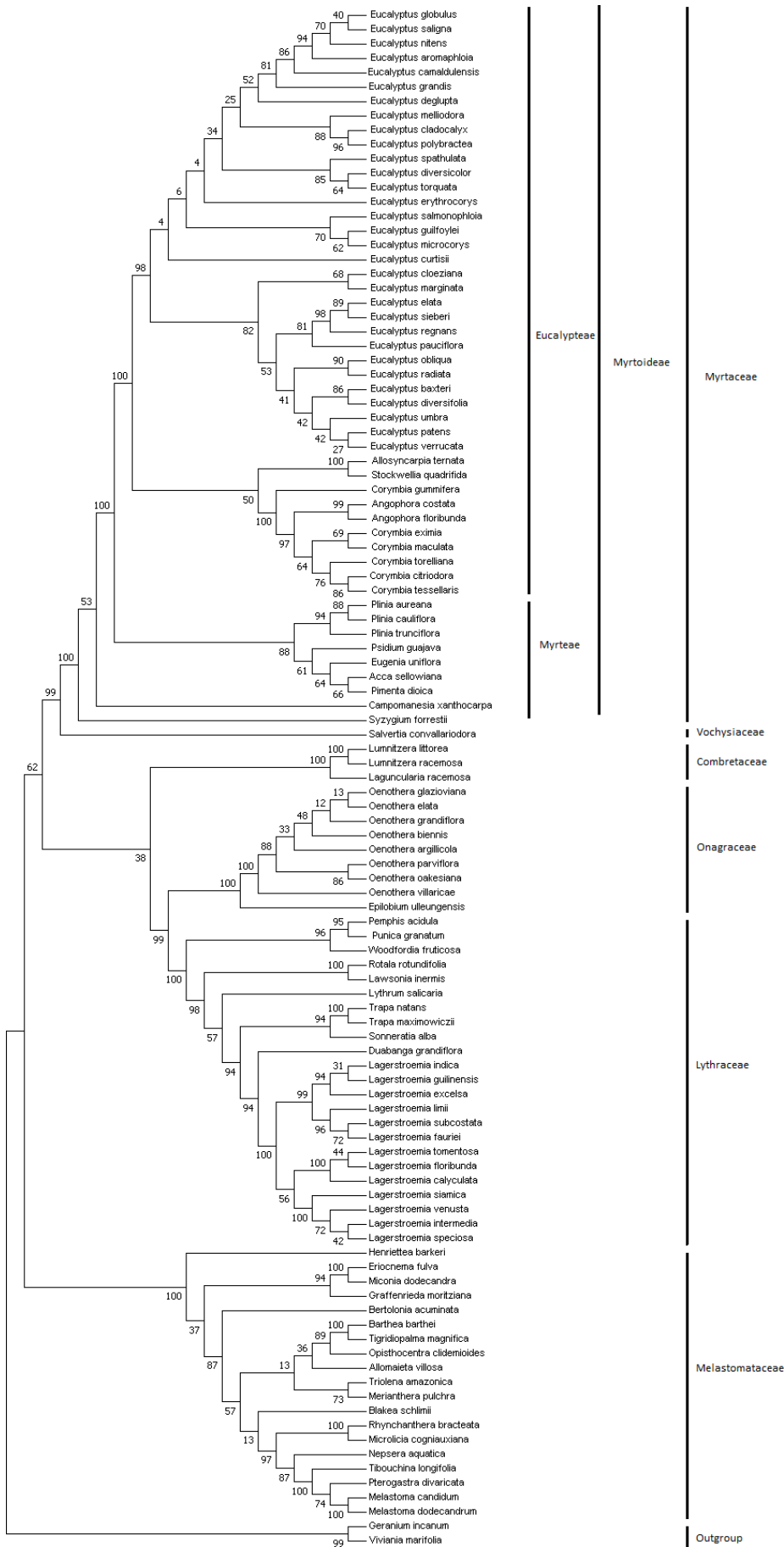


Figura 5. Filogenia da ordem Myrtales baseada em quatro genes plastidiais: *matK*, *rbcL*, *ndhF* e *atpB*. A árvore filogenética foi gerada pelo software MEGA7, usando o método de máxima verossimilhança, e o suporte de bootstrap foi calculado usando 1000 replicatas. Valores de suporte de bootstrap (%) são mostrados na interseção de cada ramo. Ordens e famílias indicadas seguem a classificação de APG- IV (2016).

Na Figura 6 é possível observar o tempo de divergência estimado acima dos pontos de calibração. Esta avaliação para a ordem Myrtales, foi obtida através do método Reltime (Tamura et al., 2012) e o método de Tamura-Nei (Tamura et al., 1993), utilizando-se dois pontos de calibração obtidos através do servidor TimeTree: sendo um ponto da ordem Geraniales (*outgroup*) e um ponto da família Myrtaceae com a espécie *P. guajava*.

Para a família Myrtaceae, o grande grupo da tribo Eucalypteae revelou um tempo de divergência curto entre as espécies, com os valores de 0,69 a 0,0 Ma (Milhões de anos). Período do Oligoceno. Este curto tempo de divergência pode ser um dos grandes desafios de identificar as diferentes espécies de eucalipto, confirmando quando se une a informação de bootstrap obtido na filogenia anterior, onde foram encontrados valores baixos, como por exemplo BS=40% entre *Eucalyptus globulus* e *Eucalyptus saligna*, espécies que o tempo de divergência entre elas foi de 0,0 Ma. O tempo de divergência entre a tribo Myrteae e Eucalypteae encontrado foi de aproximadamente 29Ma, provavelmente diferenciarem-se no Mioceno.

Para a família Combretaceae, o tempo de divergência entre as espécies foi de 13,52 a 2,49 Ma. Para a família Onagraceae, os representantes do gênero *Oenothera*, tiveram o tempo de divergência de 0,0 Ma, onde de oito espécies, cinco destas na correlação filogenética obtiveram valores de bootstrap menores que 40%; nesta família, o tempo de divergência foi de 10,8 a 0,0 Ma. Estima-se que a divergência das duas famílias se deu entre os períodos Mastrinciano e Campaniano (Time Tree, 2020). Pertencendo ao mesmo clado da família Onagraceae encontra-se a família Lythraceae, onde os valores de divergência variaram de 27,66 a 0,0 Ma. Onde mais uma vez, espécies que tiveram baixo tempo de divergência, tiveram também baixo valor de bootstrap.

A família Melastomataceae ocupou um clado diferente das demais, com tempo de divergência de 0,55 a 9,97 Ma. A linhagem que divide a família Melastomataceae das demais indica divergência a cerca de 67,46 Ma. Estima-se que a divergência das

famílias Melastomataceae e Lythraceae se deu também entre as Eras Mastrinciano e Campaniano (Time Tree, 2020).

As correlações em espécies de Myrtales, de forma geral, são avaliadas dentro de cada família de forma separada (Conti et al, 2002; Sytsma et al, 2004; Clausing e Renner, 2001; Schönenberger e Conti, 2003 e Wilson et al, 2005) e com dados do APG IV (2016). A junção desses estudos gerou uma relação filogenética que dispõe as famílias estudadas da seguinte forma: Combretaceae, Onagraceae, Lythraceae, Myrtaceae, Vochysiaceae e Melastomataceae. A posição da família Combretaceae ainda é considerada incerta (Berger e Sytsma, 2010; Kriebel et al, 2017).

As famílias nesta filogenia não seguiram a ordem de classificação proposta anteriormente, a sequência foi de Myrtaceae, Vochysiaceae, Combretaceae, Onagraceae, Lythraceae e Melastomataceae, finalizando com as duas espécies do grupo externo pertencentes à ordem Geraniales; uma relação um similar à obtida por Berger et al (2016) baseado no APG IV.

Essas famílias se diferenciam taxonomicamente por possuírem tipos diferentes de flores, como foi visto por Vasconcelos (2013), que classificou as florações em diferentes espécies da ordem em tipos A, B, C, D, E, F, G, H, I, J, K e L. Com as da família Myrtaceae sendo possuidora de flores tipo B, I e J que possuem como polinizadores morcegos, pássaros e insetos; da família Vochysiaceae com flores tipo D e L, com os polinizadores abelha e demais insetos; com a família Combretaceae com flores tipo A com anemofilia e insetos pequenos como estratégia de polinização; Onagraceae com flores tipo C, E, G e L com apenas insetos como polinizadores; Lythraceae com flores tipo J, sendo polinizada por insetos; e Melastomataceae com flores tipo C e G também polinizadas por insetos.

O trabalho de Thornhill (2015) estima o tempo de aproximadamente 40Ma para surgimento do gênero *Eucalyptus* e aproximadamente 20Ma para diferenciação do gênero *Angophora*, faixa aproximada da encontrada neste trabalho com a tribo Eucalypteae se diferenciando da tribo Myrteae por volta dos 29Ma, este mesmo valor também foi encontrado para *Syzygium forrestii*.

Para Berguer et al (2016) o tempo de divergência indica que a ordem Myrtales divergiu de Geraniales em 124 milhões de anos, ainda no Cretáceo Inferior. Dentro da ordem Myrtales a divergência ocorreu entre 67-64 milhões de anos para a família Melastomataceae, 75 milhões de anos para Myrtaceae e entre 50-45 milhões de anos para Combretaceae.

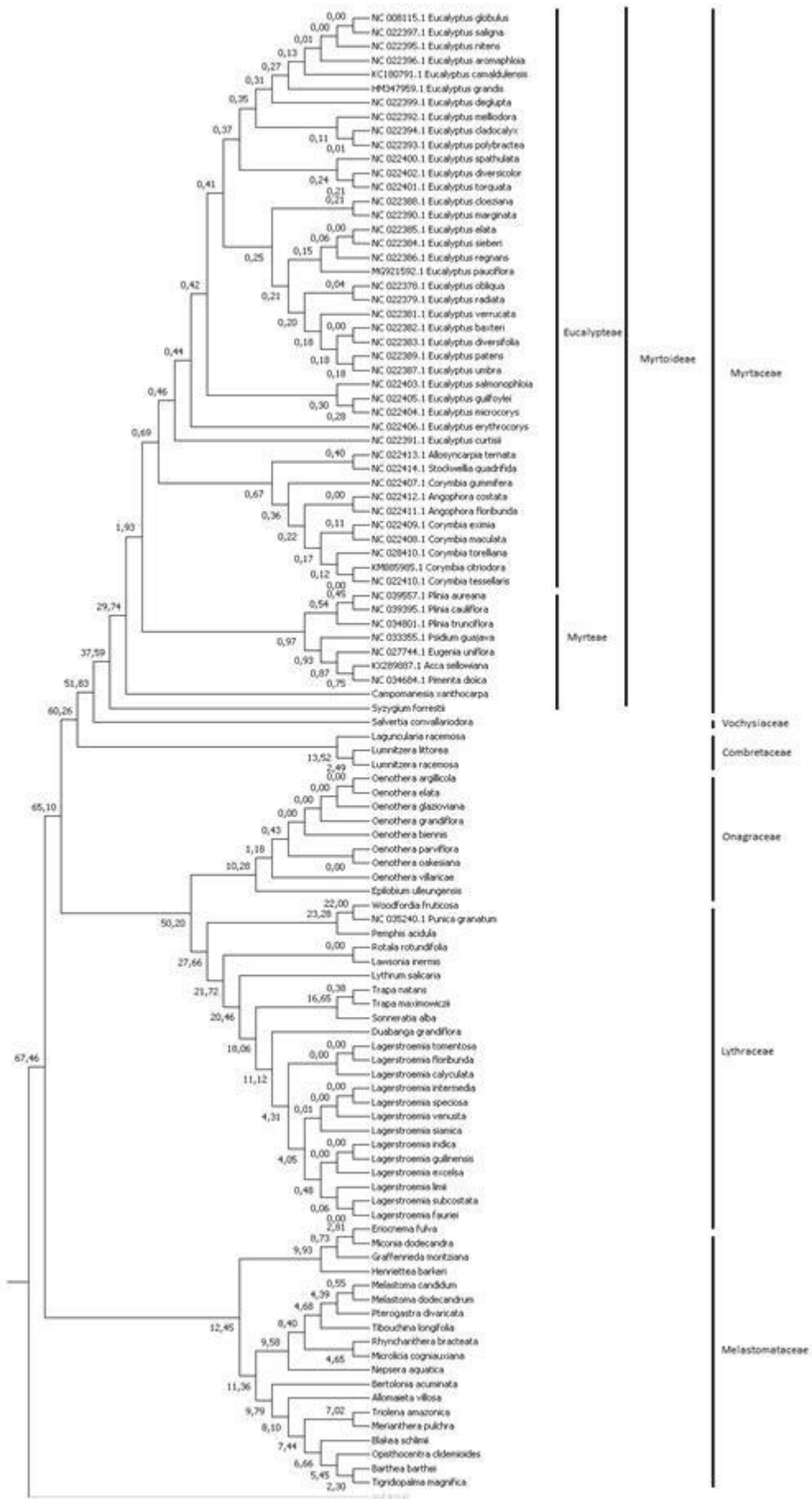


Figura 6. Filogenia e Árvore do tempo de 108 espécies da ordem Myrtales, baseada em quatro genes plastidiais completos. Genes de duas espécies foram incluídas como grupo externo. A árvore filogenética foi gerada pelo software MEGA7, usando o método de máxima verossimilhança, e o suporte de bootstrap foi calculado usando 1000 replicatas. Valores de suporte de bootstrap (%) são mostrados abaixo da interseção de cada ramo. A árvore do tempo foi inferida pelo método Realltime (Tamura et al., 2012) e pelo modelo de Tamura-Nei (Tamura et al., 1993). A árvore foi computada utilizando dois pontos de calibração obtidos pelas escalas disponíveis no site Timetree (<http://www.timetree.org/>). Classificação taxonômica de subfamílias e tribos Segundo Groppo et al.(2008). Valores de divergência do tempo entre as espécies são mostradas acima da interseção de cada ramo.

4. CONCLUSÃO

As correlações filogenéticas verificadas neste trabalho, demonstraram semelhanças com as correlações evolutivas da família Myrtaceae estabelecidas com dados morfológicos, assim como também foram obtidas correlações semelhantes para a classificação de espécies da ordem Myrtales. Verifica-se então, a possibilidade do uso dos genes plastidiais atpB, matK, rbcL e ndhF em conjunto para confirmar relações de ancestralidade nesta ordem e família de forma confiável.

5. AGRADECIMENTOS

À equipe e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pela concessão da bolsa de mestrado.

6. REFERÊNCIAS BIBLIOGRÁFICAS

1. Alves, HS. **Sequenciamento e análise do genoma cloroplastidial de eucalipto (*Eucalyptus grandis*)**. 2005. Tese de Doutorado. Universidade de São Paulo.
2. Amorim BS. Filogenia e estudos taxonômicos do clado Gomidesia (Myrtaceae, Myrcia SL) na Floresta Atlântica do Brasil. 2017.
3. APG IV. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181: 1–20, 2016.
4. Bendich A. Circular chloroplast chromosomes: the grand illusion. *Plant Cell* 16:1661–1666, 2004.
5. Berger, BA. Divergence times, historical biogeography, and shifts in speciation rates of Myrtales. *Molecular Phylogenetics and Evolution*, v. 95, p. 116-136, 2016.
6. Brudno M, Do CB, Cooper GM, Kim MF, Davydov E. 2003. LAGAN and Multi LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* 13: 721–731. DOI: [10.1101/gr.926603](https://doi.org/10.1101/gr.926603)
7. Chambers GK, Macavoy ES. Microsatellites: consensus and controversy. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, v. 126, n. 4, p. 455-476, 2000.
8. COM FLORES, Sistemática das Plantas. Filogenia das Angiospermas. 2017.
9. Cruz-Maya ME, & Barrientos-Priego AF. (2018). Phylogenetic analysis of some members of the sub genus *Persea* (*Persea*, Lauraceae) Análisis filogenético de algunos miembros del subgénero *Persea* (*Persea*, Lauraceae). *Revista Chapingo Serie Horticultura*, 24(2).
10. Dhingra A, Folta KM. ASAP: Amplification, sequencing; annotation of plastomes. *BMC Genomics*. v. 6, p.176, 2005.
11. Faria Júnior JEQD. (2014). Revisão taxonômica e filogenia de *Eugenia* sect. *Pilothecium* (*Kiaersk.*) D. Legrand (Myrtaceae).
12. Felsenstein J. 1985. Confidence limits on phylogenies: na approach using the bootstrap. *Evolution* 39: 783-791.
13. FERES JM. Diversidade genética, sistema reprodutivo e fluxo de pólen em duas populações de *Tabebuia roseo-alba* (Ridl.) Sand.: implicações para a conservação. Tese de Doutorado. Universidade de São Paulo, 2009.
14. Ferreira ME, Grattapagli A D. Introdução ao Uso de Marcadores Moleculares em Análise Genética. 2. ed. Brasília: Embrapa Cenargen. 220 p. 1998.
15. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 32:273–279.
16. Gressler E, Pizo MA. & Morellato LPC. Polinização e dispersão de sementes em Myrtaceae do Brasil. *Brazilian Journal of Botany*, p. 509-530, 2006.
17. Gu C, Dong B, Xu L, Tembrock L, Zheng S, Wu Z. The complete chloroplast genome of *Heimia myrtifolia* and comparative analysis within Myrtales. *Molecules*, v. 23, n. 4, p. 846, 2018.
18. Guimarães CT, Magalhães JV, Lanza MA, Schuster I. Marcadores moleculares e suas aplicações no melhoramento genético. Embrapa Milho e Sorgo-Artigo em periódico indexado (ALICE), 2009.

19. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets, *Molecular Biology and Evolution* 33:1870–1874. DOI:10.1093/molbev/msw054.
20. Lin CP, Huang JP, Wu CS, Hsu CY, Chaw SM. Comparative chloroplast genomics reveals The evolution of Pinaceae genera and sub families. *Genome biology and evolution*, v. 2, p. 504-517, 2010.
21. Lorenzi H; Souza WC. Botânica sistemática. Guia ilustrado para identificação das famílias de fanerógamas nativas e exóticas do Brasil, baseado em APG II. Editora Plantarum. 2ª ed. 640 p., 2008.
22. Machado, L. D. O., Sarzi, D. S., Oliveira, N. M., &Stefenon, V. M. (2012). Análises In Silico das Relações Filogenéticas entre Plantas. *Anais do Salão Internacional de Ensino, Pesquisa e Extensão*, 4(4).
23. Machado, L. O. Análise comparativa do genoma plastidial de Myrtaceae: *Accasellowiana* (O. Berg) Burret, *Eugenia uniflora* L., *Campomanesia xanthocarpa* (Mart.) O. Berg, *Plinia cauliflora* (Mart.) Kausel, *Plinia aureana* (Mattos) Mattos e *Plinia* sp. 2017.
24. Machado, L. Phylogenomic relationship of feijoa (*Acca sellowiana* (O. Berg) Burret) with other Myrtaceae based on complete chloroplast genome sequences. *Genetica*, v. 145, n. 2, p. 163-174, 2017.
25. Machado, LO. Análise comparativa do genoma plastidial de Myrtaceae: *Acca sellowiana* (O. Berg) Burret, *Eugenia uniflora* L., *Campomanesia xanthocarpa* (Mart.) O. Berg, *Plinia cauliflora* (Mart.) Kausel, *Plinia aureana* (Mattos) Mattos e *Plinia* sp. 2017.
26. Morais LMF, Conceição GM, Nascimento JM. Família Myrtaceae: análise morfológica e distribuição geográfica de uma coleção botânica. *Agrarian Academy, Centro Científico Conhecer - Goiânia*, v. 1, p. 317-346, 2014.
27. Santos MF. (2014). *Biogeografia de Myrciasl, taxonomia e filogenia do clado Sympodio myrcia (Myrtaceae)* (Doctoral dissertation, Universidade de São Paulo).
28. Takeuchi R, Kimura S, Saotome A, Sakaguchi K. Biochemical properties of a plastidial DNA polymerase of rice. *Plant molecular biology*, v. 64, n. 5, p. 601-611, 2007.
29. Tamura K, Battistuzzi FU, Billig RP, Murillo O, Filipinski A, Kumar S. 2012. Estimating divergence times in large molecular phylogenies. *Proc Natl Acad Sci USA*, 109:19333–19338. DOI:10.1073/PNAS.1213199109.
30. Tamura K, Nei M .1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *MolBiol Evol.*10:512–526. DOI:10.1093/oxfordjournals.molbev.a040023
31. Tangphatsornruang S, Sangsrakru D, Chanprasert J, Uthapaisanwong P. Yoocha, T, Jomchai, N, Tragoonrung S. The chloroplast genome sequence of mung bean (*Vignaradiata*) determined by high-through put pyro sequencing: structural organization and phylogenetic relationships. *DNA Res* v. 17, p. 11–22, 2010.
32. Temnykh S, Declerck G, Lukashova A. 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential, *Genome Res.*, 11, 1441–1452. DOI: 10.1101/gr.184001.

33. Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673-4680.
34. Thornhill A. H, Ho SY, Külheim C, Crisp MD. Interpreting the modern distribution of Myrtaceae using a dated molecular phylogeny. *Molecular Phylogenetics and Evolution*, v. 93, p. 29-43, 2015.
35. Venancio LPR. Isolamento e caracterização de microssatélites em Búfalo de rio (*Bubalus bubalis*) a partir da construção de bibliotecas genômicas parciais com hibridização seletiva, 2008.
36. Wilson PG, O'Brien MM, Heslewood MM, Quinn CJ. Relationships within Myrtaceae sensu lato based on a matK phylogeny. *Plant Systematics and Evolution*, v. 251, n. 1, p. 3-19, 2005.
37. Wyman SK, Jansen RK, Boore JL. 2004. *Bioinformatics* 20:3252–3255.

ANEXO A - Normas para preparação e submissão de trabalhos (Scientia Plena)

Condições para submissão

Como parte do processo de submissão, os autores são obrigados a verificar a conformidade da submissão em relação a todos os itens listados a seguir. As submissões que não estiverem de acordo com as normas serão devolvidas aos autores.

- A contribuição é original e inédita, não sendo avaliada para publicação por outra revista. O arquivo principal para submissão está em formato Microsoft Word (desde que não ultrapasse 2MB).
- O arquivo está preparado de acordo com os padrões de estilo e requisitos bibliográficos que constam no artigo-exemplo: **http://scientiaplena.org.br/public/journals/1/Modelo_Artigo_2015.docx**
- Os nomes completos de todos os autores, bem como sua afiliação institucional, foram cadastrados de acordo com a ordem de autoria apresentada no trabalho (campo Metadados).
- Foram indicados, no campo Comentários ao Editor, a área e subárea de conhecimento do trabalho.
- Foram indicados, no campo Comentários ao Editor, três possíveis avaliadores para a submissão (nome completo, email e afiliação institucional).

Declaração de Direito Autoral

Autores que publicam nesta revista concordam com os seguintes termos:

- Autores mantêm os direitos autorais e concedem à revista o direito de primeira publicação, com o trabalho simultaneamente licenciado sob a Licença Creative Commons Attribution que permite o compartilhamento do trabalho com reconhecimento da autoria e publicação inicial nesta revista.
- Autores têm autorização para assumir contratos adicionais separadamente, para distribuição não-exclusiva da versão do trabalho publicada nesta revista (ex.: publicar em repositório institucional ou como capítulo de livro), com reconhecimento de autoria e publicação inicial nesta revista.
- Autores têm permissão e são estimulados a publicar e distribuir seu trabalho online (ex.: em repositórios institucionais ou na sua página pessoal) a qualquer ponto antes ou durante o processo editorial, já que isso pode gerar alterações produtivas, bem como aumentar o impacto e a citação do trabalho publicado.

Política de Privacidade

Os nomes e endereços informados nesta revista serão usados exclusivamente para os serviços prestados por esta publicação, não sendo disponibilizados para outras finalidades ou a terceiros.

Taxas para autores


Este periódico cobra as seguintes taxas aos autores.

- Publicação de artigo: 200,00 (BRL)
- Caso o manuscrito submetido seja aceito para publicação, será necessário o pagamento de uma Taxa de Publicação de Artigo para auxiliar nos custos de publicação.

Terá desconto na taxa de submissão:

- 1) Sócio adimplente da Associação Sergipana de Ciência;
- 2) Avaliador que tenha emitido parecer solicitado pela revista.

Estrutura do artigo:



SCIENTIA PLENA
www.scientiaplena.org.br

VOL. X, NÚM. X 20XX
doi: 10.14808/sci.plena.20XX.XXXXXX

Título do trabalho
XXXXX

X. X. Sobrenome¹, X. X. Sobrenome²

Nome de Departamento/Cabocabinho/Gen; Nome de Instituição, CEP, Cidade-Estado, País
Nome de Departamento/Cabocabinho/Gen; Nome de Instituição, CEP, Cidade-Estado, País

*email@dominio@provedor@brasil.com
(E-mail em dia de mar do ano, avião em dia de mar de ano)

O resumo deve ser iniciado aqui e não pode ultrapassar 250 palavras.
Palavras-chave: palavra-chave 1, palavra-chave 2, palavra-chave 3

Insert the abstract here.
Keywords: keyword 1, keyword 2, keyword 3

1. INTRODUÇÃO

Na seção de Introdução do artigo, o autor deve descrever o estado-da-arte do problema, além de justificar e apresentar os objetivos do seu trabalho.

Neste modelo, que está formatado segundo o modelo adotado pela revista, aproveitaremos esta seção para apresentar algumas informações sobre a submissão de artigos à **Scientia Plena**.

A **Scientia Plena** é uma publicação científica mensal e aceita manuscritos originais e inéditos, redigidos em português, inglês ou espanhol. Artigos de revisão não são aceitos para publicação.

Trabalhos que utilizam seres humanos como objeto de estudo ou experimentação animal devem indicar no texto o número da aprovação pelos respectivos Comitês de Ética.

O trabalho não deverá estar sendo avaliado simultaneamente por outra revista e todos os autores devem estar cientes da submissão.

O trabalho deve ser submetido pelo sistema eletrônico da revista em formato "doc", com tabelas e figuras incluídas no corpo do texto. Todo o corpo do texto deve ser redigido em Times New Roman, tamanho 11, justificado e com espaçamento simples. As margens das páginas devem ser de 2,5 cm (superior e inferior) e 3,0 cm (esquerda e direita). Todos os parágrafos devem apresentar tabulação de 0,5 cm e as tabelas e figuras devem ser citadas por extenso no corpo do texto (ex: Figura 1; Tabela 1). Ao longo do texto deve ser utilizado o sistema internacional de unidades (SI) para indicação de medidas.

Para citação das referências, utilizar o Estilo Vancouver, com a numeração entre colchetes e alinhada ao texto. Exemplos: "... para determinados valores [1]..."; "... Segundo **Medeira**, et al. (2005) [2]..."; "... estudos de raios de sílex [3]..."; "... o tamanho da amostra [4]..."; "... o uso de drogas para alívio da dor [5, 6]...".

A lista de referências deve ser apresentada ao final do texto, em seção específica. Não usar notas de rodapé.

2. MATERIAL E MÉTODOS

A metodologia deve ser descrita com as informações necessárias para permitir a repetição do estudo por outro pesquisador.

3. RESULTADOS E DISCUSSÃO

Resultados e Discussão podem ser apresentados em conjunto ou em subtítulos separados.

X. X. Sobrenome/Sobrenome et al., Scientia Plena X, XXXXX (20XX) XX

Tabelas e figuras devem ser centralizadas, com legenda objetiva e autoexplicativa. Tabelas não devem apresentar linhas verticais secundárias. Devem-se evitar tabelas e/ou figuras com poucas informações, que podem ser facilmente substituídas por texto corrido.




Figura 1. Legenda da figura

Tabela 1. Exemplo de modelo de tabela

Título	Título		
	Coluna 1	Coluna 2	Coluna 3
Linha 1	XXX	XXX	XXX
Linha 2	XXX	XXX	XXX
Linha 3	XXX	XXX	XXX
Linha 4	XXX	XXX	XXX

4. CONCLUSÃO

Uma conclusão deve ser apresentada com as principais contribuições do estudo.

5. AGRADECIMENTOS

Apresentar os agradecimentos pertinentes, se houver.

6. REFERÊNCIAS BIBLIOGRÁFICAS

1. **Bello** DB, **Crooks** VC, **DeGroot** JG, **Chiu** V. Blood pressure levels before dementia. *Arch Neurol*. 2005; 62(11):112-6. doi:10.1001/archneur.62.1.112.
2. **Murphy** P, **Wassenaar** X, de **Wolff** HE, **MacGregor** GA. Links between dietary salt intake, renal salt handling, blood pressure, and cardiovascular diseases. *Hypertens Rev*. 2005; 4(2):670-715. doi: 10.1152/hyperev.00056.2003
3. **Jenkins** PF. Making sense of the chest x-ray: a hands-on guide. New York: Oxford University Press; 2005. 194 p.
4. **DeGroot** RH. *Statistics in medicine*. 2nd ed. Amsterdam (Netherlands): Elsevier Academic Press; 2006. Chapter 24, Regression and correlation methods. p. 447-86. doi: 10.1016/B978-0-12-364864-2.00025-1
5. **Zhao** C. Development of **apoptosis** and application to protein research and drug discovery [dissertation]. Buffalo (NY): State University of New York at Buffalo; 2005. 276 p.
6. **Rice** AS, **Francher-Smith** WP, **Becker** D, **Brooks** JW, **Capobianco** and *et al.* In: **DeGroot** JG, **Go** DB, **Wassenaar** M, editors. *Proceedings of the 10th World Congress on Fairs*; 2002. Aug 17-22; San Diego, CA. Seattle (WA): LAMP Press; c2003. p. 437-68.

CAPÍTULO III

A ser enviado para a revista GMR – Genetics and Molecular Research (Qualis – A2)

Caracterização *in silico* e análise filogenética de uma 1-aminociclopropano-1-carboxilato sintase em representantes de Dicotiledôneas

G. E. Santos^{1,3}; H. J. Jimenez²; R. M. M. Filho^{1,3}; A.V.V. Montarroyos²; L. S. S. Martins^{1,3}

¹Departamento de Biologia; Universidade Federal Rural de Pernambuco; Recife - PE - Brasil

²Departamento de Agronomia; Universidade Federal Rural de Pernambuco; Recife - PE - Brasil

³Programa de Pós Graduação em Melhoramento Genético de Plantas; Universidade Federal Rural de Pernambuco; Recife - PE - Brasil

RESUMO: Myrtaceae é uma família de grande importância econômica e ecológica no Brasil, onde se destacam os eucaliptos, por sua grande importância na produção de madeira para uso industrial, e diversas espécies da tribo Myrteae como a goiaba, pitanga, araçá e jaborcaba, que são utilizados para produção de alimentos sejam eles *in natura* ou industrializados. Com a evolução das ferramentas de sequenciamento de DNA e proteínas, análises baseadas nestas moléculas tem sido cada vez mais usadas para estabelecer correlações evolutivas entre organismos em conjunto dados morfológicos. A 1-aminociclopropano-1-carboxilato sintase, também conhecida como ACC sintase, é uma enzima que desempenha um papel significativo na síntese de etileno. Este hormônio gasoso é de extrema importância para o desenvolvimento de frutos e atua também como um regulador de vários processos fisiológicos e de desenvolvimento, como dormência de sementes, germinação, crescimento vegetativo, floração e senescência. Com o objetivo de caracterizar, comparar e identificar seus domínios conservados, sequências das proteínas ACC sintase foram recuperadas do banco de dados GenBank através da ferramenta BlastP. Os parâmetros físico-químicos foram avaliados através da ferramenta ProtParam. A identificação de domínios funcionais foi avaliada por meio do servidor ProDom e a estimativa de efeitos mutacionais foi feita por meio do servidor SNAP2. O alinhamento de sequência das proteínas foi realizado por meio do algoritmo ClustalW e as árvores filogenéticas produzidas com o software MEGA 7 utilizando o método de máxima verossimilhança. Para a predição de estrutura das proteínas em modelos 3D foi utilizado o servidor Phyre2. As árvores filogenéticas das espécies de dicotiledôneas construídas por meio da sequência da proteína da ACC sintase apresentam consistência com proposta de classificação apresentada na literatura. No entanto os métodos utilizados se mostraram suficientes para a separação até o nível de tribo dentro das famílias Malvaceae, Rutaceae, Euphorbiaceae, Vitaceae, Moraceae, Cannabaceae, Solanaceae, Compositae, Myrtaceae, Juglandaceae, Fagaceae, Rosaceae e Leguminosae. O uso de ferramentas *in silico* se mostrou viável para construção de um modelo 3D de proteínas por meio de modelagem por homologia. Os modelos 3D gerados demonstraram estabilidade quanto ao nível energético, possibilitando a análise de suas estruturas terciárias e funções moleculares.

ABSTRACT: Myrtaceae is a botanical family of great economic and ecological importance in Brazil, where eucalyptus stands out, for its great importance in the production of wood for industrial use, and several species of the Myrteae tribe such as guava, pitanga, araçá and jaboticaba, which they are used for food production whether fresh or processed. With the evolution of DNA and protein sequencing tools, analyzes based on these molecules have been increasingly used to establish evolutionary correlations between organisms together with morphological data. 1-aminocyclopropane-1-carboxylate synthase, also known as ACC synthase, is an enzyme that plays a significant role in the synthesis of ethylene. This gaseous hormone is extremely important for fruit development and also acts as a regulator of several physiological and developmental processes, such as seed dormancy, germination, vegetative growth, flowering and senescence. In order to characterize, compare and identify their conserved domains, sequences of the ACC synthase proteins were retrieved from the GenBank database using the BlastP tool. The physical-chemical parameters were evaluated using the ProtParam tool. The identification of functional domains was evaluated using the Prodom server and the mutational effects were estimated using the SNAP2 server. The protein sequence alignment was performed using the ClustalW algorithm and the phylogenetic trees produced with the MEGA 7 software using the maximum likelihood method. To predict the structure of proteins in 3D models, the Phyre2 server was used. Phylogenetic trees of dicotyledonous species built using the ACC synthase protein sequence are consistent with the classification proposal presented in the literature. However, the methods used proved to be sufficient for separation up to the tribe level within the Malvaceae, Rutaceae, Euphorbiaceae, Vitaceae, Moraceae, Cannabaceae, Solanaceae, Compositae, Myrtaceae, Juglandaceae, Fagaceae, Rosaceae and Leguminosae families. The use of *in silico* tools proved to be feasible for building a 3D model of proteins through homology modeling. The generated 3D models demonstrated stability in terms of energy level, enabling the analysis of their tertiary structures and molecular functions.

INTRODUÇÃO

A 1-aminociclopropano-1-carboxilato sintase, também conhecida como ACC sintase, é uma enzima que desempenha um papel significativo na síntese de etileno, um hormônio vegetal produzido em uma grande variedade de plantas superiores (Bakshi et al., 2015). Este hormônio gasoso é um regulador de vários processos fisiológicos e de desenvolvimento, como dormência de sementes, germinação, crescimento vegetativo, floração, amadurecimento climático dos frutos e senescência (Van de Poel et al., 2015). Além disso, demonstrou desempenhar um papel importante na defesa da planta contra fatores de estresse biótico e abiótico (Van de Poel et al., 2015; Wen, 2015).

A importância da ACC sintase no processo de amadurecimento de frutas foi verificada por experimentos utilizando RNA anti-sentido em tomates. Frutos de tomate transgênicos que expressam o RNA anti-sentido da ACC sintase não amadurecem, a menos que expostos ao etileno exógeno (Capitani et al., 1999). A ACC sintase foi identificada pela primeira vez na cultura do tomate (Boller et al., 1979; Kende 1993), onde posteriormente foram encontrados em outras espécies de plantas como abobrinha, *Arabidopsis*, arroz, trigo, abóbora, maçã, batata, pêra, banana (Ge et al., 2000) entre outros.

A ACC sintase pertence à família das enzimas dependentes de piridoxal-5'-fosfato (PLP) e possui similaridade de sequência (441 a 496 aminoácidos) com outros membros dessa família, como aspartato amino transferase (AATase) e tirosina aminotransferase (TATase) (Feng et al., 2000), onde a expressão de cada um ocorre de modo diferente em resposta a diferentes tipos de estímulo como a injúria mecânica, a infecção viral, a injúria pelo frio, a anaerobiose e o tratamento com auxina (Purgatto, 2001).

Com o rápido crescimento das bases de dados de sequências, o número de sequências pertencentes a uma família de proteínas funcionalmente relacionada está aumentando acentuadamente. Como consequência, é cada vez mais necessário analisar as relações existentes entre os numerosos membros de uma família de proteínas (Watson et al., 2005). Para entender o comportamento das proteínas nos grupos de organismos, análises filogenéticas podem auxiliar questões sobre como as proteínas estão relacionadas em diferentes espécies, e se elas podem ter evoluído a partir de um ancestral comum (Kasap et al., 2010). Trabalhos tem sido realizados para elucidar a estrutura, relações evolutivas e filogenéticas da ACC sintase (Jakubowicz et al., 2002), onde atualmente os resultados filogenéticos podem ser continuamente aperfeiçoados devido à crescente disponibilidade de dados biológicos e a novas abordagens para analisá-los (Andrade et al., 2011).

Este trabalho teve como objetivos: (1) Caracterizar, comparar e identificar domínios conservados, em sequências de aminoácidos de uma 1-aminociclopropano-1-carboxilato sintase em espécies de dicotiledôneas disponíveis em bancos de dados públicos. (2) Realizar uma análise filogenética baseada nas sequências encontradas, para compreensão das relações entre espécies das famílias Myrtaceae, Euphorbiaceae, Vitaceae, Leguminosae, Rutaceae, Malvaceae, Cannabaceae, Solanaceae, Rosaceae, Moraceae, Compositae, Fagaceae e Jugladiaceae. (3) Desenvolver modelos tridimensionais da 1-aminociclopropano-1-carboxilato sintase de representantes de dicotiledôneas, com base na metodologia de modelagem por homologia, de forma a possibilitar uma melhor compreensão de suas estruturas e funções moleculares. Esta análise comparativa fornecerá informações teóricas importantes para posteriores análises envolvendo estas proteínas, considerando seu efetivo potencial para uso em estudos filogenéticos, amadurecimento climático dos frutos e na defesa da planta contra fatores de estresse biótico e abiótico.

MATERIAL E MÉTODOS

Recuperação das sequências

As sequências da ACC sintase foram obtidas através do algoritmo de busca BLAST (Basic Local Alignment Search Tool) na plataforma do NCBI (National Center for Biotechnology Information). Sequências de proteínas em espécies de dicotiledôneas foram recuperadas pela ferramenta BLASTp para procurar os homólogos, onde um total de 32 sequências foram encontradas (Tabela 1).

Análise das Sequências

Parâmetros físico-químicos das ACC sintase presentes em espécies das famílias Myrtaceae, Euphorbiaceae, Vitaceae, Leguminosae, Rutaceae, Malvaceae, Cannabaceae, Solanaceae, Rosaceae, Moraceae, Fagaceae, Compositae e Jungladaceae foram analisadas pelo ProtParam (<http://web.expasy.org/protparam>) (Gasteiger et al., 2005). A presença de sítios de clivagem de peptídeos de sinal foi investigada usando o servidor TOPCONS (<http://topcons.cbr.su.se/pred>) (Tsirigos et al., 2015). A identificação dos domínios funcionais da proteína, sua classificação e ontologia foram realizadas com o uso do servidor Prodom (<http://prodom.prabi.fr/prodom/>) (Servant et al., 2002), uma base de dados de família de domínios de proteínas de segmentos homólogos. A estimativa dos efeitos funcionais ocasionados por mutações de sequências de aminoácidos foram obtidas por meio do servidor SNAP2 (<https://roslab.org/services/snap2web/>) (Hecht et al., 2015).

Análise Filogenética e Agrupamento

A predição estrutura terciária dos modelos 3-D das ACC sintase de *P. Guajava* (AFU90953.1) e *G. Max* (XP_003534540.1) foram realizadas pelo servidor Phyre2 (<http://www.sbg.bio. ic.ac.uk/phyre2>) (Kelley et al., 2015) no modo multi-template. As sequências da ACC sintase foram alinhadas utilizando-se o algoritmo ClustalW e as árvores filogenéticas foram produzidas no software MEGA 7.0.21 (Kumar et al., 2016). As árvores filogenéticas foram construídas usando o método de Máxima verossimilhança (MV) com teste de bootstrap com 1000 réplicas.

Previsão, Avaliação e Validação da Estrutura Terciária dos Modelos

A predição estrutura terciária dos modelos 3-D das ACC sintase foram realizadas pelo servidor Phyre2 (<http://www.sbg.bio. ic.ac.uk/phyre2>) (Kelley et al., 2015) no modo multi-template. Este servidor utiliza-se métodos avançados de detecção de homologia para construir um modelo 3D. A estrutura 3D da proteína foi então visualizada pelo pacote UCSF Chimera (Pettersen et al., 2004).

A qualidade do modelo foi avaliada usando o servidor Molprobitry (<http://molprobitry.biochem.duke.edu/>) (Chen et al., 2010) pela análise de amachandran. O Z-score foi calculado utilizando o servidor interativo ProSA-web (<https://prosa.services.come.sbg.ac.at/prosa.php>) para reconhecer erros em estruturas tridimensionais (Wiederstein et al., 2007). Minimização de energia e correção de pequenos erros no modelo tridimensional foram realizados pelo servidor Yasara force Field (Krieger et al., 2009) e pelo software KiNG (Chen et al., 2009).

RESULTADOS E DISCUSSÃO

Neste estudo, analisamos sequências da ACC sintase em espécies de dicotiledôneas, a partir de sequências de proteínas que foram recuperadas do banco de dados do NCBI. As propriedades físico-químicas das sequências de proteína foram analisadas pelo servidor *ProtParam*. A Tabela 1 mostra o peso molecular, o ponto teórico isoelétrico (pI) e a média de hidropaticidade (GRAVY) das sequências proteicas avaliadas.

Tabela1. Análise da estrutura primária e localização subcelular das seqüências de ACC sintase de espécies de Dicotiledôneas

Espécie	GB-ID	Nº aa	PM (kDa)	pl	Gravy	Familia
<i>Psidium guajava</i>	AFU90953.1	491	55.59	6.53	-0.342	Myrtaceae
<i>Rhodamnia argentea</i>	XP_030547978.1	491	55.49	7.01	-0.322	Myrtaceae
<i>Syzygium oleosum</i>	XP_030468262.1	491	55.28	6.81	-0.314	Myrtaceae
<i>Eucalyptus grandis</i>	XP_010055081.1	486	54.71	7.21	-0,302	Myrtaceae
<i>Jatropha curcas</i>	XP_012083309.1	485	54.44	6.51	-0.226	Euphorbiaceae
<i>Vitis vinifera</i>	RVW47060.1	481	53.99	6.81	-0.227	Vitaceae
<i>Hevea brasiliensis</i>	XP_021644099.1	483	54.24	5.74	-0.244	Euphorbiaceae
<i>Manihot esculenta</i>	XP_021612461.1	482	54.09	6.33	-0.233	Euphorbiaceae
<i>Trifolium repens</i>	AJF36210.1 1	484	55.17	7.57	-0.338	Leguminosae
<i>Ricinus communis</i>	XP_002511916.1	482	54.29	8.26	-0.228	Euphorbiaceae
<i>Citrus unshiu</i>	GAY51435.1	477	53.35	6.28	-0.195	Rutaceae
<i>Citrus sinensis</i>	NP_001275873.2	477	53.38	6.28	-0.190	Rutaceae
<i>Gossypium raimondii</i>	XP_012490018.1	493	55.89	8.04	-0.343	Malvaceae
<i>Gossypium hirsutum</i>	NP_001314151.1	493	55.84	8.08	-0.322	Malvaceae
<i>Gossypium australe</i>	>KAA3475241.1	488	55.30	6.48	-0.321	Malvaceae
<i>Arachis hypogaea</i>	XP_025649534.1	485	55.21	8.90	-0.349	Leguminosae
<i>Glycine max</i>	XP_003534540.1	485	54.75	8.41	-0.286	Leguminosae
<i>Cannabis sativa</i>	XP_030487921.1	522	58.04	7.01	-0.248	Cannabaceae
<i>Durio zibethinus</i>	XP_022770501.1	486	55.21	6.58	-0.248	Malvaceae
<i>Cajanus cajan</i>	XP_020220146.1	484	54.61	8.15	-0.277	Leguminosae
<i>Medicago truncatula</i>	XP_003624084.2	484	55.18	7.21	-0.351	Leguminosae
<i>Cicer arietinum</i>	XP_004492797.1	483	55.05	8.20	-0.319	Leguminosae
<i>Vigna unguiculata</i>	XP_027927049.1	484	54.80	6.48	-0.239	Leguminosae
<i>Juglans regia</i>	XP_018841214.1	484	54.21	6.25	-0.180	Juglandaceae
<i>Morus notabilis</i>	XP_010089045.1	486	54.67	6.47	-0.144	Moraceae
<i>Malus domestica</i>	XP_008364813.1	503	56.24	6.05	-0.255	Rosaceae
<i>Nicotiana tabacum</i>	NP_001313190.2	491	55.31	6.40	-0.311	Solanaceae
<i>Herrania umbratica</i>	XP_021281361.1	491	55.42	7.94	-0.246	Malvaceae
<i>Quercus suber</i>	XP_023874217.1	473	53.38	6.57	-0.238	Fagaceae
<i>Helianthus annuus</i>	XP_022022519.1	497	56.12	6.44	-0.265	Compositae
<i>Capsicum baccatum</i>	PHT40845.1	486	54.83	7.92	-0.303	Solanaceae
<i>Abrus precatorius</i>	XP_027352117.1	483	54.71	6.05	-0.307	Leguminosae

GRAVY: Grand Average of Hydropathicity; pl: ponto isoelétrico.

O ponto isoelétrico das seqüências variou entre 5,74 (*Hevea brasiliensis*) e 8.90 (*Arachis hypogaea*), mostrando que provavelmente esta proteína exibe variação de pH para atuação entre as espécies analisadas. O ponto isoelétrico (pl) de uma proteína é normalmente considerado como uma mera propriedade físico-química que resulta das propriedades de carga ambivalente somadas de seus constituintes de aminoácidos (Stekhoven et al., 2008). De fato, os valores de ponto isoelétrico das proteínas nem sempre são registrados nos dados, apesar da estrutura primária da proteína ser totalmente conhecida (Stekhoven et al., 2008). O ponto isoelétrico (pl) é importante para a solubilidade e localização subcelular, revelando ser uma ferramenta poderosa para prever e entender as interações entre proteínas, proteínas e membranas (fosfolipídios) ou para determinar a presença de isoformas proteicas (Stekhoven et al., 2008).

A média de hidropaticidade variou de -0.351 (*Medicago truncatula*) a -0.144 (*Morus notabilis*), o que demonstra que esta proteína é de caráter hidrofílico para todas as espécies. Muitas aplicações importantes do uso da hidropaticidade das estruturas proteicas têm sido relatadas na literatura. Dentre essas, algumas aplicações pioneiras incluem a compreensão da estrutura 3D das proteínas, relacionamento evolutivo e detecção de proteínas análogas e distantes (Lolkema, 1998). O peso molecular variou de 53.35kDa (*Citrus unshiu*) a 58.04kDa (*Cannabis sativa*), onde tem sido utilizado para isolar proteínas por meio de técnicas de eletroforese e cromatografia (Wu, 2003).

O alinhamento múltiplo de seqüências é mostrado na Figura 1. No alinhamento observado, foi possível verificar regiões de pouca similaridade, onde a maior parte das posições apresentou pouca conservação entre as seqüências.

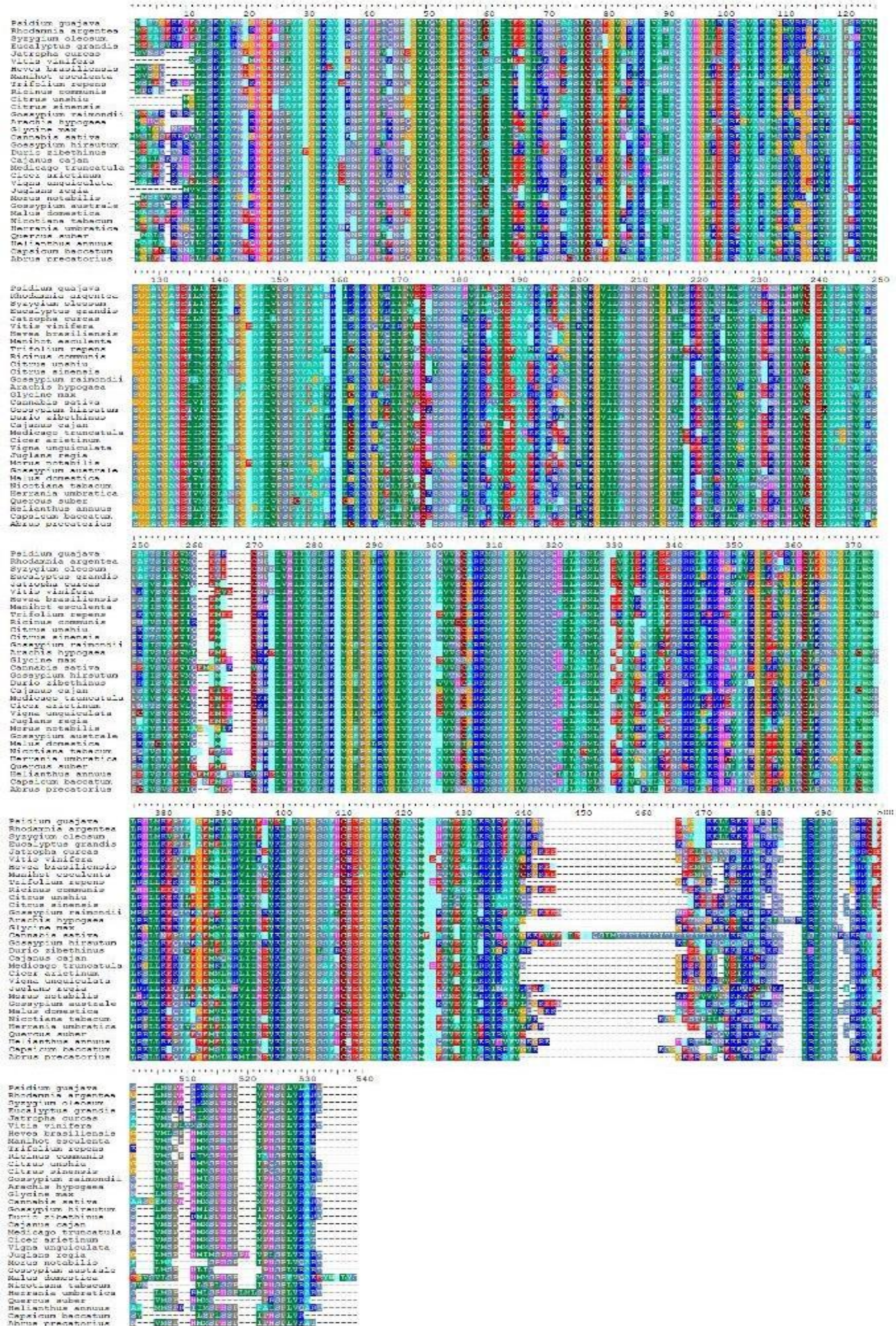


Figura1: Alinhamento de seqüências de ACC sintase avaliadas. As seqüências foram alinhadas por ClustalW, onde os resíduos idênticos e similares são exibidos na mesma cor.

Para a caracterização funcional da ACC sintase e previsão dos efeitos funcionais de mutações, foram selecionadas quatro espécies (Figura 2). Foi identificada a existência de dois domínios funcionais: PD339834 e PD981186 nas sequências de aminoácidos da ACC sintase. O domínio funcional PD339834 foi encontrado em todas as sequências observadas e está associado com funções de síntese da piridoxal-5'-fosfato (PLP), que é um cofator essencial orgânico para a ativação do etileno. Foram encontradas as mesmas informações no servidor sobre as atividades do domínio funcional PD981186.

O servidor SNAP2 revelou que os domínios funcionais PD339834 e PD981186 são as regiões da proteína mais vulneráveis a ocorrência de mutações. Um mapa foi gerado pelo servidor mostrando a presença de colorações diferentes com provável substituição em cada colocação dessas proteínas. A coloração vermelha indica tendência forte para mutação, coloração branca apresenta sinal de baixo efeito e a coloração azul aponta neutralidade. A ocorrência de uma maior quantidade de sítios neutros a mutações no domínio PD339834 ilustra sua conservação e presença em todas as sequências verificadas.

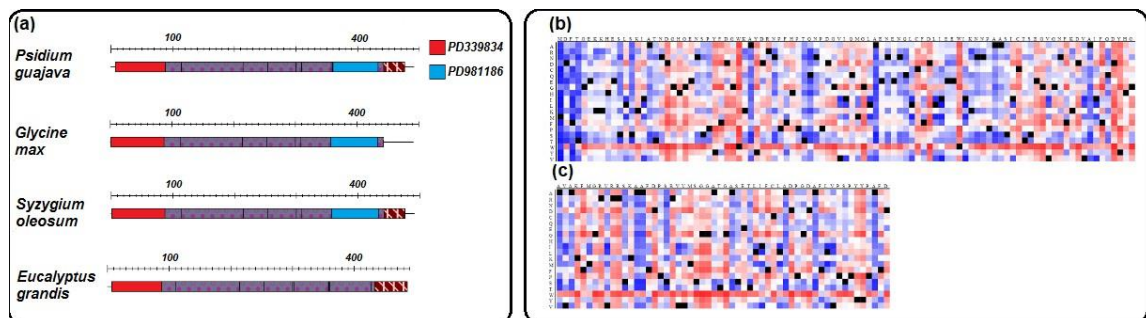


Figura2. Previsão dos efeitos funcionais das mutações em aminoácidos pelo servidor SNAP2 e Domínios funcionais da 1-aminociclopropano-1-carboxilato sintase observados pelo servidor ProDom.

Nas análises filogenéticas houve a formação de dois clados (Figura 3), sendo o primeiro composto pelas famílias Malvaceae (*Gossypium raimondii*, *Gossypium hirsutum*, *Gossypium australe*, *Durio zibethinus* e *Herrania umbratica*) (BS=100%), Rutaceae (*Citrus unshiu* e *Citrus sinensis*) (BS=100%), Euphorbiaceae (*Jatropha curcas*, *Ricinus communis*, *Hevea brasiliensis* e *Manihot esculenta*) (BS=100%), Vitaceae (*Vitis vinífera*), Moraceae (*Morus notabilis*), Cannabaceae (*Cannabis sativa*), Solanaceae (*Nicotiana tabacum* e *Capsicum baccatum*), Compositae (*Helianthus annuus*) e Myrtaceae (*Eucalyptus grandis*, *Syzygium oleosum*, *Rhodamnia argentea* e *Psidium guajava*) (BS=100%) e o segundo clado englobou as famílias Juglandaceae (*Juglans regia*), Fagaceae (*Quercus suber*), Rosaceae (*Malus domestica*) e Leguminosae (*Arachis hypogaea*, *Medicago truncatula*, *Trifolium repens*, *Cicer arietinum*, *Abrus precatorius*, *Cajanus cajan*, *Glycine max* e *Vigna unguiculata*) (BS=100%).

A família Malvaceae (BS=100%) (Figura 3) é constituída por nove subfamílias, onde as espécies *Gossypium raimondii*, *Gossypium hirsutum* e *Gossypium australe* fazem parte da subfamília Malvoideae, *Durio zibethinus* da subfamília Bombacoideae e *Herrania umbratica* da subfamília Byttenerioideae. Embora Malvoideae mostre representantes em um grupo isolado e bem suportado (BS=100%), alguns estudos indicam que as subfamílias Malvoideae e Bombacoideae formam o clado “Malvatheca”, sustentado por sequencias de matK e ndhF e morfologicamente pelas anteras altamente modificadas (Baum et al., 2004).

Na família Rutaceae (BS=100%) (Figura 3), as espécies *Citrus unshiu* e *Citrus sinensis* fazem parte da tribo Citrae, e se configuram por revelar semelhanças entre si, como por exemplo, fraca barreira reprodutiva resultando na geração de híbridos tornando-os taxonomicamente complexos (Guerra et al. 2000).

A família Euphorbiaceae (BS=99%) (Figura 3) engloba quatro subfamílias com fraco suporte filogenético (Stevens, 2014). A subfamília Crotonoideae abrange as espécies *Jatropha curca* e *Manihot esculenta*, e é tipicamente representada por árvores, arbustos e ervas, tendo como propriedade importante à presença de látex em diversos gêneros (Stevens, 2014). São distribuídas pantropicalmente e algumas em regiões temperadas (Stevens, 2014). Acalyphoideae contém *Ricinus communis*, e não é monofiletica com distribuição pantropical (Stevens, 2014). É composta por árvores, arbustos ou ervas e raramente apresentam látex (Stevens, 2014). Euphorbioideae traz *Hevea brasiliensis* como representante, sendo que as principais características que a distingue das demais subfamílias é a presença de látex de coloração branca. É distribuída pantropicalmente e algumas em regiões temperadas (Stevens, 2014).

A família Leguminosae (BS=100%) (Figura 3) é composta por seis subfamílias, nos quais as espécies *Arachis hypogaea*, *Medicago truncatula*, *Trifolium repens*, *Cicer arietinum*, *Abrus precatorius*, *Cajanus cajan*, *Glycine max* e *Vigna unguiculata* fazem parte da subfamília Papilionoideae (Schrire et al., 2005). Análises moleculares indicam que Papilionoideae é monofiletica, com os seus representantes encontrados desde as florestas úmidas e secas até os desertos secos e frios (Schrire et al., 2005). Papilionoideae são reconhecidos e diferenciados das outras subfamílias por possuírem folhas frequentemente pinadas, flores papilionadas com pétalas diferenciadas e sementes com hilo bem delimitado (Lewis et al., 2005).

A família Myrtaceae (BS=100%) (Figura 3) apresenta duas subfamílias, onde as espécies *Psidium guajava*, *Rhodamnia argentea*, *Syzygium oleosum* e *Eucalyptus grandis* fazem parte da subfamília Myrtoideae Sweet com 15 tribos, onde estão inseridos todos os gêneros americanos, com exceção do monoespecífico chileno *Tepualia* Griseb. As espécies da subfamília Myrtoideae são caracterizados por possuírem frutas carnudas (baga ou drupa) e folhas opostas (Wagner et al. 1990; Vasconcelos et al., 2017). Os números de cromossomos diplóides de membros desta subfamília variam de $2n = 22$ a $2n = 88$. A subfamília é mais desenvolvida na América tropical, e os maiores gêneros são *Eugenia* (600 espécies), *Myrcia* (300 espécies), *Syzygium* (200 espécies) e *Psidium* (100 espécies) (Vasconcelos et al., 2017).

A Filogenia em espécies de dicotiledôneas com base em proteínas tem sido amplamente utilizadas para entender a importância das proteínas dentro de grupos de organismos. Análises filogenéticas podem elucidar questões sobre como as proteínas estão relacionadas em diferentes espécies, e se elas podem ter evoluído a partir de um ancestral comum (Van Holle et al., 2017; Moraes Filho et al., 2016). Filogenias baseadas em sequências de ACC sintase demonstram similaridade com a filogenia das angiospermas (Nascimento et al., 2014).

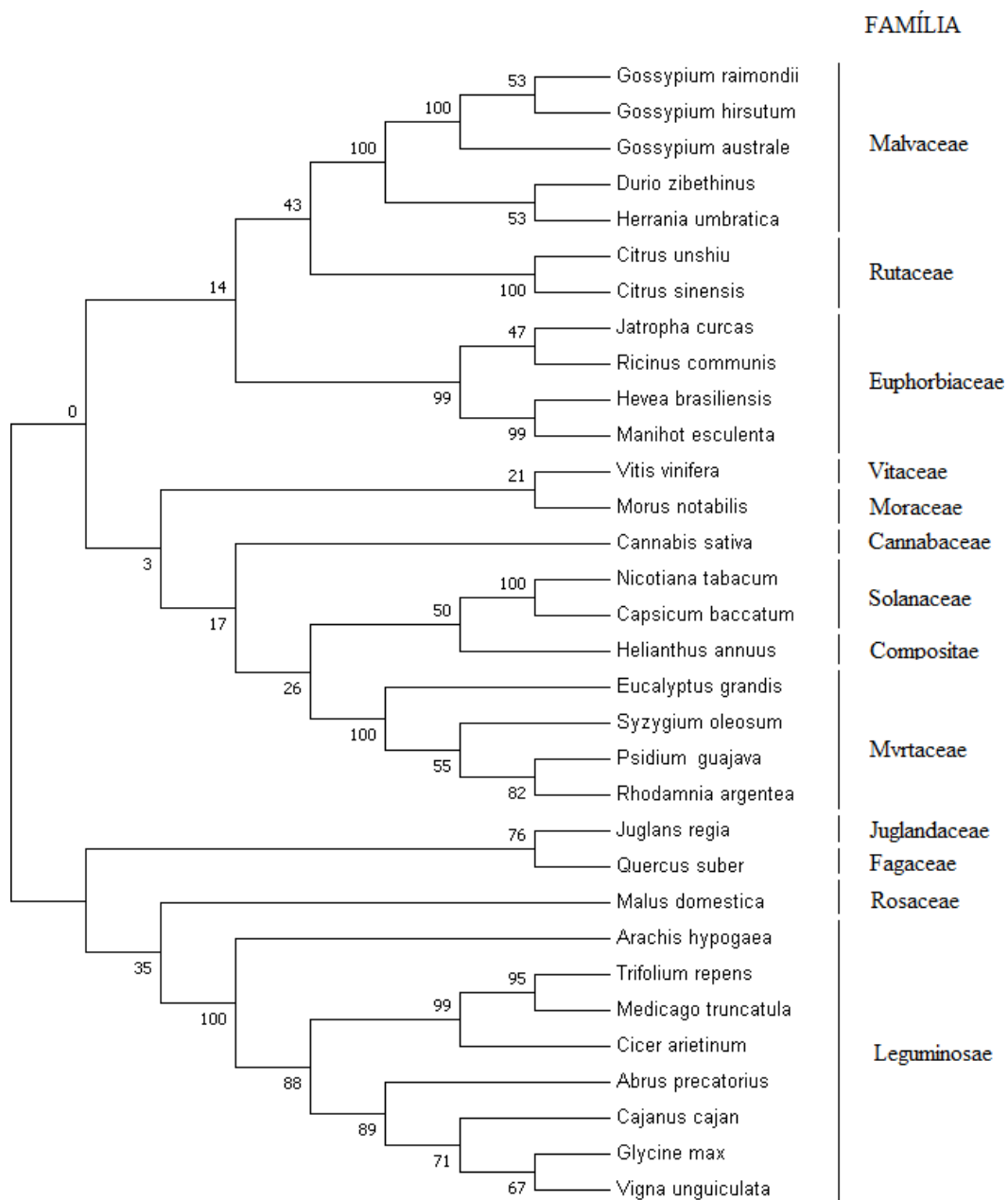


Figura 3. Árvore filogenética das espécies de dicotiledôneas, geradas com base nas seqüências de ACC sintase pelo método máxima verossimilhança (MV).

Para predição da estrutura terciária da proteína pelo servidor Phyre2, foram selecionadas as sequências da lectina das espécies *Glycine max* e *Psidium guajava*, representantes das dicotiledôneas contempladas no banco de dados (Figura 4).

Depois do refinamento das estruturas 3D pelo software KiNG e pelo Servidor Yasara Force Field, os modelos foram examinados pelo servidor MOLPROBITY para a verificação de erros na previsão da estrutura terciária. Os modelos refinados apresentaram 93,75% e 96,51% de resíduos de aminoácidos nas regiões favorecidas por análise do gráfico de Ramachandran para as espécies *Glycine max* e *Psidium guajava* respectivamente. Os valores de Z-score avaliados pelo servidor ProSa foram de -10,53 e -10,46 para *Glycine max* e *Psidium guajava* respectivamente (Figura 5). Os Z-score das estruturas de proteínas são importantes porque caracterizam a energia conformacional das proteínas, e sugerem qualidade geral do modelo com respeito à distribuição de energia derivada de configurações geradas aleatoriamente. Para a avaliação da estrutura de uma proteína, é esperado que o seu valor de Z-Score esteja dentro da variação esperada para proteínas de tamanho similar (regiões em azul).

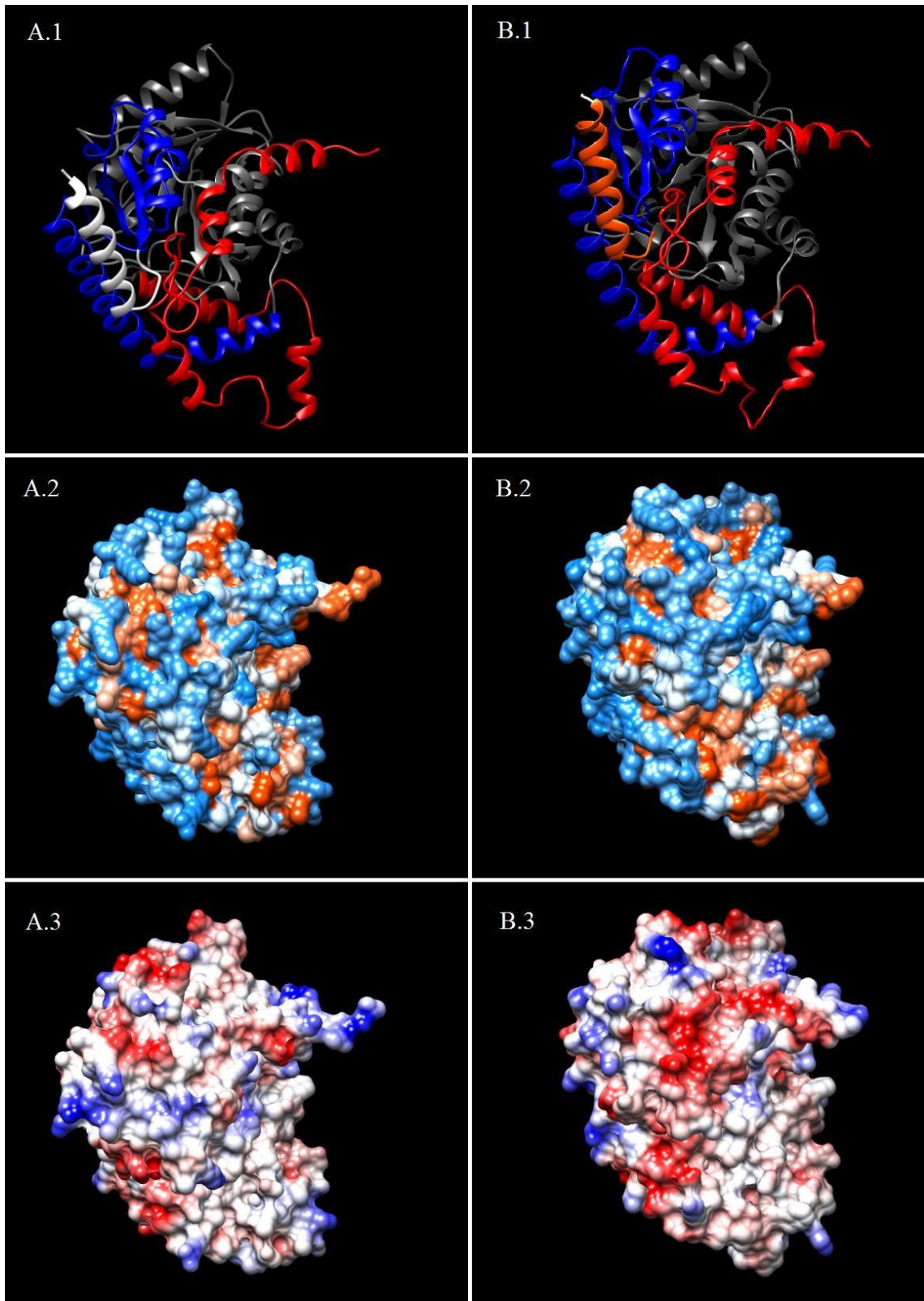


Figura 4. Estrutura tridimensional predita pelo servidor Phyre2 para ACC sintase da espécie *Glycine max* (**A1**) e *Psidium guajava* (**A2**). Hidrofobicidade representada como gradiente de cor, sendo o azul o mais hidrofílico e vermelho alaranjado para os mais hidrofóbicos da espécie *Glycine max* e (**B2**) *Psidium guajava* (**B1**). Superfície eletrostática representada como um gradiente de cor, a partir domais carregado negativamente (vermelho) para o mais positivamente carregado (azul) da espécie *Glycine max* (**A3**) e *Psidium guajava* (**B3**).

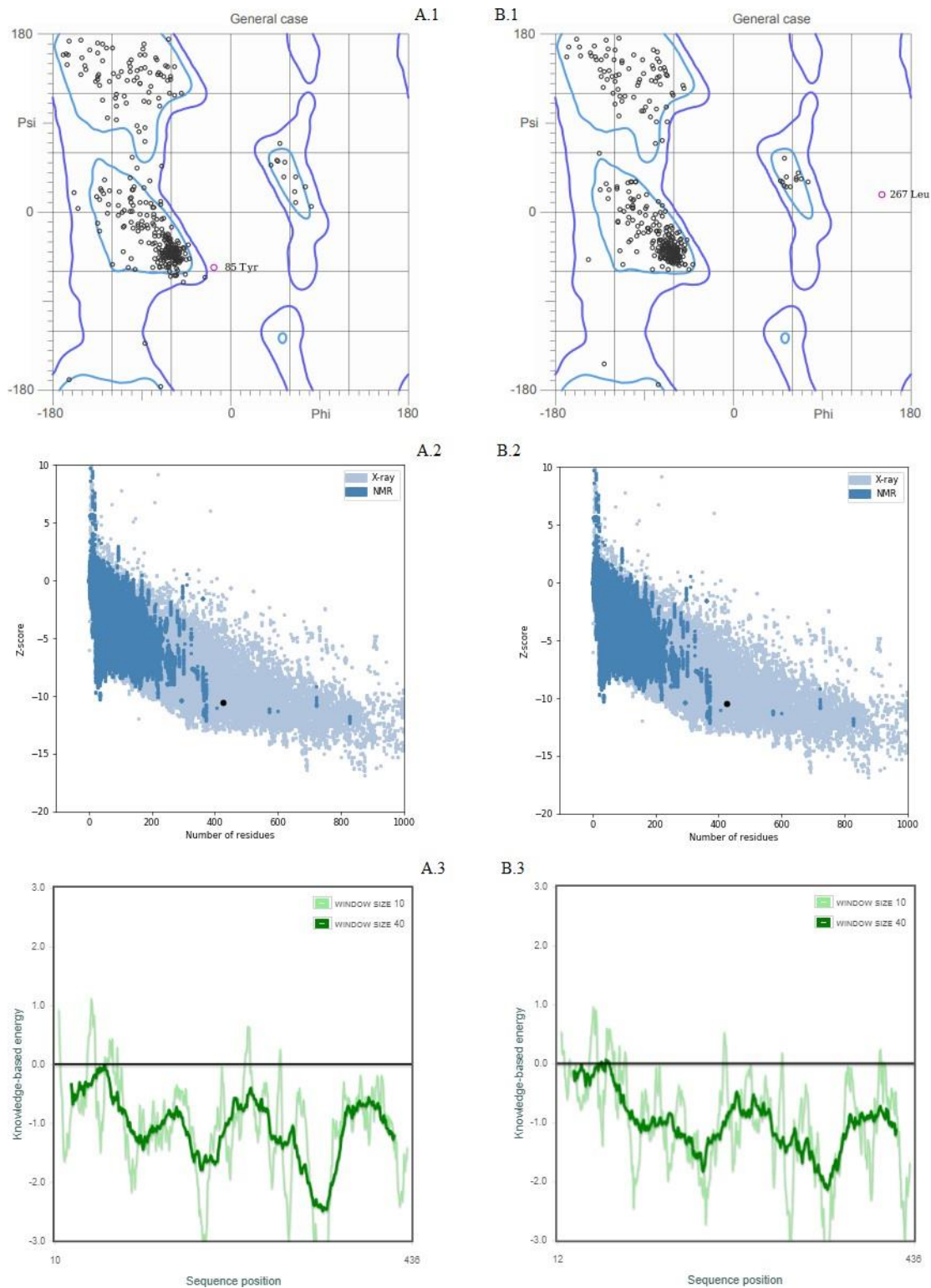


Figura 5. Gráfico de Ramachandran gerado pelo servidor MolProbity, a partir do modelo tridimensional construído pelo servidor Phyre2 para *Glycine max* (A1) e *Psidium guajava* (B1). Valor do Z-score (ponto preto) pelo ProSA-web para *Glycine max* (A2) e *Psidium guajava* (B2). Uso do ProSA-web para 1-aminociclopropano-1-carboxilato sintase mostrando o gráfico de energia de escores de resíduos da estrutura de proteína nativa para *Glycine max* (A3) e *Psidium guajava* (B3).

A ACC sintase realiza um papel fundamental na síntese de etileno (Bakshi et al., 2015), desempenhando uma função importante na regulação do crescimento e processos de desenvolvimento da planta. A via biossintética do etileno envolve a conversão de S-adenosil metionina (AdoMet) em ácido 1-aminociclopropano-1-carboxílico (ACC), que é sintetizado pela ACC sintase (ACS) e depois converte o ACC em etileno pelo ACO (1-aminociclopropano-1- ácido carboxílico oxidase) (Bleecker, 2000; Kende, 1993; Yang & Hoffman, 1984).

Gu et al., (2019) realizando uma análise de expressão de todo o genoma de genes homeobox (*HB*) e PCR quantitativo em tempo real em frutos maduros de pêsego, revelou que gene do pêsego ACC sintase (PpACS1) afeta a produção de etileno, desempenhando um papel importante no desenvolvimento e amadurecimento dos frutos. Ji et al., (2016) utilizando a reação em cadeia da polimerase (PCR) para amplificar as sequências gênicas da ACC sintase de 25 coleções de germoplasma de melancia, mostrou que o gene da melancia ACC sintase 4 (CitACS4) determina a andromonoiccia (forma sexual em que as variedades apresentam flores masculinas e hermafroditas) em melancia, sendo importante para a determinação do sexo em varias espécies de Cucurbitáceas. Com essas informações, a ACC sintase podem ser fortes postulantes para a engenharia de plantas para uso na regulação do nível de biossíntese de etileno, mas também para poder gerar novas moléculas em espécies em que a atividade do etileno é limitada.

A atividade das proteínas está correlacionada com a sua configuração tridimensional, sendo, portanto assim definida por sua sequência de aminoácidos. Embora seja difícil a análise das sequências das proteínas, a modelagem computacional por homologia é uma possibilidade exequível para a geração de modelos 3D de proteínas com arranjos ainda desconhecidos, com base em proteínas homologas com estruturas já determinadas.

Os modelos 3D apresentado neste trabalho revelam poucas diferenças importantes na estrutura da ACC sintase de *Glycine max* e *Psidium guajava* observados na Figura 5. Verifica-se a presença de beta-folhas na região do domínio PD339834 e PD981186 no modelo de *Glycine max* e *Psidium guajava*, embora o domínio PD339834 apresente um comprimento maior em *Glycine max* do que em *Psidium guajava*. Os domínios PD339834 e PD981186 apresentam conformação similar, sugerindo que estes domínios apresentam a mesma função nas duas espécies.

De acordo com o nosso conhecimento, este trabalho apresenta os dois primeiros modelos 3D da ACC sintase para *Glycine max* e *Psidium guajava*. A modelagem 3D e a análise comparativa desta proteína possibilitarão informações importantes na compreensão de sua evolução e funções moleculares nas espécies desse gênero.

CONCLUSÕES

As árvores filogenéticas das espécies de dicotiledôneas construídas por meio da sequência da proteína da ACC sintase apresentam consistência com proposta de classificação apresentada na literatura. No entanto os métodos utilizados se mostraram suficientes para a separação até o nível de tribo dentro das famílias Malvaceae, Rutaceae, Euphorbiaceae, Vitaceae, Moraceae, Cannabaceae, Solanaceae, Compositae, Myrtaceae, Juglandaceae, Fagaceae, Rosaceae e Leguminosae. O uso de ferramentas *in silico* se mostrou viável para construção de um modelo 3D de proteínas por meio de modelagem por homologia. Os modelos 3D gerados demonstraram estabilidade quanto ao nível energético, possibilitando a análise de suas estruturas terciárias e funções moleculares.

AGRADECIMENTOS

"O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. G. E, Santos foi financiado por uma bolsa de mestrado CAPES, Moraes Filho RM foi financiado por uma bolsa de pós doutorado CAPES.

REFERÊNCIAS

- Andrade LF, Nahum LA, Avelar LGA. (2011). Eukaryotic protein kinases (ePKs) of the helminth parasite *Schistosoma mansoni*. *BMC genomics*. 12(1): 215. <https://doi.org/10.1186/1471-2164-12-215>
- Baum DA, Smith, SD, Yen A, Alverston WS, Nyfeler R Whitlock, A. & Oldham, R. Phylogenetic relationship of *Malvatheca* (Bombacoideae and Malvoideae; Malvaceae sensu lato) as inferred from plastid DNA sequences. *Am. J. Bot.* 2004. 91(11): 1863-1871.
- Bakshi A., Shemansky J. M., Chang C. R., Binder B. M. History of research on the plant hormone ethylene. *J. Plant Growth Regul.* 2015. 34 809–827. 10.1007/s00344-015-9522-9
- Bleecker AB, and Kende H. Ethylene: A gaseous signal molecule in plants. *Annu. Rev. Cell Dev. Biol.* 2000. 16, 1–18.
- Boller T, Hener RC, and Kende H. Assay for and enzymatic formation of an ethylene precursor, 1-aminocyclopropane-1-carboxylic acid. *Planta*.1979. 145, 293–303. doi: 10.1007/BF00454455
- Capitani G, Hohenester E, Feng L, Storici P, Kirsch JF, Jansonius J. N. Structure of 1-aminocyclopropane-1-carboxylate synthase, a key enzyme in the biosynthesis of the plant hormone ethylene. *J. Mol. Biol.* 1999. 294 745–756. 10.1006/jmbi.1999.3255
- Chen VB, Arendall WB, Headd JJ, Keedy DA. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 2010. 66: 12-21. <https://doi.org/10.1107/S0907444909042073>.
- Chen VB, Davis IW, Richardson DC. KING (Kinemage, Next Generation): a versatile interactive molecular and scientific visualization program. *ProteinSci.* 2009. 18: 2403-2409. <http://dx.doi.org/10.1002/pro.250>
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S. Protein identification and analysis tools on the ExPASy server. In: *The proteomics protocols handbook* (Walker JM, ed.). 2005. 571-607. <https://doi.org/10.1385/1-59259-890-0:571>
- Ge L, Liu J, Wong W, Hsiao W, Chong K, Xu Z. Identification of a novel multiple environmental factor-responsive 1-aminocyclopropane-1-carboxylate synthase gene, NT-ACS2, from tobacco. *Plant Cell Environ.* 2000. 23(11): 1169-1182.
- Guerra M, Dos Santos KGB, Barros e Silva AE, Ehrendorfer F (2000) Heterochromatin Banding Patterns In Rutaceae-Aurantioideae - A Case Of Parallel Chromosomal Evolution, *American Journal of Botany* 87: 735–747.
- Gu C, Guo ZH, Cheng HY, Zhou YH, Qi KJ, Zhang SA HD-ZIP II HOMEBOX transcription factor, PpHB.G7, mediates ethylene biosynthesis during fruit ripening in peach. *Plant Science*, 2019. V. 278, P:12-19.
- Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variant. *BMC Genomics*. 2015. 16: S1. <https://doi.org/10.1186/1471-2164-16-S8-S1>
- Hui-Chun Wu, Hua-Ming Chen, Chyuan-Yuan Shiau, Free amino acids and peptides as related to antioxidant properties in protein hydrolysates of mackerel (*Scomber austriasicus*), *Food Research International*, V 36: 9–10, p. 949-957, 2003.

Feng L, Geck MK, Eliot AC, Kirsch JF. Aminotransferase activity and bioinformatic analysis of 1-aminocyclopropane-1-carboxylate synthase. 2000. *Biochemistry.*; 39: 15242–9.

Jakubowicz M. Structure, catalytic activity and evolutionary relationships of 1-aminocyclopropane-1-carboxylate synthase, the key enzyme of ethylene synthesis in higher plants. *Acta Biochimica Polonica.* 2002: 757-774

Ji G, Zhang J, Zhang H, Sun H, Gong G, Shi J, Tian S, Guo S, Ren Y, Shen H, Gao J, Xu Y (2016). Mutation in the gene encoding *1-aminocyclopropane-1-carboxylate synthase 4 (CitACS4)* led to andromonoecy in watermelon. *J Integr Plant Biol.* 2016. 58: 762– 765.

Kasap S, Benkrid K. (2010). High performance phylogenetic analysis with maximum parsimony on reconfigurable hardware. *IEEE Transactions on Very Large Scale Integration VLSI Systems.* 99: 1–13. <https://doi.org/10.1109/TVLSI.2009.2039588>

Kelley LA, Mezulis S, Yates CM, Wass MN. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 2015. 10: 845-858. <https://doi.org/10.1038/nprot.2015.053>.

Kende, H. Ethylene biosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 1993. 44, 283–307

Krieger E, Joo K, Lee J, Raman S, et al. Improving physical realism, stereochemistry and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins.* 2009. 77:114-122. <http://doi:10.1002/prot.22570>.

Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol.* 2016. 33(7):1870-4. <https://doi.org/10.1093/molbev/msw054>

Lolkema S, DJ Slotboom. Estimation of structural similarity of membrane proteins by hydrophathy profile alignment. *Mol Membr. Biol.*, 15 (1998), pp. 33-42.

Lewis GP, Schrire BD, Mackinder BA, Lock M. *Legumes of the world.* Kew: Royal Botanic Gardens, 2005. 577p.

Moraes Filho RM, Martins LSS. In silico comparative analysis of tylenchid nematodes pectate lyases. *Genetics and Molecular Research. GMR.* 2016. 15038402. <http://dx.doi.org/10.4238/gmr.15038402>

Nascimento. F. X., Rossi. M. J., Soares. C. R. F. S., McConkey. B. J., Glick. B. R. New Insights into 1-Aminocyclopropane-1-Carboxylate (ACC) Deaminase Phylogeny, Evolution and Ecological Significance. *PLoS One*, 2014, V.9 (6).

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004. 25(13):1605-12. <https://doi.org/10.1002/jcc.20084>

Purgatto, E; et al. Inhibition of B-amylase activity, starch degradation and sucrose formation by indole-3-acetic acid during banana ripening. *Planta.* 2001, v.212, n:5-6, p.823-828.

Yang, S.F., and Hoffman, N.E. Ethylene biosynthesis and its regulation in higher-plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 1984. 35, 155–189.

Schirire BD, Lavin M, Lewis GP. Global distribution patterns of the Leguminosae: Insights from recent phylogenies. In: FRIIS I, Balslev H. Plant diversity and complexity patterns: Local, regional and global dimensions. *Biologiske Skrifter*, 2005, V. 55, P. 375-422.

Stekhoven S, F.M.A.H., Gorissen, M.H.A.G. & Flik, G. The isoelectric point, a key to understanding a variety of biochemical problems: a minireview. *Fish Physiol Biochem*. 2008. 34, 1–8. <https://doi.org/10.1007/s10695-007-9145-6>

Sekeli R, Abdullah J., Namasivayam, P., Muda, P., Bakar, U., Yeong, W., et al. (2014). A interferência do RNA da expressão dos genes da oxidase do ácido 1-aminociclopropano-1-carboxílico (aco1 e aco2) prolonga a vida útil do mamão exotika (*Carica papaya* L.). *Molecules* 19, 8350–8362. doi: 10.3390 / moléculas19068350

Servant F, Bru C, Carrère S, Courcelle E, Gouzy J, et al. ProDom: Automated clustering of homologous domains. *Briefings in Bioinformatics*. 2002. 3: 246-251.

Stevens PF. Angiosperm Phylogeny. Website. Acesso em 2014: (<http://www.mobot.org/MOBOT/research/APweb/>).

Tsirigos KD, Peters C, Shu N, Käll L. The TOPCONS web server for combined membrane protein topology and signal peptide prediction. *Nucleic Acids Res*. 2015. 43: 401-407. <https://doi.org/10.1093/nar/gkv485>.

Wagner, W. L., D. R. Herbst, and S. H. Sohmer (eds.), *Manual of the flowering plants of Hawai'i*. University of Hawai'i Press and Bishop Museum Press, Honolulu. *Eucalyptus*. 1990. Pp. 948-959.

Watson JD, Laskowski RA, Thornton JM. (2005). Predicting protein function from sequence and structural data. *Curr. Opin.Struct.Biol*. 15: 275–284.

Wiederstein M, Sippl M J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids*. 2007. 35:407-410. <https://doi.org/10.1093/nar/gkm290>.

Wen, C.-K. (ed.). *Ethylene in Plants*. Dordrecht: Springer. 2015. doi: 10.1007/978-94-017-9484-8

Wu, Y.-T., Huang, W.-Y., Lin, T.-C. and Sheu, S.-J. Determination of moutan tannins by high-performance liquid chromatography and capillary electrophoresis. *J. Sep. Science*. 2003. 26: 1629-1634. doi:10.1002/jssc.200301587

Van de Poel, B., Smet, D., and Van Der Straeten, D. Ethylene and hormonal cross talk in vegetative growth and development. *Plant Physiol*. 2015. 169, 61–72. doi: 10.1104/pp.15.00724

Van Holle S, De Schutter K, Eggermont L, Tsaneva M, Dang L, Van Damme EJM. Comparative Study of Lectin Domains in Model Species: New Insights into Evolutionary Dynamics. Komatsu S, ed. *International Journal of Molecular Sciences*. 2017. 18(6):1136. <https://doi:10.3390/ijms18061136>.

Vasconcelos T. N.C., Proença C. E.B., Ahmad. B., et al. Myrteae phylogeny, calibration, biogeography and diversification patterns: Increased understanding in the most species rich tribe of Myrtaceae, *Molecular Phylogenetics and Evolution*. 2017, V. 109, P. 113-137.