



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE  
PRÓ-REITORIA DE PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS FARMACÊUTICAS

DAISY SOTERO CHACON

**Avaliação do transcriptoma e perfil metabólico na investigação da biossíntese de alcaloides de  
*Erythrina velutina* Willd.**

NATAL - RN

2020

DAISY SOTERO CHACON

**Avaliação do transcriptoma e perfil metabólico na investigação da biossíntese de alcaloides de  
*Erythrina velutina* Willd.**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas da Universidade Federal do Rio Grande do Norte, como parte dos requisitos para a obtenção do título de mestre em Ciências Farmacêuticas.

Orientadora: Profa. Dra. Raquel Brandt Giordani

NATAL - RN

2020

Universidade Federal do Rio Grande do Norte - UFRN  
Sistema de Bibliotecas - SISBI  
Catalogação de Publicação na Fonte. UFRN - Biblioteca Setorial do Centro Ciências da Saúde - CCS

Chacon, Daisy Sotero.

Avaliação do transcriptoma, proteoma e perfil metabólico na investigação da biossíntese de alcaloides de *Erythrina velutina* Willd / Daisy Sotero Chacon. - 2020.

145f.: il.

Dissertação (Mestrado) - Universidade Federal do Rio Grande do Norte, Centro de Ciências da Saúde, Programa de Pós-graduação em Ciências Farmacêuticas. Natal, RN, 2020.

Orientadora: Raquel Brandt Giordani.

1. *Erythrina* - Dissertação. 2. Transcriptoma de *Erythrina velutina* - Dissertação. 3. Biossíntese de alcaloides benzilisoquinolina - Dissertação. 4. Bioma Caatinga - Dissertação. 5. Perfil metabólico direcionado - Dissertação. I. Giordani, Raquel Brandt. II. Título.

RN/UF/BS-CCS

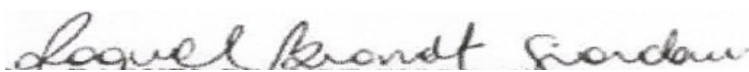
CDU 582.736.3

DAISY SOTERO CHACON

**Avaliação do transcriptoma e perfil metabólico na investigação da biossíntese de alcaloides de  
*Erythrina velutina* Willd.**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas da Universidade Federal do Rio Grande do Norte, como parte dos requisitos para a obtenção do título de mestre em Ciências Farmacêuticas.

**Orientadora:** Profa. Dra. Raquel Brandt Giordani



Presidente: Raquel Brandt Giordani, UFRN



Examinador externo: Maysa Furlan, UNESP



Examinador interno: Leandro de Santis Ferreira, UFRN

## AGRADECIMENTOS

A Deus, por me ter dado a capacidade de realizar esse trabalho;

A CAPES, pela bolsa concedida o que tornou possível minha dedicação exclusiva;

A UFRN, por todo aprendizado pessoal e científico que foi proporcionado;

Ao instituto Serrapilheira, por todo suporte financeiro que tornou viável a realização do trabalho;

A minha orientadora, Profa. Dra. Raquel Giordani, pelo incentivo em todas as etapas de minha vida na universidade. Sua dedicação, humildade, ética e bondade é admirável e tornou possível esse trabalho, és um grande exemplo de professora, pesquisadora e ser humano. Obrigada pela amizade, por ser essa mãe que és, por acreditar em mim, por corrigir quando estou errada e pela confiança de todos esses anos;

Ao Professor Taffarel Melo de Torres, pela dedicação, grande paciência e por compartilhar tudo que sei acerca da bioinformática;

A todos os colaboradores do Centro Nacional de Pesquisa em Energia e Materiais (CNPEM); Centro de Facilidades de Apoio à Pesquisa (CEFAP-USP); Departamento de biociências (UFERSA); Laboratório de Farmacognosia (GPNBio-UFRN); Laboratório de Fisiologia Vegetal (UFRGS); Instituto de Química (UNESP-Araraquara); Laboratório de bioquímica e química de proteínas (UnB); Departamento de Biologia Celular e Genética (UFRN);

A minha família, colegas e amigos que contribuíram até aqui.

## RESUMO

As plantas pertencentes ao gênero *Erythrina* são fontes de alcaloides tetracíclicos do tipo benzilisoquinolínicos, podem ser encontradas no nordeste brasileiro, no bioma da Caatinga e utilizadas na medicina popular principalmente como depressor do sistema nervoso central (sedativo, calmante). Um dos fatores limitantes que muitos produtos naturais apresentam como medicamentos é seu baixo rendimento e difícil obtenção em quantidades suficientes para abastecer a indústria farmacêutica. Dessa forma, de acordo com estudos de fronteira na área é importante compreender aspectos da biossíntese de alcaloides por meio de ferramentas moleculares. Este trabalho tem como objetivo contribuir com o estudo na área de biossíntese no que se refere aos alcaloides presentes nas folhas e sementes da espécie *Erythrina velutina* Willd através das tecnologias “ômicas” (transcriptoma e perfil metabólico) de forma a compreender a dinâmica de produção dessas moléculas com potencial bioativo. Portanto, para a investigação do transcriptoma realizamos o sequenciamento *de novo* pela primeira vez em uma espécie do gênero e em uma planta nativa do bioma Caatinga com essa abordagem. A análise do perfil metabólico direcionado foram realizados com uma estratégia de Cromatografia Líquida acoplada à Espectrometria de Massas [LC-MS/MS] aliado à ferramenta do *Molecular Networking* para os metabólitos secundários. Como resultado, foram identificados 17 transcritos candidatos a compor a via biossintética dos alcaloides isoquinolínicos em sementes e folhas desta espécie. Além disso, 25 alcaloides foram anotados pertencendo às duas classes de alcaloides já reportadas para o gênero: dienoide e alcenoide, sendo 8 alcaloides comum a folhas e sementes, 8 encontrados exclusivamente em semente e 9 exclusivos em folhas. Esses dados contribuem com o conhecimento científico no delineamento de estratégias de intervenção biotecnológica para estudos futuros, a fim de melhorar a produção dessas moléculas bioativas na planta, além da valorização de uma planta da Caatinga do RN como fonte de alcaloides com potencial interesse farmacêutico.

**Palavras-Chave:** Transcriptoma de *Erythrina velutina*; Análise multi-ômicas; Biossíntese de alcaloides benzilisoquinolina; Bioma Caatinga; Perfil metabólico direcionado.

## ABSTRACT

Plants belonging to the genus *Erythrina* are sources of tetracyclic alkaloids of the benzyloquinoline type, can be found in northeastern Brazil, in the Caatinga biome and used in folk medicine mainly as a central nervous system depressant (sedative, sedative). One of the limiting factors that many natural products present as medicines is their low yield and difficult to obtain in sufficient quantities to supply the pharmaceutical industry. Thus, according to border studies in the area, it is important to understand aspects of alkaloids biosynthesis through molecular tools. This work aims to contribute to the study in the area of biosynthesis with regard to the alkaloids present in the leaves and seeds of the species *Erythrina velutina* Willd through “omics” technologies (transcriptome and metabolic profile) in order to understand the dynamics of production of these molecules with bioactive potential. Therefore, for the investigation of the transcriptome we performed the sequencing again for the first time on a species of the genus and on a plant native to the Caatinga biome with this approach. The analysis of the targeted metabolic profile were performed with a Liquid Chromatography strategy coupled to Mass Spectrometry [LC-MS/MS] combined with the Molecular Networking tool for secondary metabolites. As a result, 17 transcripts were identified to form the biosynthetic pathway of isoquinoline alkaloids in seeds and leaves of this species. In addition, 25 alkaloids were noted belonging to the two classes of alkaloids already reported for the genus: dienoid and alkenoid, with 8 alkaloids common to leaves and seeds, 8 found exclusively in seeds and 9 exclusively in leaves. These data contribute to scientific knowledge in the design of biotechnological intervention strategies for future studies, in order to improve the production of these bioactive molecules in the plant, in addition to the enhancement of a plant in the Caatinga of RN as a source of alkaloids with potential pharmaceutical interest.

**Keywords:** *Erythrina velutina*; Multi-omics analysis; Transcriptome; Targeted metabolite profile; Biosynthesis of Benzyloquinoline Alkaloids; Caatinga Biome.

## LISTA DE FIGURAS

<b>Figura 1</b> - Visão e detalhes morfológicos da espécie <i>Erythrina velutina</i> .....	15
<b>Figura 2</b> - Via do chiquimato.....	17
<b>Figura 3</b> - Linha do tempo da evolução da informação genética.....	19
<b>Figura 4</b> - Mapa de localização com pontos de coleta.....	27
<b>Figura 5</b> - Estratégia experimental e integração das análises ômicas.....	34
<b>Figura 6</b> - Mapa do Rio Grande do Norte (Brasil) e características climáticas monitoradas.....	38
<b>Figura 7</b> - Distribuição dos alcaloides e rede Molecular MS/MS obtido a partir de extratos bruto de <i>Erythrina velutina</i> .....	43
<b>Figura 8</b> - Características químicas estruturais dos alcaloides do cluster 1 identificados em amostras de folhas e sementes com análise de LC-MS/MS (modo de ionização positiva para amostras de sementes e folhas de <i>Erythrina velutina</i> ).....	45
<b>Figura 9</b> - Características químicas estruturais dos alcaloides do cluster 2 identificados em amostras de folhas e sementes com análise de LC-MS/MS (modo de ionização positiva para amostras de sementes e folhas de <i>Erythrina velutina</i> ).....	46
<b>Figura 10</b> - Cromatograma dos extratos brutos de folhas e sementes de <i>Erythrina velutina</i> .....	48
<b>Figura 11</b> - Resultado da avaliação do BUSCOv3.....	52
<b>Figura 12</b> - Quantidade de transcrito baseado em domínio e sua respectiva proteína.....	57
<b>Figura 13</b> - Primeiras etapas da biossíntese dos alcaloides eritrínicos.....	60
<b>Figura 14</b> - Etapa intermediária 1 da biossíntese dos alcaloides eritrínicos.....	63
<b>Figura 15</b> - Etapa intermediária 2 da biossíntese dos alcaloides eritrínicos.....	65
<b>Figura 16</b> - Etapa intermediária 3 da biossíntese dos alcaloides eritrínicos.....	66
<b>Figura 17</b> - Etapa final da biossíntese dos alcaloides eritrínicos.....	67
<b>Figura 18</b> - Mapa de calor dos transcritos identificados a participarem da via de biossíntese dos alcaloides eritrínicos.....	68
<b>Figura 19</b> - Proposta da rota de biossíntese para alcaloides de <i>Erythrina</i> .....	70
<b>Figura 20</b> - Resumo dos resultados obtidos.....	73

## LISTA DE TABELAS

<b>Tabela 1</b> - Estrutura básica e subgrupos dos alcaloides presentes em <i>Erythrina</i> spp .....	14
<b>Tabela 2</b> - Solo coletado de aproximadamente 0 a 20 cm de camada superficial em áreas da vegetação da Caatinga, na estação seca .....	39
<b>Tabela 3</b> - Visão geral da análise e montagem do transcriptoma .....	51
<b>Tabela 4</b> - Transcritos e seus respectivos valores de identidade identificados por BLASTx .....	58
<b>Tabela 5</b> - Transcritos identificados de acordo com o nº EC .....	58

## LISTA DE ABREVIATURAS E SIGLAS

SNC – Sistema nervoso central  
AchE – Acetilcolinesterase  
BIAs – Alcaloides benzilisoquinolínicos  
EPSP – 5-enolpiruvilchiquimato-3-fosfato  
EPSP sintase - 5-enolpiruvilchiquimato-3-fosfato sintase  
PEP – Fosfoenolpiruvato  
DNA - Ácido desoxirribonucleico  
RNA - Ácido ribonucleico  
RNAm – RNA mensageiro  
RNAr – RNA ribossômico  
RNAt – RNA transportador  
NGS – Sequenciamento de próxima geração  
cDNA – DNA complementar  
MS – Espectrometria de massas  
GNPS – Rede molecular social natural global  
Da – Dalton  
RIN – Integridade do RNA  
pb – pares de base  
TPM – Transcrito por milhão  
EC – Número de comissão de enzimas  
IUBMB – União internacional de bioquímica e biologia molecular  
 $m/z$  – Relação massa/carga  
mg – Miligrama  
 $\mu\text{L}$  – Microlitro  
kHz – Hertz  
mL – Mililitros  
L – Litros  
 $\mu\text{m}$  - Micrometro

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>11</b>
<b>2 REVISÃO BIBLIOGRÁFICA</b> .....	<b>13</b>
2.1 Da família à espécie <i>Erythrina velutina</i> Willd .....	13
2.2 Biossíntese de alcaloides benzilisoquinolínicos .....	16
2.3 Análises “ômicas” .....	18
2.3.1 Análise do transcriptoma .....	18
2.3.2 Análise metabolômica .....	23
<b>3 OBJETIVOS</b> .....	<b>26</b>
3.1 OBJETIVO GERAL .....	26
3.2 OBJETIVOS ESPECÍFICOS .....	26
<b>4. METODOLOGIA</b> .....	<b>27</b>
4.1 Material vegetal .....	27
4.2 Dados edafoclimáticos e análise do solo .....	29
4.3 Análise do perfil metabólico direcionado para alcaloides eritrínicos .....	30
4.3.1 Extração de metabólitos .....	30
4.3.2 Análise por LC/MS e DAD .....	30
4.3.3 Análise da rede molecular MS/MS ( <i>Molecular Networking</i> ) .....	30
4.4 Análise do transcriptoma .....	31
4.4.1 Análise de RNA-Seq .....	33
4.4.1.1 Extração do RNA total e preparação da biblioteca de cDNA .....	33
4.4.1.2 Análise e anotação do RNAseq .....	33
<b>5 RESULTADO E DISCUSSÃO</b> .....	<b>37</b>
5.1 Dados edafoclimáticos e análise de solos coletados .....	37
5.2 Análise do perfil metabólico .....	41
5.3 Análise do transcriptoma .....	50
5.3.1 Montagem <i>de novo</i> do transcriptoma .....	50
5.3.2 Proposta de via de produção dos alcaloides eritrínicos .....	54
5.3.2.1 Plantas produtoras de BIAS consideradas como referência para este estudo .....	54
5.3.2.2 Prováveis genes e proteínas envolvidos na biossíntese de alcaloides eritrínicos .....	55
<b>6 CONCLUSÃO</b> .....	<b>75</b>
<b>REFERÊNCIA BIBLIOGRÁFICA</b> .....	<b>77</b>

## 1 INTRODUÇÃO

No Brasil, as pesquisas para descoberta de protótipos de fármacos e/ou fitofármacos, além de propiciar o avanço da pesquisa básica multidisciplinar, podem contribuir também para o desenvolvimento tecnológico nacional, levando em conta que a diversidade micromolecular dos inúmeros biomas brasileiros é ainda muito pouco explorada como uma fonte de substâncias de interesse farmacológico (BARREIRO; BOLZANI, 2009; BRAZ FILHO, 2010). Esse contexto se aplica ao potencial da biodiversidade presente no Nordeste do Brasil, o qual pode ser considerado um celeiro de novas moléculas bioativas. A Caatinga é o único bioma exclusivamente brasileiro, ocupa 11% do território nacional e abriga fauna e flora únicas; já foram identificadas cerca de 1500 espécies vegetais, embora se estima que possa chegar a 3000 espécies (PRADO, 2003). O estudo e a conservação da biodiversidade da Caatinga constituem um dos maiores desafios do conhecimento científico brasileiro visto que é, proporcionalmente, a região menos estudada, além de ser a menos protegida.

As plantas são fonte de uma variedade de substância, dentre estas os metabólitos secundários (ou, metabólitos especializados) ocupam um papel de grande destaque, sobretudo quando consideramos aos alcaloides. Os alcaloides bioativos de fonte vegetal têm sua importância tradicionalmente conhecida no contexto das Ciências Farmacêuticas. Com o advento das estratégias “ômicas” de estudo destaca-se a importância de conhecer como a planta produz um alcaloide bioativo através dos genes e proteínas envolvidos em sua biossíntese e isso abre a nova perspectiva de modular e otimizar a produção em larga escala dessa molécula de interesse. Novas informações sobre genes e suas funções, em parte, são responsáveis por uma revolução tecnológica dirigida para as ciências da vida. Diversas abordagens são utilizadas na pesquisa em química de produtos naturais, em decorrência de sua extensão e longa história de desenvolvimento, que caracterizam esta ciência de ampla interdisciplinaridade (BORGES; BERLINCK; SCOTTI; VIEIRA, 2017).

Autores sugerem que os alcaloides do gênero *Erythrina* atuam no sistema nervoso central (SNC) por alteração da neurotransmissão do glutamato (FAGGION et al., 2011) ou alteração colinérgica (DECKER et al., 1995; SETTI-PERDIGÃO et al., 2013). Eles apresentam efeitos anticonvulsivantes e ansiolíticos (SANTOS ROSA et al., 2012) e em especial os alcaloides isoquinolínicos em diferentes famílias demonstraram atividade como inibidores da acetilcolinesterase (AChE) (LÓPEZ et al., 2002; PAGLIOSA et al., 2010).

As técnicas para o estudo da biossíntese de alcaloides estão evoluindo drasticamente nos últimos anos, a investigação das etapas ainda não caracterizadas podem fornecer

informações importantes que permitam uma maior produção desses metabólitos de maneira induzida por diferentes ferramentas biotecnológicas, como por exemplo, por plantas ou organismos modelos modificados geneticamente, abrindo perspectivas para uma variedade de estudos, desde a caracterização destas moléculas, como suas potenciais atividades biológicas.

Nesse contexto, este trabalho contempla como foco central a análise de dados oriundos do transcriptoma, proteoma e perfil metabólico direcionado no que se refere em especial, à biossíntese de alcaloides benzilisoquinolínicos de folhas e sementes de *Erythrina velutina* coletadas em habitat natural da Caatinga com o intuito de contribuir para a caracterização da produção desses peculiares alcaloides em um importante bioma exclusivamente brasileiro.

## 2 REVISÃO BIBLIOGRÁFICA

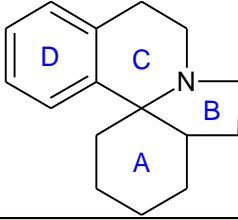
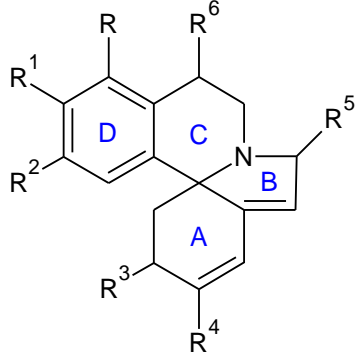
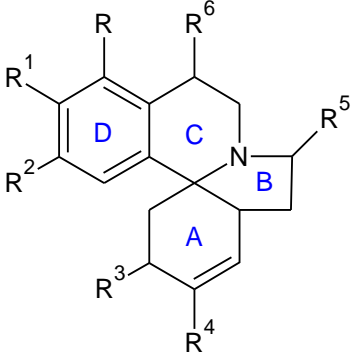
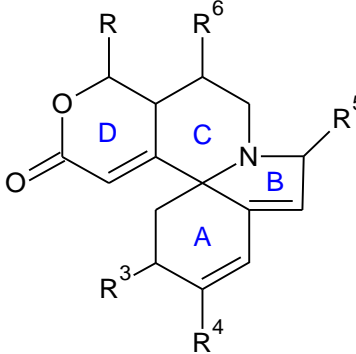
### 2.1 Da família à espécie *Erythrina velutina* Willd

A família Fabaceae ou Leguminosae é considerada a terceira maior família de angiospermas, constituída por 751 gêneros, contendo aproximadamente 19.500 espécies com distribuição global e alto nível ecológico (LPWG et al., 2013), além de estar incluída entre as principais famílias vegetais fixadoras de nitrogênio (WINK, 2013). Ainda a respeito desta família, ela é dividida em três grandes subfamílias tradicionalmente descritas, são elas: Caesalpinioideae, Mimosoideae e Papilionoideae (OLIVEIRA; IWAZAKI; OLIVEIRA, 2014), esta última considerada a maior com 478 gêneros e 13.800 espécies, incluindo as espécies de importância econômica (LPWG et al., 2017), como *Erythrina velutina*.

O gênero *Erythrina* compreende cerca de 110 espécies de árvores e arbustos, sendo considerada uma potente fonte de constituintes químicos com propriedades medicinais (HUSSAIN et al., 2016). O nome tem uma origem grega da palavra “eritros”, que significa cor vermelha, referindo-se à cor das flores (GLEDHILL, 2008). O gênero tem sido extensivamente estudado desde 1990 (RAMBO et al., 2019). Estudos realizados em diferentes partes vegetais demonstraram que os alcaloides de *Erythrina* spp. têm potencial para atuar no sistema nervoso central, com propriedades sedativa, anticonvulsivante, neuroprotetora e ansiolítica (DANTAS et al., 2004; RAUPP et al., 2008; RODRIGUES et al., 2017; VASCONCELOS et al., 2007).

De maneira geral, o gênero apresenta em sua composição química uma ampla diversidade de metabólitos secundários. Estudos na área da fitoquímica demonstram vários metabólitos bioativos em *Erythrina* spp, tais como flavonas, flavonoides, saponinas, taninos e triterpenoides (CARVALHO et al., 2009; DEVI; MANOHARAN, 2011), entretanto, os alcaloides tetracíclicos são os principais marcadores químicos deste gênero (PARSONS; PALFRAMAN, 2010). Os alcaloides eritrínicos tetracíclicos são formados por uma espiroamina tetracíclica, subdividem-se em três categorias: os dienoides, onde apresentam um sistema diênico nos anéis A e B, os alcenoides, que possuem uma dupla ligação geralmente no anel A e o terceiro grupo representado por uma lactona do anel D (AMER; SHAMMA; FREYER, 1991), conforme é apresentado na Tabela 1.

**Tabela 1** - Estrutura básica e subgrupos dos alcaloides presentes em *Erythrina* spp

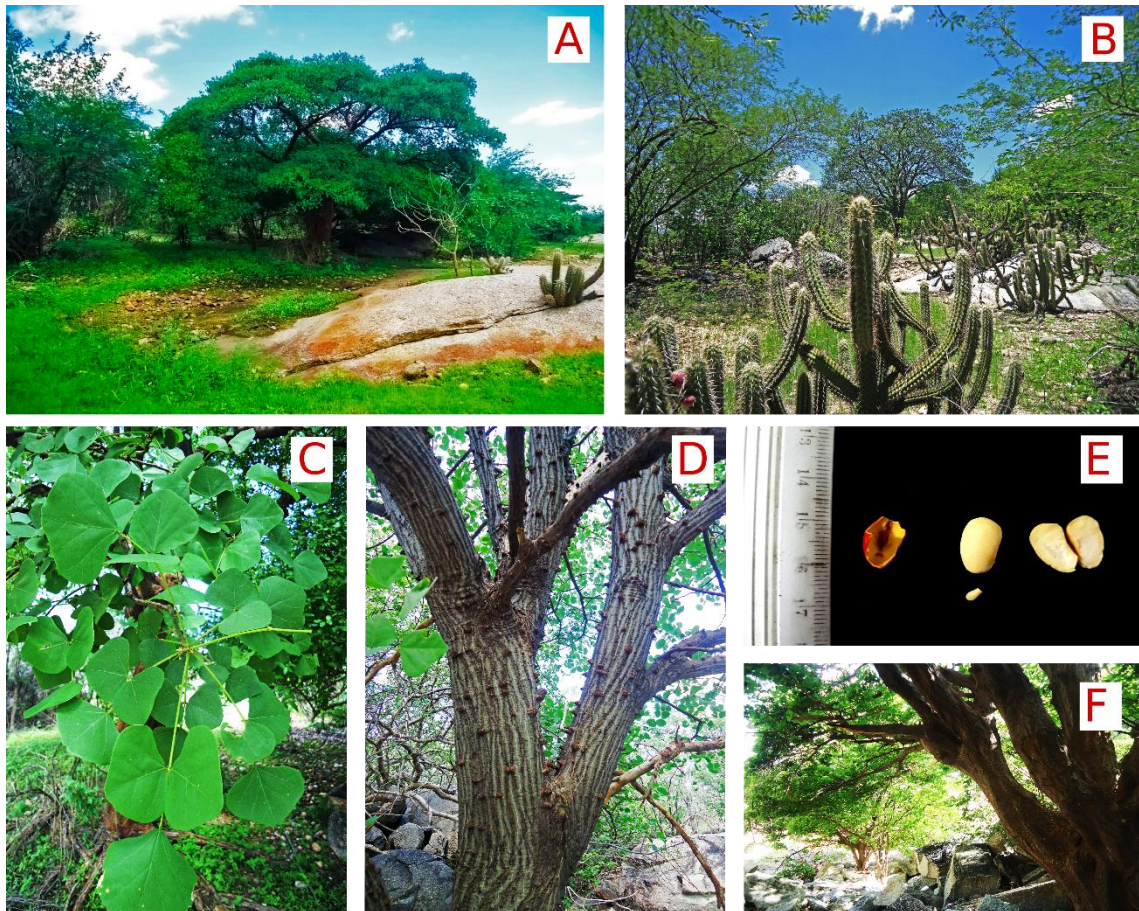
Estrutura química dos alcaloides eritrínicos		
		
Classe dienoide	Classe alcenoide	Classe lactona
		

Os alcaloides eritrínicos são característicos do gênero *Erythrina* e apresentam em suas estruturas diferentes substituintes e ligações em posições específicas, o que os torna diversificados estruturalmente. R a R6: diferentes substituintes na estrutura dos alcaloides eritrínicos; A-B-C-D: Letras referentes aos anéis presente na estrutura básica dos alcaloides. Fonte: elaborado pelo autor (2020).

No que se refere a *Erythrina velutina*, esta é uma espécie localizada na região costeira de todos os estados do norte e nordeste do Brasil (DANTAS et al., 2004), encontrada na Caatinga que é a principal fitofisionomia que ocorre na região semiárida brasileira, considerada a maior floresta tropical seca da América do Sul ocupando mais de 70% da região nordeste e a (RODRIGUES et al., 2018; CARVALHO, 2008).

Conhecida popularmente como bucaré, mulungu e mulungu-da-flor-vermelha, as árvores podem chegar próximas a 15 m de altura na idade adulta (Figura 1), o tronco é reto a levemente tortuoso, as folhas são trifoliadas, os folíolos são orbiculares e de cor verde mais clara e sementes são bicolors, denominadas miméticas, de coloração vermelho-escura e vermelho-alaranjada (CARVALHO, 2008).

**Figura 1** - Visão e detalhes morfológicos da espécie *Erythrina velutina*



A – Planta adulta; B- detalhes da Caatinga e um dos locais próximos da coleta; C - folhas; D – caule com detalhes espinhosos; E – partes constituintes da semente: casca, embrião e cotilédone; F – Tronco. Fonte: elaborado pelo autor (2020).

Com o intuito de estudar os efeitos do extrato de *Erythrina velutina*, alguns estudos demonstram a presença de alcaloides com atividade antioxidante, anticolinesterásica e antinociceptivas nas folhas (MARCHIORO et al., 2005; SANTOS ROSA et al., 2012), além de reversão de sintomas semelhantes à esquizofrenia (XIMENES et al., 2018), neuroproteção (RODRIGUES et al., 2017) e atividade ansiolítica na casca de caule (RAUPP et al., 2008). Da mesma forma, há estudos que abordam a importância do processo de descoberta de fármacos a partir de produtos naturais, pela identificação e isolamento de compostos bioativos do gênero *Erythrina* (FLAUSINO et al., 2007; OZAWA et al., 2009; SANTOS ROSA et al., 2012).

Com certo destaque, as principais espécies do gênero *Erythrina* têm atraído interesses no país, sobretudo pelas características farmacológicas e químicas que elas apresentam. A *Erythrina velutina*, foco deste trabalho, é a única espécie do gênero localizada no bioma da Caatinga, o que faz intensificar ainda mais a importância deste bioma como fonte de moléculas bioativas. No momento, a Farmacopeia Brasileira está desenvolvendo parâmetros para balizar

o controle da qualidade de *E. velutina* em produtos farmacêuticos e medicamentos visando a produção de uma monografia oficial da espécie.

## 2.2 Biossíntese de alcaloides benzilisoquinolínicos

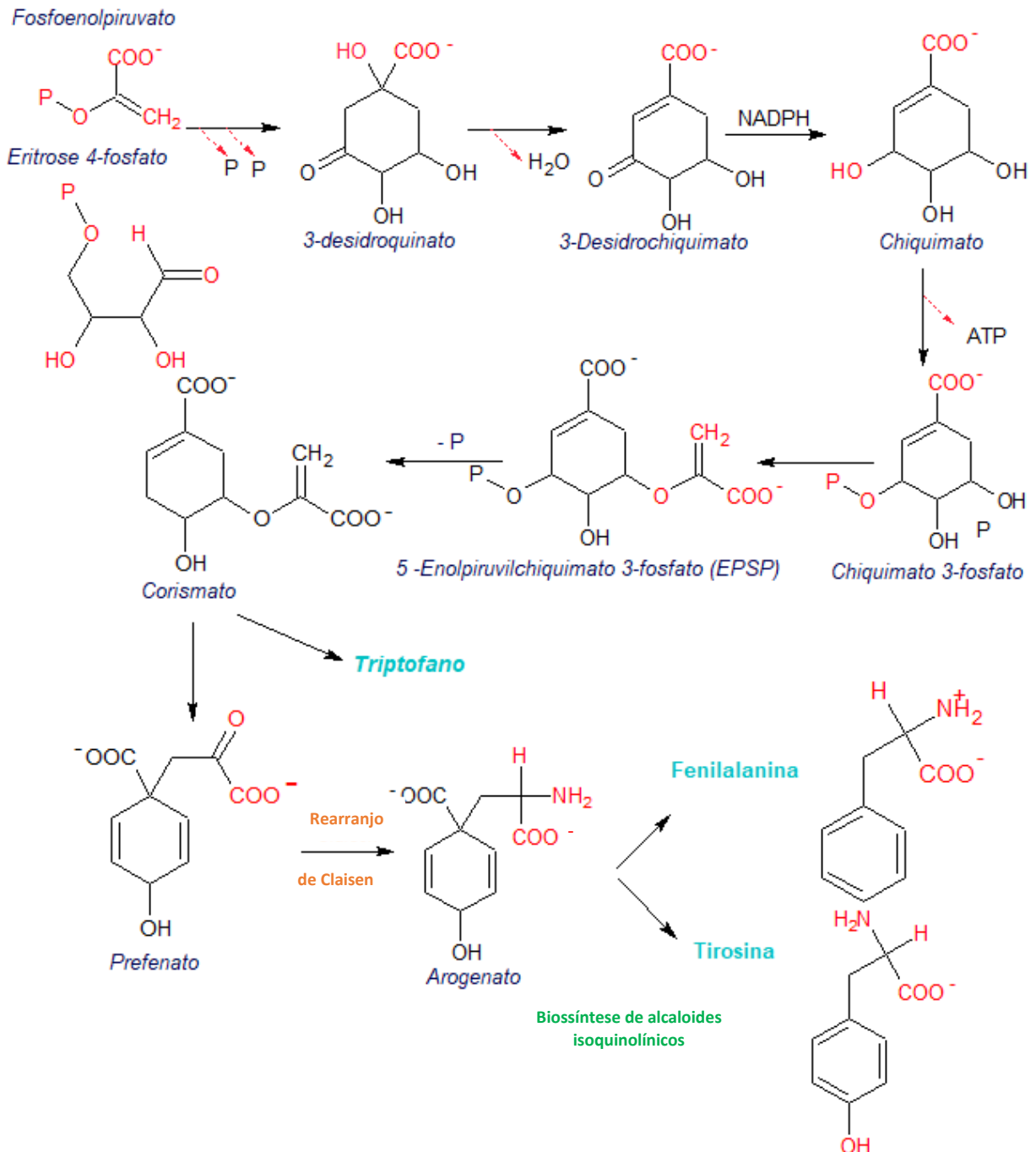
Constituindo uma classe estruturalmente importante, os alcaloides são moléculas que possuem o átomo de nitrogênio heterocíclico e são derivados de aminoácidos (RAJBIR KAUR; SAROJ ARORA, 2015). Em consonância, os alcaloides presentes na espécie *Erythrina velutina* são do tipo benzilisoquinolínicos (BIAs), uma família constituída de aproximadamente 2500 alcaloides que além de seu papel proeminente na medicina tradicional, têm uma ampla variedade de aplicações farmacológicas com potencial valor farmacêutico (NARCROSS et al., 2016), atuando como analgésicos (AMIRKIA; HEINRICH, 2014), antitumorais (KANG et al., 2015), antimicrobianos (PENG et al., 2015), entre outros.

De forma geral, sabe-se que eles derivam dos aminoácidos fenilalanina e tirosina, que de reações químicas e enzimáticas levam à formação inicial de um esqueleto básico característico do gênero e posteriormente em alcaloides mais específicos. Por outro lado, essas reações ainda não são suficientemente elucidadas, dados a respeito da biossíntese da espécie *Erythrina velutina* ainda não foram descritos.

A rota de biossíntese em questão dos aminoácidos aromáticos: fenilalanina e tirosina, inicia-se com a formação do ácido chiquímico a partir de uma reação de condensação entre eritrose-4-fosfato, proveniente da via das pentoses, e fosfoenolpiruvato (PEP), oriundo da glicólise, com liberação de ambos os grupos fosfatos seguido da remoção da água e da redução do grupamento carbonila (figura 2) (HELDT; PIECHULLA, 2011; RANDHIR; LIN; SHETTY, 2004). O ácido chiquímico (chiquimato) é a molécula base para formação dos aminoácidos percussores dos alcaloides eritrínicos.

Após a formação do chiquimato, uma reação química inicial de fosforilação ATP-dependente, seguida de uma reação enzimática pela 5-enolpiruvilchiquimato-3-fosfato sintase (EPSP sintase), que incorpora o fosfoenolpiruvato (PEP) na cadeia lateral do ácido-3-fosfochiquímico, forma um novo intermediário [5-enolpiruvilchiquimato-3-fosfato (EPSP)], que será transformado em ácido corísmico através de uma reação de eliminação. O corismato passa por uma reação de rearranjo de Claisen promovido pela corismato mutase favorecendo a formação de ácido prefênico e em seguida dois caminhos podem ser seguidos, um que leva à formação dos aminoácidos fenilalanina e tirosina e outro que leva à formação do triptofano (DEWICK, 2002).

**Figura 2** - Via do chiquimato



O chiquimato é um importante composto intermediário em vias bioquímicas de plantas e faz parte de umas das principais rotas metabólicas de produção de metabolismo secundário, em especial alcaloides. Fonte: elaborado pelo autor (2020).

A rota biossintética dos alcaloides eritrínicos tem uma compreensão bastante limitada, estudos são escassos e foram publicados há muito tempo. A primeira investigação biossintética dos alcaloides eritrínicos proposta (BARTON et al., 1968) evidenciou a (S)-norprotosinomenina como precursor dos alcaloides eritrínicos a partir de análise experimental utilizando principalmente técnicas de marcação isotópica de possíveis precursores. Estudos

posteriores sob a mesma condição experimental, mostraram que a (*S*)-reticulina e (*S*)-norcoclaurina foram metabolizados em alcaloides eritrínicos (MAIER et al., 1999), demonstrando a mobilização de diferentes precursores para diferentes esqueletos químicos dos alcaloides eritrínicos, o que trouxe um novo entendimento da biossíntese desses intrigantes alcaloides, pois ficou comprovado que (*S*)-norprotosinomenina, até o momento assumido como precursor de todos os alcaloides eritrínicos, não havia sido incorporado nas estruturas finais.

Além da espécie em estudo, *Papaver somniferum*, *Coptis japonica*, *Thalictrum flavum*, *Carthamus tinctorius* e *Eschscholzia californica* também são produtoras de BIAs (HAGEL et al., 2015; SOHRABI; ISMAILI; NAZARIAN-FIROUZABADI, 2018; ZHANG et al., 2019) e podem fornecer informações relevantes com estudos já estabelecidos na literatura, o que nos leva a ressaltar a importância da contribuição da quimiotaxonomia e intensificar a investigação desses peculiares alcaloides por ferramentas moleculares.

Se por um lado a pequena quantidade de informações se torna um grande desafio, por outro, a iniciativa de estudos como este, poderá possibilitar uma ferramenta fundamental para o entendimento metabólico em se tratando de produção de alcaloides e viabilizar modificação genética em plantas, abrindo novas perspectivas no âmbito das ciências farmacêuticas no Brasil.

### **2.3 Análises “ômicas”**

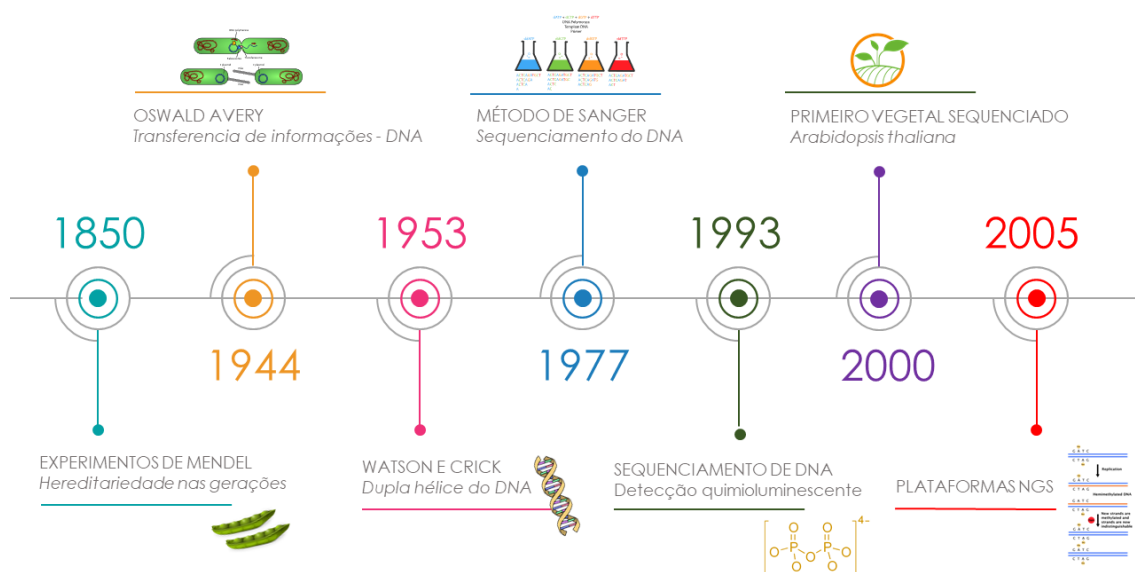
A prospecção de moléculas a partir da biodiversidade é uma forma viável e eficiente, na qual, as “ômicas” são uma importante ferramenta para o processo. Mesmo com a revolução que proporcionou o advento da genômica, somente seu emprego não garante a compreensão do sistema biológico. A era atual, denominada pós-genômica é capaz de informar e integrar os transcritos, proteínas e metabólitos de um organismo, estas ferramentas integram o que é denominado de “ômicas”, que mesmo sendo um fator que corrobora para compreensão de um organismo em nível sistêmico, ainda é uma tarefa árdua, pois existem especializações e interações celulares complexas na própria célula, em tecidos ou mesmo na interação com o meio ambiente (SOUZA; RHODEN; PAMPHILE, 2014).

#### **2.3.1 Análise do transcriptoma**

Constata-se que nos últimos 100 anos houve uma grande evolução acerca do conhecimento da informação genética e molecular e quais suas características. Ao verificar esses fatos históricos compreende-se melhor a importância de tecnologias indispensáveis para pesquisa biológica, para solucionar problemas que até então não eram possíveis.

Nesta escala histórica, uma das primeiras contribuições pode ser atrelada a Mendel em 1850 com a descoberta que as características genéticas são transmitidas por hereditariedade ao longo das gerações, fortalecendo e corroborando com a teoria evolutiva de Darwin, que se mantém firmemente até os dias atuais (SCHWARZBACH et al., 2014). O grupo do pesquisador Oswald Avery (AVERY; MACLEOD; MC CARTY, 1994) demonstrou que a transferência de informações entre organismos (bactérias) é viável e que ocorre pelo material genético, o DNA (figura 3).

**Figura 3** - Linha do tempo da evolução da informação genética



Fonte: elaborado pelo autor (2020).

Em 1953 a estratégia empregada por Watson e Crick apresentada como um modelo compatível com os resultados experimentais que haviam sido obtidos até aquele momento foi construir um modelo molecular de estudos de difração de raios X possibilitou chegar à conclusão que a molécula de DNA é formada por duas longas cadeias paralelas, constituídas por nucleotídeos dispostos em sequência (HUTCHISON, 2007; SILVA, 2007).

Estas descobertas deram início a um período investigação para uma melhor compreensão a respeito material genético, porém, tecnologias que proporcionassem o sequenciamento de genes pareciam distantes e quase impossíveis. Não obstante, em 1977, Sanger desenvolveu seu método para determinar as sequências de fragmentos relativamente pequenos de DNA humano. A partir desse período, esforços foram realizados e diferentes métodos de sequenciamento começaram a ser desenvolvidos (SALZBERG; POP, 2008).

Ainda nesse contexto, um dos primeiros métodos desenvolvidos na década de 90 foi o método de sequenciamento de DNA em fase sólida para detecção não radioativa de mudanças

de base sem a necessidade de eletroforese (NYRÉN; PETTERSSON; UHLÉN, 1993), que se baseava em uma reação mediada pela DNA polimerase por um ensaio de detecção quimioluminescente de pirofosfato inorgânico.

No ano 2000 foi iniciado o desenvolvimento de uma nova forma de sequenciamento e o primeiro vegetal a ter sequência genômica completa - *Arabidopsis thaliana* (Iniciativa do Genoma de *Arabidopsis*). Mais tarde em 2005, foi introduzida uma nova era de análises de transcriptoma de alto rendimento – o “sequenciamento de próxima geração (NGS)” (ANSORGE, 2009). Avanços recentes no sequenciamento de alto rendimento (RNA-Seq) forneceram um meio para rápida caracterização e quantificação de transcriptomas, envolvendo o sequenciamento direto de DNAs complementares (cDNAs), seguido pelo mapeamento das leituras de sequenciamento para o genoma ou conjuntos de genes de referência para análise da expressão gênica e detecção de polimorfismos (KUMAR et al., 2012).

O transcriptoma consiste no conjunto completo de transcritos (RNAs mensageiros, RNAs ribossômicos, RNAs transportadores e os micro RNAs) de um dado organismo, parte, tecido ou linhagem celular. (COSTA et al., 2010).

Várias tecnologias foram desenvolvidas para quantificar o transcriptoma, incluindo abordagens baseadas em sequência (WANG; GERSTEIN; SNYDER, 2009), que define diretamente a sequência de cDNA e foram utilizadas neste estudo pela plataforma *Illumina NextSeq 550*, tecnologia de 2ª geração.

A primeira etapa envolvida na tecnologia citada na plataforma *Illumina* é a preparação das bibliotecas contendo fragmentos de cDNA de tamanho compatível a serem sequenciadas que serão ligados à adaptadores em ambas as extremidades. Após a fixação, o adaptador da extremidade livre liga-se ao *primer* complementar adjacente na placa de vidro (*flowcell*) povoada por adaptadores complementares dos fragmentos (FIETTO; MACIEL, 2015). A alta densidade de *primers* ligados a *flowcell* facilita esta ligação que tem a função de iniciar a extensão da fita pela RNA polimerase, ao encontrar um OH livre na extremidade do *primer* livre (MARDIS, 2008).

Ainda a respeito, são fornecidos nucleotídeos não marcados que proporcionarão a extensão de uma fita de cDNA complementar a fita molde fixado na placa. A extremidade desta fita recém sintetizada se anela ao *primer* na extremidade da fita molde formando uma estrutura em ponte que dá nome ao processo de amplificação. Posteriormente, ocorre desnaturação reiniciando o processo.

Na segunda etapa, após vários ciclos de sequenciamento, são adicionados ao longo de toda a extensão da placa uma solução contendo os quatro tipos de nucleotídeos terminadores

marcados com fluorescência. Quando necessário estes nucleotídeos contendo os fluoróforos são adicionados às respectivas cadeias em cada um dos *clusters*. Posteriormente, ocorre a excitação do fluoróforo por feixes de raios laser fazendo com que os nucleotídeos emitam uma fluorescência dos clusters (FIETTO; MACIEL, 2015; KIRCHER; STENZEL; KELSO, 2009).

Ao final uma imagem contendo a cor da fluorescência é capturada para cada posição dos *clusters* na placa de vidro. Este processo se repete até que todas as bases de um fragmento sejam determinadas. Como consequência do sequenciamento, são gerados milhões de fragmentos menores (*reads*) que se alinham baseando-se em regiões de sobreposição entre eles para produzir sequências únicas e maiores denominadas *contigs*.

Diversos trabalhos com transcriptoma de espécies da família Fabaceae são citados na literatura, como é o caso da espécie *Pisum sativum* L., conhecida como ervilha de jardim. Cinco bibliotecas de cDNA obtidas a partir de nódulos ou raízes noduladas foram usadas neste estudo para sequenciamento de próxima geração (ZHERNAKOV et al., 2017). Esta espécie é considerada uma das leguminosas mais importantes do ponto de vista agrícola no mundo e uma planta modelo versátil para estudar as bases genéticas de interações benéficas planta-microrganismo (SHTARK et al., 2010).

Do mesmo modo, foi aplicado o RNA-Seq para investigar em sete tecidos em sete estágios no desenvolvimento de sementes em *Glycine max* (soja), comparando as leituras de transcrição com a mais recente sequência do genoma da mesma espécie, investigando a relação entre a estrutura e expressão gênica para tecidos (SEVERIN et al., 2010).

Mais recentemente, a espécie *Vicia faba* L., o feijão fava, também teve seu transcriptoma sequenciado (KHAN et al., 2019). Neste trabalho foram utilizadas as folhas das plantas controle e submetidas a estresse hídrico no estado vegetativo e as folhas das plantas controle e com estresse hídrico na fase de florescimento. Após montagem e anotação dos transcritos, foi observado como processos mais significativos: os processos metabólicos, celulares, a resposta ao estímulo, sinalização e desintoxicação, além de diferentes transcritos expressos nos estágios vegetativo e em floração.

Além de espécies da família Fabaceae, autores relatam também a investigação de transcritos por métodos de sequenciamento em espécies produtoras de alcaloides benzilisoquinolínicos. Entre elas: *Eschscholzia californica*, *Glaucium flavum*, *Papaver bracteatum*, *Sanguinaria canadensis*, *Thalictrum flavum*, *Berberis canadensis* e *Cocculus trilobus* (FARROW; HAGEL; FACCHINI, 2012). Com destaque a espécie *Papaver somniferum* que possui o genoma sequenciado (CELIK et al., 2014).

Desde tempos imemoriais, o ópio e os seus derivados, além de exercerem ponderável influência sobre o comportamento dos seres humanos, têm sido empregados como sedativo e como analgésico. A partir do século XIX, com o isolamento dos alcaloides do ópio e as facilidades para o emprego, houve aumento do interesse pelo uso (DUARTE, 2005). Um dos fatores mais significativo nessa escala história foi a descoberta da morfina, o primeiro alcaloide extraído e somando 10% dos alcaloides presente na espécie do ópio.

O isolamento da morfina foi a porta de entrada para estabelecer estruturas químicas de outros derivados naturais ampliando o conhecimento para diferentes espécies, e hoje, é considerada como uma planta modelo em estudos tanto taxonômicos e quimiotaxonômicos, sendo relatados na literatura vários transcritos e proteínas identificados e caracterizados na rota de biossíntese de alcaloides benzilisoquinolínicos, entre eles, as acetiltransferase, oxidoredutases e metiltransferases que estão entre os mais evidenciados (GROTHER; LENZ; KUTCHAN, 2001; LISCOMBE; FACCHINI, 2007), fornecendo valiosas informações acerca da formação do esqueleto químico benzilisoquinolínicos.

### 2.3.2 Análise metabolômica

O aprimoramento de tecnologias de sequenciamento e da proteômica permitiu que abordagens de metabolômica chegassem como forma complementar ao conhecimento, pois mesmo que investigado o conjunto de transcritos e proteínas de um organismo não é garantia de funcionalidade, então, desta forma os metabólitos surgem como norteadores da funcionalidade gênica. Metabólitos são os produtos finais dos processos regulatórios celulares como resposta final dos sistemas biológicos às mudanças genéticas ou ambientais e estão associados aos termos transcriptoma e proteoma (FIEHN, 2002).

A tecnologia voltada para uma visão geral compreensiva dos metabólitos presentes em um organismo é denominada metabolômica, que está prevista para desempenhar um papel significativo, se não indispensável, em preencher o intervalo fenótipo-genótipo (HALL, 2006). O uso do termo perfil metabólico como uma nova ferramenta comparativa da função do gene foi introduzido principalmente nos anos 2000 (FIEHN et al., 2000). Enquanto o metaboloma está associado ao conjunto de todos os metabólitos pertencentes a um organismo, o perfil metabólico por outro lado, envolve uma análise rápida, muitas vezes não quantitativa, com o objetivo de identificar um perfil metabólico específico que caracterize uma determinada amostra (VILLAS BÔAS et al., 2005).

A análise do metaboloma pode ser dividida em duas categorias: *targeted* e *untargeted* ou *nontargeted* (FRENCH; HARVEY; MCCULLAGH, 2018; LEI; HUHMANN; SUMNER, 2011). Enquanto que na primeira abordagem a análise é direcionada para moléculas alvos, que já são previamente conhecidas, a *untargeted* faz um perfil total dos metabólitos presentes no sistema (DE VOS et al., 2007; DUDLEY et al., 2010; LIN et al., 2011).

Após a extração, a identificação de metabólitos se dá geralmente por metodologias refinadas, pois detectar essas moléculas em organismos complexos se torna um desafio, sobretudo em organismos vegetais. O uso da espectrometria de massas (MS) e espectroscopia de ressonância magnética nuclear (RMN) são as metodologias mais utilizadas e vem sendo recentemente empregada em análises metabolômicas (LIU et al., 2019; YOBİ et al., 2012, 2013).

A espectroscopia de RMN fornece informações detalhadas sobre a estrutura molecular, para compostos puros e em misturas complexas, mas também pode ser usada para sondar a dinâmica molecular e a mobilidade dos metabólitos por meio da interpretação dos tempos de spin-relaxamento de RMN e pela determinação dos coeficientes de difusão molecular (LINDON; NICHOLSON, 2008). Por mais que o RMN seja geralmente aceito como padrão-

ouro na elucidação da estrutura, a técnica possui uma sensibilidade relativamente menor. Desta forma, a espectrometria de massas fornece informações químicas altamente específicas que estão diretamente relacionadas à estrutura química, tais como massa precisa, padrões de distribuição de isótopos para determinação de fórmulas e íons de fragmentos característicos para elucidação estrutural via correspondência espectral para dados compostos autênticos (LEI; HUHMAN; SUMNER, 2011).

Estudos relatam a metabolômica baseada em duas estratégias MS: análise direta por MS e análise por MS acoplado a técnicas de separação. As análises diretas não necessitam de separação cromatográfica e podem ser realizadas por infusão direta por ESI, MALDI, e por técnicas de pressão atmosférica como DESI (Desorption Electrospray Ionization), ESSI (Electro-Sonic Spray Ionization), EASI (Easy Ambient Sonic-Spray Ionization), e mais recentemente por Imaging MS, enquanto que as análises que necessitam de pré-separação podem ser realizadas por GC-MS (*Gas Chromatography – Mass Spectrometry*), LC-MS (*Liquid Chromatography - Mass Spectrometry*) e CE-MS (*Capillary electrophoresis - Mass Spectrometry*) (ABDELNUR, 2011).

Frente ao breve exposto, após análise por espectrometria de massas, a mineração de dados pode ser realizada por meio de banco de dados para comparação espectral. O *Global Natural Social Molecular Networking* (GNPS) é uma plataforma de espectrometria de massas que visa ser uma base de conhecimento de acesso aberto e compartilhamento de dados de espectrometria de massa.

Este banco de dados utiliza como ferramenta de análise a pesquisa de rede molecular, com exibições visuais do espaço químico presente em experimentos de espectrometria de massa. Esta abordagem agrupa conjuntos de espectros de moléculas relacionadas (conhecidas como famílias moleculares) mesmo quando os próprios espectros não são identificados (não correspondem a nenhum composto conhecido) (<https://gnps.ucsd.edu/>), e foi a plataforma utilizada neste estudo.

Desde a década de 90 os estudos fitoquímicos acerca de alcaloides presentes em espécies do gênero *Erythrina* aumentam consideravelmente. Em uma recente revisão da literatura (RAMBO et al., 2019) demonstra os dados atuais sobre composição química do gênero de 1990 a 2017 presente em 30 espécies vegetais. Esse estudo relata a identificação de 91 alcaloides durante o período citado, apresentando propriedades ansiolíticas e anticonvulsivantes já estabelecidas associadas a esses alcaloides.

Recentemente, foram isolados os alcaloides majoritários da semente de *Erythrina velutina*, a erisodina (C<sub>18</sub>H<sub>21</sub>NO<sub>3</sub>) e eritralina (C<sub>18</sub>H<sub>19</sub>NO<sub>3</sub>) (ALMEIDA et al., 2020). Alcaloides

como 8-oxoeritralina, 10,11-dioxoeritralina, eritralina, erisotrina, eritratina, 11 $\beta$ -hydroxierisodina e os intermediários reticulina e norreticulina foram identificados em flores e sementes das espécie de *Erythrina corallodendron*, *Erythrina stricta* e *Erythrina abyssinica* (ZHAO et al., 2018; WU et al., 2018; MOHAMMED et al., 2012).

Em 1991, alcaloides glicosilados do tipo dienoide também foram isolados no gênero por Amer e colaboradores, são eles: 11b-metoxiglucoerisodina, e glucoerisodina em *Erythrina lysistemon*, o isolado 11- $\beta$ -metoxiglucoerisovina dos frutos de *Erythrina vespertilio* em 2012 e eritralina-11 $\beta$ -*O*-glucopiranosideo das sementes de *Erythrina crista-galli* (TAN et al., 2017).

### **3 OBJETIVOS**

#### **3.1 OBJETIVO GERAL**

Contribuir para o entendimento da biossíntese de alcaloides presente em folhas e sementes da espécie *Erythrina velutina* utilizando as tecnologias “ômicas” - transcriptoma e perfil metabólico.

#### **3.2 OBJETIVOS ESPECÍFICOS**

3.2.1 Detectar e identificar o perfil de alcaloides eritrínicos presente em folhas e sementes de *Erythrina velutina* oriundas da Caatinga do RN;

3.2.2 Analisar os dados referentes ao metabolismo de nitrogênio no que se refere a alcaloides eritrínicos oriundos do transcriptoma total de folhas e sementes de *E. velutina*;

3.2.3 Integrar os dados de genes expressos a via biossintética de alcaloides eritrínicos;

3.2.4 Sugerir uma possível rota biossintética para alcaloides em folhas e sementes de *E. velutina* coletada na Caatinga do RN a partir da análise integrativa;

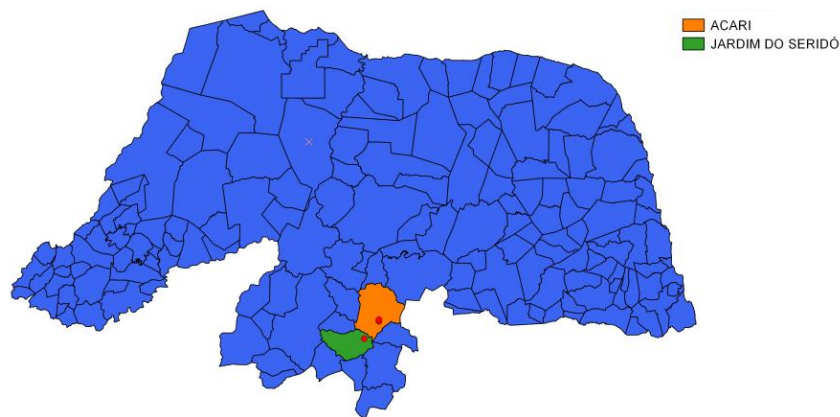
3.2.5 Desenvolver conhecimento para a área de pesquisa em produtos naturais com enfoque multidisciplinar permeando por aspectos químicos e biológicos, contribuindo com o entendimento da dinâmica de produção dessas moléculas com potencial bioativo e subsidiar futuros estudos biotecnológicos.

## 4. METODOLOGIA

### 4.1 Material vegetal

Folhas e sementes de plantas de *Erythrina velutina* Willd foram coletadas em populações naturais em quatro diferentes localizações, duas na cidade de Acari (localização A: S 06,47207°/W 036,64236° e localização B: S 06,47205°/W 036,64236° com altitude de 270 metros acima do nível do mar) e duas em Jardim do Seridó (localização C: S 06,56349°/W 036,71972° e localização D: S 06.46235°/ W 036.64183° com altitude de 218 metros acima do nível do mar) no Rio Grande do Norte (figura 4), em março de 2018, durante o período diurno e seco. Para cada população (com uma distância entre 2 a 8 km entre si), cinco plantas adultas com tamanho médio de 5-10 metros foram selecionadas. A coleta se deu pela seleção de folhas verde-escuras, totalmente expandidas, sem sinais de herbivoria, localizadas nas extremidades dos ramos na parte superior da árvore e mais expostas a radiação solar, enquanto que apenas sementes com cor vermelha intensa presas a vagens maduras ainda na planta matriz foram coletadas.

**Figura 4** - Mapa de localização com pontos de coleta



O Seridó do Rio Grande do Norte constitui uma importante região natural nordestina, possui características únicas se comparadas a outras regiões do RN pois é influenciada por fatores edafoclimáticos particulares, além da diversidade natural que a região apresenta. Fonte: elaborado pelo autor (2020).

Em campo, os tecidos vegetais foram lavados rapidamente com água destilada, imediatamente congelados em nitrogênio líquido e armazenados a -80 °C até o uso. A espécie

foi identificada pelo botânico MSc. Alan Araújo Roque e um exemplar de cada ponto de coleta foi depositado no Herbário da Universidade Federal do Rio Grande do Norte, sob o número de referência (UFRN16079). A autorização para a coleta do material vegetal foi concedida pelo SISBIO (327493) e a permissão para acesso ao patrimônio genético brasileiro foi fornecida pelo SISGEN (A8E4663).

As amostras foram organizadas em *pools* de 1 a 8, onde os *pools* de 1 a 4 correspondiam a amostras de sementes (p1 e p3 – duas localizações de Acari; p2 e p4 – duas localizações de Jardim do Seridó, cada amostra sendo composta por material de 5 indivíduos diferentes) e *pools* de 5 a 8 correspondiam a amostras de folhas (p6 e p8 – duas localizações de Jardim do Seridó; p5 e p7 – duas localizações de Acari, cada amostra sendo composta por material de 5 indivíduos diferentes).

Para entender melhor a *Erythrina velutina* em seu ambiente nativo na Caatinga brasileira e fatores que podem contribuir na produção de metabólitos secundários, em especial os alcaloides, abordagens adicionais foram reunidas por meio de análises ambientais e análise química dos solos dos locais de coleta desse estudo.

## 4.2 Dados edafoclimáticos e análise do solo

Os dados climáticos foram obtidos a partir de uma colaboração com a Dra. Renata Libonati, do Instituto de Geociências da Universidade Federal do Rio de Janeiro (IJEU-UFRJ). O desenvolvimento de mapas de estimativas de precipitação total, área de queima, fogo ativo, radiação solar líquida da superfície, temperatura e umidade relativa foram realizados utilizando dados de satélites meteorológicos para avaliação e caracterização climática da Caatinga do Rio Grande do Norte em um período de 1 ano de antecedência da coleta, de março de 2017 a fevereiro de 2018.

Parâmetros e detalhes na análise podem ser encontrados a seguir: **Precipitação total** (tp): Era-Interim 0.25°x0.25° - mensal - <https://apps.ecmwf.int/codes/grib/param-db?id=228>; **Área de queima**: AQM-MODIS V6 0.01°x0.01° - mensal - <http://prodwww-queimadas.dgi.inpe.br/aq1km/> and - MCD64A1 C6 0.01°x0.01° - mensal - <https://search.earthdata.nasa.gov/>; **Fogo ativo**: VIIRS 375m: pontual e 0.01°x0.01° - diário - <https://firms.modaps.eosdis.nasa.gov/download/>; **Características da vegetação**: MYD13A3 (Aqua) C6 0.01°x0.01°, layer '1\_km\_monthly\_NDVI', reamostragem por 'Nearest Neighbor' - mensal - <https://search.earthdata.nasa.gov/>; **Cobertura e uso da terra**: - MCD12Q1 C6 0.01°x0.01° - anual - <https://search.earthdata.nasa.gov/>; **Radiação solar** (MJ m-2): **radiação solar líquida**, clear sky (ssrc) Era-Interim 0.25°x0.25° - mensal - <https://apps.ecmwf.int/codes/grib/param-db?id=210>; **Temperatura** (K): Temperatura cutânea (skt) Era-Interim 0.25°x0.25° - mensal - <https://apps.ecmwf.int/codes/grib/param-db?id=235>; **Umidade relativa** (%): - Combinação de Dew 2m Temperatura (d2m) + 2m Temperatura (t2m) + Pressão da superfície (sp) Era-Interim 0.25°x0.25° - mensal.

A caracterização química do solo para determinação das concentrações de metais pesados, micronutrientes e fertilidade foram realizadas no Laboratório de Análise de Água da EMPARN, seguindo a metodologia proposta por EMBRAPA (1997).

### 4.3 Análise do perfil metabólico direcionado para alcaloides eritrínicos

#### 4.3.1 Extração de metabólitos

Quanto à abordagem do perfil metabólico, as sementes e folhas congeladas de *Erythrina velutina* foram trituradas com almofariz e pistilo em nitrogênio líquido, seguida de liofilização. Ao pó da planta foram adicionadas as misturas de solventes na proporção de 10 mg da planta a 100 µL de etanol:água [70:30 (v/v)] (ANHESINE, 2018), misturados com vórtex por 10 segundos, seguido de 30 minutos de banho de ultrassom (40 kHz, 200W) e filtração em papel filtro qualitativo de 12,5 cm de comprimento e 80g/m<sup>2</sup> de gramatura.

O procedimento de extração foi repetido três vezes para cada amostra. Essas mesmas amostras foram usadas na análise em etapas adicionais de secagem (sistema *SpeedVac*) e solubilização apropriada a ferramenta analítica LC-MS, que constituiu de uma mistura de etanol e água *Milli-Q* suficiente para atingir a concentração de 10 mg/mL, 3,0 µL de cada uma foram injetados no sistema LC-MS.

A análise e processamento dos dados foram realizados no Núcleo de Bioensaios, Biossíntese e Ecofisiologia (NuBBE-UNESP) em colaboração com o Professor Alberto José Cavalheiro. O procedimento de extração foi repetido três vezes para cada amostra.

#### 4.3.2 Análise por LC/MS e DAD

A separação e detecção de compostos polares médios foram realizadas em um controlador Shimadzu Classe-LC 10, CBM 20A e um espectrômetro de massa *micrOTOF-QII* (Bruker Daltonics) usando o electropulverização (ESI) para ionização e tempo de voo (TOF) como analisador. Utilizou-se uma coluna XB C18 Kinetex (100 x 2,1 mm, 2,6 µm, 100 Å), a 40 °C, vazão de 0,400 mL/min e um detector UV-DAD operando entre 190 e 400 nm.

O método cromatográfico foi baseado nos estudos de Vos (2007) e Lisec (2006). A fase móvel foi composta por H<sub>2</sub>O (A) e acetonitrila (B), ambos com ácido fórmico a 0,1% (v/v), utilizando o seguinte gradiente: 5 - 35% B (23 min); 35,0 - 100% B (17 min), mantendo 100% B por 8 minutos, seguiu a redução de 100 - 5% B (2 min) e manteve a condição inicial (5% B) por 10 minutos, totalizando 60 minutos de análise.

Soluções de teste de mistura (mistura de padrões - aminoácidos, flavonóides, etc.) e controles de qualidade (QC - mistura de amostras - 40 µL cada) foram injetados para garantir a reprodutibilidade do método, instrumento e condições cromatográficas. As experiências de HRMS foram realizadas em um espectrômetro de massa *micrOTOF-QII* (Bruker Daltonics) no

modo de ionização positiva. Os parâmetros de aquisição foram: tensão capilar e cone, 4500 V e 500 V, respectivamente, temperatura capilar a 200 °C, N<sub>2</sub> como gás de secagem a 4,5 bar e vazão de 9,0 L/min. Os espectros foram adquiridos entre 100 - 1500 Da e para análise de massa de alta resolução ou instrumentos e as experiências foram calibradas usando uma solução Na-TFA padrão interna (4 mg/mL) no final de cada análise. As experiências MS/MS foram realizadas no modo Auto-MS / MS (modo positivo), selecionando até 5 íons precursores dentro de um ciclo de 3,0 s. Para a fragmentação dos íons precursores, foi utilizado um gradiente de energia de 20 a 50 eV. Inicialmente, as amostras em branco foram analisadas para identificar interferências e contaminantes do instrumento e da coluna cromatográfica. Estes foram então excluídos do ciclo de seleção e fragmentação do Auto-MS/MS por uma função que permite suas exclusões.

### 4.3.3 Análise da rede molecular MS/MS (*Molecular Networking*)

Para realizar a análise da Rede Molecular, todos os dados do MS/MS foram inicialmente convertidos para o formato *mzXML* usando o *software MSConvert* e processados usando o pacote *MzMine* (script java) para detecção e calibração de massa, deconvolução de pico, agrupamento de massa (padrão isotópico) e alinhamento espectral (PLUSKAL et al., 2010). Depois que todos os dados foram processados usando o software *MzMine* (KATAJAMAA E ORESIC, 2005), os espectros foram exportados em formato de texto (.mgf) e carregados na plataforma de Rede Molecular Social dos Produtos Naturais Globais (GNPS) para inicializar o cálculo da similaridade entre espectros (pontuação de similaridade de cosseno). Em teoria, os espectros são convertidos em vetores que são pareados e comparados por um cálculo de produto escalar. A pontuação de similaridade de cosseno atribuída a cada par de vetores varia de 0 a 1, em que 0 representa completamente diferente e 1 representa espectros idênticos. Os dados foram filtrados removendo todos os picos de MS<sub>2</sub> dentro de  $\pm 17$  u do valor do precursor *m/z*. Os espectros de MS<sub>2</sub> foram filtrados escolhendo apenas os seis picos superiores na janela de  $\pm 50$  u em todo o espectro. Os dados foram agrupados em cluster com *MSCluster* com uma tolerância de massa pai de 1,0 Da e uma tolerância de íons de fragmento MS<sub>2</sub> de 0,5 Da para criar espectros de consenso. A rede foi criada filtrando as arestas para ter uma pontuação de cosseno acima de 0,65 e mais de 6 picos correspondentes.

Os espectros na rede foram pesquisados nas bibliotecas espectrais do GNPS (GNPS: <https://gnps.ucsd.edu/ProteoSAFe/libraries.jsp>), usando a mesma configuração dos dados de entrada. A visualização da rede foi realizada no *Cytoscape* 2.8.3 e 3.4.3. Para evitar a interpretação incorreta de contaminantes por HPLC, as injeções em branco (fase móvel) foram

carregadas como um grupo de amostra distinto no fluxo de trabalho do GNPS e excluídas das redes.

Além disso, os nós da rede molecular coloridos com base na estrutura da planta (folha e/ou semente) na espécie de *Erythrina* e na espessura da borda atribuída aos escores de similaridade do cosseno (espessura de 0,95 a 1). Linhas mais grossas indicam maior semelhança. Sub-redes foram geradas no *Cytoscape* a partir de partes isoladas da rede maior, a fim de melhorar a análise da dinâmica metabólica, fornecer mais elucidação de moléculas e melhor visibilidade da espessura da borda e conectividade dos nós.

## **4.4 Análise do transcriptoma**

### **4.4.1 Análise de RNA-Seq**

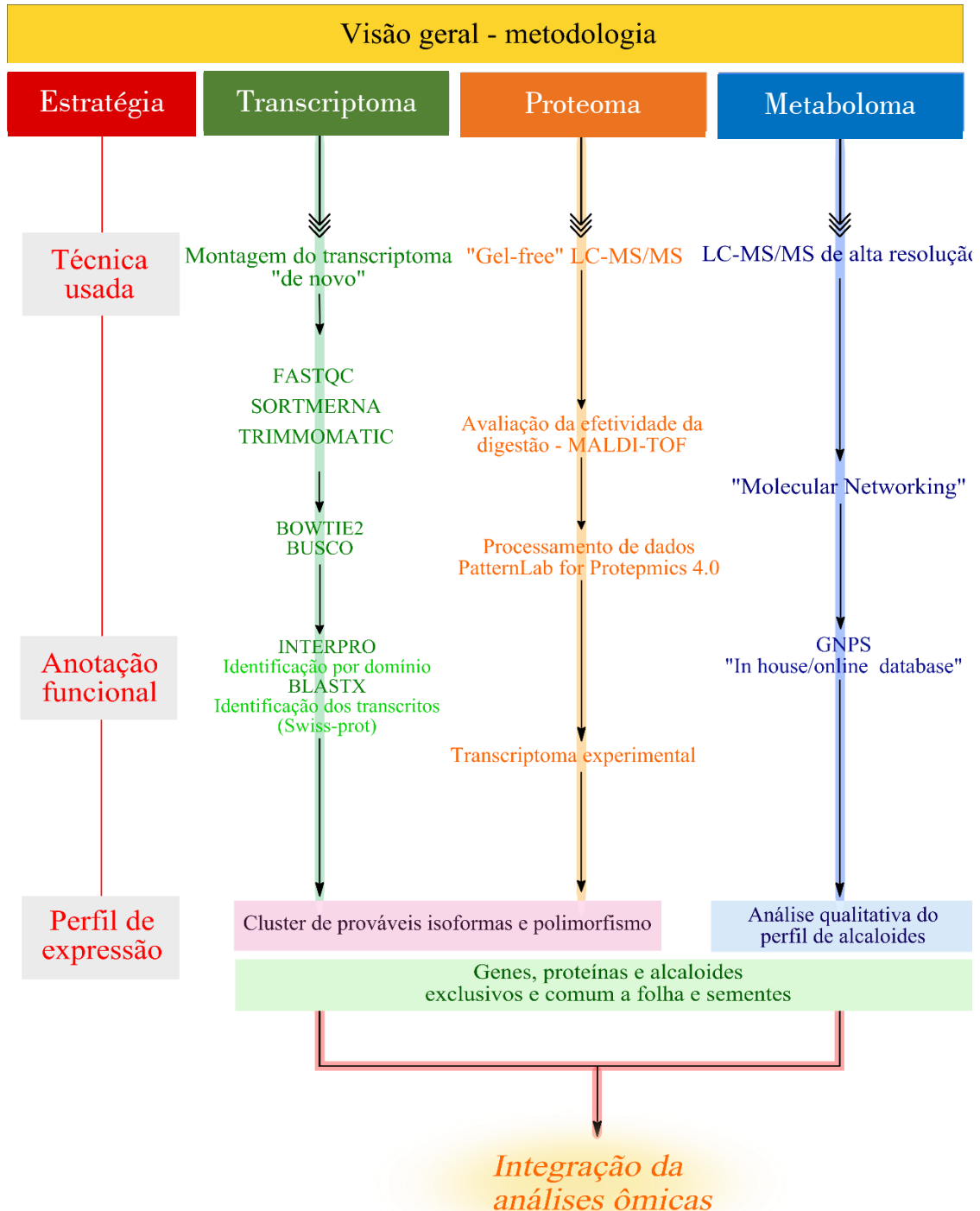
#### **4.4.1.1 Extração do RNA total e preparação da biblioteca de cDNA**

O RNA total de sementes e folhas foi isolado usando o kit de purificação de RNA *Trizol™ Plus* (Life Technologies) de acordo com o protocolo do fabricante. Para obter amostras de alta qualidade, foram tratadas com o kit *DNase Free RNase* (Qiagen) e posteriormente purificadas com o kit *RNeasy Mini* (Qiagen). A quantidade do RNA foi avaliada usando o espectrofotômetro *NanoDrop ND-1000* e a qualidade foi avaliada em *Bioanalyzer 2100* (Agilent). Todas as amostras de RNA que tinham um valor de número de integridade do RNA (RIN) acima de 6,0 foram escolhidas para quantificação no fluorômetro *Qubit* (Thermo Fischer Scientific) e preparação de bibliotecas de cDNA. Para a preparação de bibliotecas de cDNA cada pool foi composto por três amostras de RNA igualmente reunidos com um total de 0,5µg de RNA. As bibliotecas foram preparados com o Kit *mRNA TruSeq Stranded* de acordo com o protocolo *Illumina* e quantificadas usando o kit *Kapa Library Quantification* para a plataforma de Genoma da Illumina (*Kapa Biosystems*) na máquina *Applied Biosystems ViiA7* (*Life Technologies*) de acordo com as instruções do fabricante, estas etapas foram realizadas no Centro Nacional de Pesquisa em Energia e Materiais (CNPEM – Campinas/SP), em colaboração com o pesquisador Leandro Vieira da Silva. A mesma quantidade de cada biblioteca de RNAseq foi reunida para equalizar a concentração do índice e, em seguida, submetida à *Facility GENIAL* no Centro de Facilidades de Apoio a Pesquisa (CEFAP - USP) para o sequenciamento de acordo com os protocolos padrão da plataforma *NextSeq500* (*Illumina*).

#### **4.4.1.2 Análise e anotação do RNAseq**

O RNA total foi sequenciado de acordo com os protocolos padrão da plataforma *NextSeq500* (Illumina), gerando leituras de 150 pb. A metodologia do transcriptoma bem como das demais abordagens “ômicas” estão resumidas na Figura 5. As amostras foram organizadas por *pools* de amostras (de 1 a 8), onde de 1 a 4 correspondiam ao RNA das sementes e de 5 a 8 correspondiam ao RNA das folhas, onde os números sinalizam diferentes pontos de coleta de pelo menos 3 amostras.

**Figura 5** - Estratégia experimental e integração das análises ômicas



A prospecção de substâncias a partir da biodiversidade é uma fonte de bastante eficiência e as ferramentas que podem auxiliar nesse processo se torna cada vez mais importante onde a tendência é utilizar dados ômicos para prospecção de compostos bioativos de origem natural. Integralizar esses dados contribui para melhor compreensão do sistema biológico, direcionando para um maior entendimento de produção de metabólitos pela planta. Fonte: elaborado pelo autor (2020).

A qualidade das sequências foi avaliada pelo software *FastQC* (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), que permite verificar alguns parâmetros de qualidade nos dados brutos da sequência de sequenciamento de alto rendimento. A remoção do RNAr foi realizada usando o software *SortMeRNA* (v2.1) para todas as leituras que combinam as cadeias *Forward* (R1) e *Reverse* (R2). Essa etapa foi necessária para melhorar a qualidade das leituras. O RNAr de outros eucariotos, procariotos e arqueias foi identificado e removido. As leituras das bibliotecas foram cortadas por *Trimmomatic* (BOLGER; LOHSE; USADEL, 2014) para remover adaptadores e leituras de baixa qualidade, o corte da janela foi 2:20. O intervalo final de leituras estava entre 21 e 151 pares de bases e a pontuação *phred* acima de 20.

É importante ressaltar que *Erythrina velutina* é uma espécie brasileira nativa na qual o genoma não está disponível, portanto, nossos dados de RNA-*seq* devem ser reunidos “*de novo*”, o que significa que as sequências foram montadas pela sobreposição de leituras para formar *contigs*, em seguida comparando-se com outras espécies de plantas. Diante disso, os arquivos *FASTQ* aparados descritos acima foram usados em uma montagem *de novo* no software *Trinity* (WATERHOUSE et al., 2017), utilizou-se uma faixa de tamanho de 21 a 31 kmer, sendo que os melhores resultados para mapeamento de genes de ortólogos de plantas no software *BUSCO* (WATERHOUSE et al., 2017; SIMÃO et al., 2015) foram obtidos com 21kmer.

Para validar as montagens geradas no *Trinity*, usamos o software *BUSCO* v3 (*Benchmarking Universal Single-Copy Orthologs*), funcionando como uma ferramenta que fornece medidas quantitativas para a avaliação da montagem do genoma, conjunto de genes e integridade do transcriptoma, com base em expectativas evolutivas e informadas sobre o conteúdo de genes de ortólogos de cópia única quase universais selecionados no *OrthoDB* v9 e cruzando com o banco de dados Unigene.

Em seguida, nas próximas etapas foi utilizado o software *Bowtie2* (LANGMEAD; SALZBERG, 2012) usado também como uma ferramenta para alinhar as leituras de sequenciamento com as sequências montadas finais para validar o transcriptoma. O *Bowtie* pode executar os modos de alinhamento *gaped*, local e emparelhado, no entanto, para nosso propósito, o alinhamento foi feito entre nossas leituras limpas usadas para a montagem e a montagem final para cada leitura lida ou combinada.

Os transcritos montados foram submetidas ao banco de dados do *InterProScan* (v5.30-69.0) (<http://www.ebi.ac.uk/interpro/>) que combinam assinaturas de vários bancos de dados

(*PROSITE*, *Pfam*, *PRINTS*, *ProDom*, *SMART*, *TIGRFAMs*, *PIRSF*, *SUPERFAMILY*, *GENE3D* e *PANTHER*) para uma primeira anotação. Nessa etapa foi usado como parâmetro de identificação de no mínimo 50 % de cobertura e 70 % de identidade do domínio funcional frente à plataforma dos bancos de dados citados.

Logo após a anotação e identificação considerando o domínio, foram selecionados manualmente os transcritos com domínios proteicos das proteínas alvo candidatas a participarem do estudo de biossíntese de acordo com informações em banco de dados público e informações da literatura. Essas sequências foram submetidas a uma segunda anotação *BLASTx* contra o banco *UniProtKB/SWISS-PROT* (<https://www.uniprot.org/>) (CONSORTIUM, 2018), com 469.154 sequências. Como parâmetros foram utilizados no mínimo 50 % de cobertura e  $10^{-10}$  de *e-value*, além de uma análise manual do alinhamento, onde foi verificado tamanho e cobertura da sequência para que fosse possível fazer uma análise mais detalhada sobre presença de sítio de substrato/ativo, permitindo uma melhor seleção dos possíveis candidatos à rota da biossíntese dos alcaloides de *Erythrina*, bem como possíveis polimorfismos. Todas as transcrições foram quantificadas como número de transcrições por milhão (TPM) pelo software *Salmon* (PATRO et al., 2017).

## 5 RESULTADO E DISCUSSÃO

A escolha da espécie *Erythrina velutina* para análise do transcriptoma, proteoma e perfil metabólico deu-se principalmente pelo interesse em ter um estudo a nível molecular de uma planta nativa e medicinal do Brasil, sobretudo oriunda do bioma Caatinga. Isso contribui com o entendimento em uma nova perspectiva de como as plantas se comportam neste cenário edafoclimático peculiar. Além disso, *Erythrina velutina* dispõe de outras contribuições por ser uma espécie produtora de flavonoides e alcaloides, ambos metabólitos especializados com várias atividades biológicas (ALMEIDA et al., 2020; RAMBO et al., 2019; FAHMY et al., 2019). Através dos resultados deste trabalho, genes e proteínas candidatos, potencialmente envolvidos na biossíntese de alcaloides benzilisoquinolínicos são indicados, bem como a expressão gênica e proteica associados com a acumulação de alcaloides em diferentes partes vegetais serão apresentados.

### 5.1 Dados edafoclimáticos e análise de solos coletados

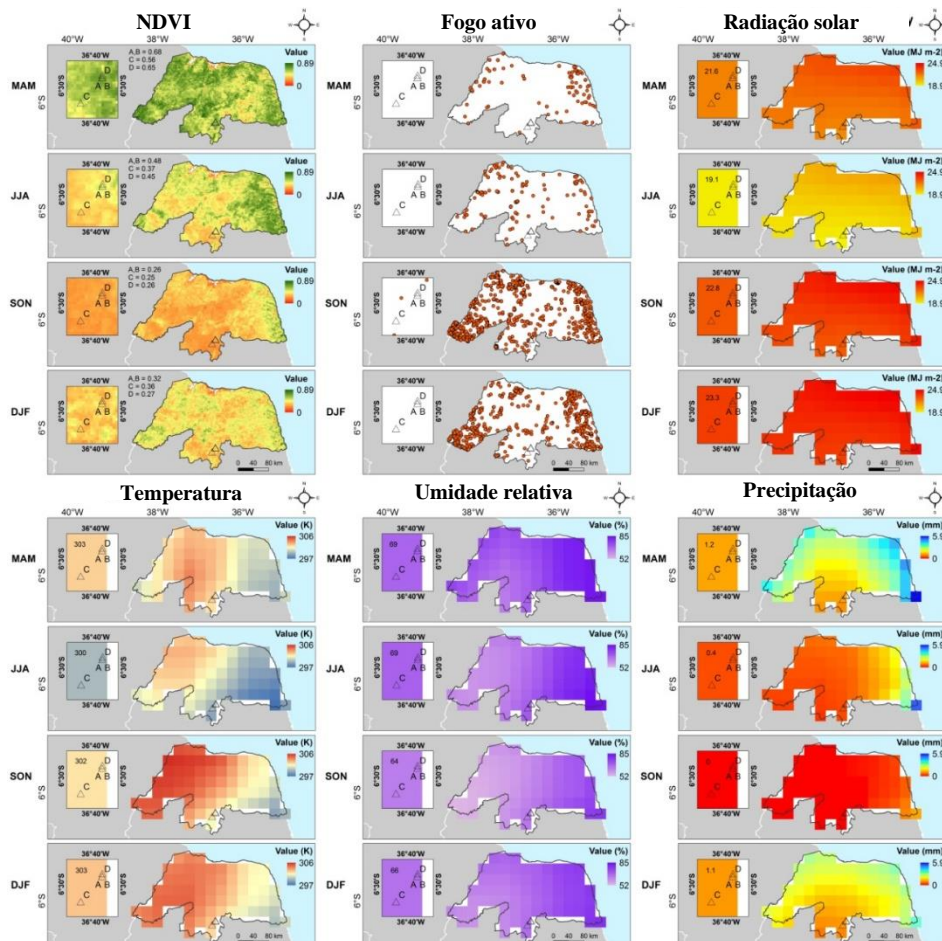
As plantas em seu ambiente natural são um desafio para a investigação ômica, visto que estão constantemente sendo influenciadas por fatores bióticos e abióticos em condições não controladas (habitat natural), resultando em especializações e interações complexas. Visando contribuir para a contextualização dos dados experimentais gerados neste trabalho, foi realizada uma caracterização dos parâmetros climáticos do estado do Rio Grande do Norte, principalmente focados nos pontos de coleta e análise parcial do solo. A Figura 6 mostra mapas trimestrais de março de 2017 a fevereiro de 2018 referentes à temperatura, fogo ativo, precipitação total e radiação solar nos 12 meses antecedentes em que as plantas foram coletadas em seus habitats naturais.

O fogo ativo foi detectado principalmente no semestre anterior, principalmente nos meses de setembro, outubro e novembro de 2017, de forma mais expressiva na região leste e oeste do estado e pouco detectado no semestre que foi realizada a coleta do material vegetal. A menor umidade relativa e um NDVI (Índice de Vegetação por Diferenças Normalizadas) próximo a 0,30 foram detectados. O parâmetro NDVI é importante porque reflete a planta saudável mediante a análise da energia absorvida e refletida em uma região espectral que mostra as condições das estruturas celulares. Portanto, no semestre anterior à colheita, o pequeno NDVI é coerente com as características da Caatinga e pode indicar uma senescência inicial das principais árvores. Na época da coleta, a espécie *Erythrina velutina* apresentava simultaneamente frutos e folhas, no entanto, 15 dias após o dia da colheita, nenhum fruto foi

encontrado e a senescência das folhas começou. Esta planta em comparação com outras árvores da Caatinga, mostra um maior porte e um atraso na perda de folhas na estação seca. As temperaturas mais altas foram detectadas em setembro, outubro e novembro de 2017 principalmente na região oeste e central do estado, e a maior radiação em dezembro de 2017, janeiro e fevereiro de 2018, e essa alta incidência de radiação solar ocorreu em todo o estado. Finalmente, o nível de precipitação acumulada no trimestre anterior à colheita foi próximo de 1,1 mm.

Ambas as regiões (Jardim do Seridó e Acari) fazem parte da microrregião de Seridó, localizada no estado do Rio Grande do Norte. Esse subespaço é caracterizado pelo clima semiárido do bioma Caatinga, com secas prolongadas, um relevo com formações antigas, solos facilmente erodíveis e um sistema hidrológico com baixo potencial (BEZERRA JÚNIOR e SILVA (2007). Dessa maneira, as características edafoclimáticas apresentadas no contexto deste trabalho, o qual aborda amostras oriundas de ambiente natural, compõem um caráter descritivo do estudo.

**Figura 6** - Mapa do Rio Grande do Norte (Brasil) e características climáticas monitoradas



O Monitoramento climático do estado do Rio Grande do Norte no bioma da Caatinga durante o período de 1 ano (dividido por trimestre). Os pontos ABCD representam o campo de colheita das quatro populações de *Erythrina velutina* utilizadas neste estudo. MAM – março, abril e maio; JJA – junho, julho e agosto; SSN – setembro outubro e novembro; DJF – dezembro, janeiro e fevereiro. Fonte: Elaborado pelo autor (2020).

Além das características climáticas, é importante conhecer as propriedades do solo para cada ponto de coleta, a fim de compreender e contextualizar a interação planta-solo-clima. Análises químicas foram desenvolvidas (tabela 2) para caracterização do solo.

**Tabela 2** - Solo coletado de aproximadamente 0 a 20 cm de camada superficial em áreas da vegetação da Caatinga, na estação seca

Parâmetros	Ponto A	Ponto B	Ponto C	Ponto D
pH em água (1:2,5)*	6,0	5,2	7,9	5,7
Cálcio (cmol <sub>c</sub> /dm <sup>-3</sup> )	1,85	1,80	18,38	1,65
Magnésio (cmol <sub>c</sub> /dm <sup>-3</sup> )	1,23	2,0	3,50	1,60
Alumínio (cmol <sub>c</sub> /dm <sup>-3</sup> )	0,0	0,69	0,0	0,0
Hidrogênio e alumínio** (cmol <sub>c</sub> /dm <sup>-3</sup> )	3,68	4,08	0,0***	2,29
Fósforo (mg/dm <sup>-3</sup> )	136	13	37	68
Potássio (mg/dm <sup>-3</sup> )	83	117	366	183
Sódio (mg/dm <sup>-3</sup> )	27	75	77	49
Ferro (mg/dm <sup>-3</sup> )	24,44	117,25	2,0	39,90
Zinco (mg/dm <sup>-3</sup> )	1,44	1,83	0,99	2,20
Cobre (mg/dm <sup>-3</sup> )	0,28	0,21	0,19	0,28
Manganês (mg/dm <sup>-3</sup> )	9,85	14,75	8,23	11,91
Níquel (mg/dm <sup>-3</sup> )	0,15	0,35	0,60	0,52
Chumbo (mg/dm <sup>-3</sup> )	0,25	0,20	0,95	0,25
Cádmio (mg/dm <sup>-3</sup> )	0,0	0,0	0,10	0,05
Cromo (mg/L)	0,0	0,0	0,25	0,0

\*Proporção de solo e água destilada (a cada 1g de solo, adiciona-se 2,5mL de água destilada para determinação do pH) \*\*O hidrogênio extraível corresponde à diferença entre a acidez potencial e o alumínio trocável do solo. \*\*\*O hidrogênio extraível é calculado apenas em amostras com pH em água abaixo de 7,0. Fonte: elaborado pelo autor.

As amostras de solo foram coletadas em quatro pontos próximos aos indivíduos de *Erythrina velutina* utilizados para coleta do material vegetal. Cada ponto de coleta possui características peculiares, duas em Jardim do Seridó (pontos A e D) e os outros dois pontos em Acari (pontos B e C). A análise do solo foi realizada no Laboratório de Análises do Solo da Companhia de Pesquisa Agropecuária do Rio Grande do Norte S/A (EMPARN).

Os micronutrientes estão envolvidos em vários processos metabólicos disseminados entre o metabolismo primário e especializado, defesa celular, transdução de sinal, percepção hormonal e regulação gênica (VATANSEVER, OZYIGIT, FILIZ, 2017). Muitos elementos metálicos são importantes para sistemas redox, como ativadores de enzimas, incluindo as envolvidas na síntese de lignina e na reação de radicais superóxido, bem como na síntese de proteínas e do fito-hormônio ácido indol-3-acético (IAA) (BROADLEY et al, 2012).

Sabe-se que a Caatinga apresenta sedimentos pedregosos e superficiais com afloramentos muito extensos de rochas maciças, processos erosivos, com predomínio de intemperismo físico ao longo da formação geológica por milhões de anos na região Nordeste do Brasil (PRADO, 2003).

Uma análise completa para avaliação da fertilidade do solo deve incluir as seguintes determinações: pH, fósforo, potássio, cálcio, magnésio, enxofre, zinco, manganês, cobre, ferro, boro, alumínio, hidrogênio mais alumínio, teor de matéria orgânica e granulometria. As plantas são capazes de absorver nutrientes em formas iônicas dissolvida na fase aquosa que compõe o solo, assim, para que a planta absorva o nutriente pela raiz, precisa primeiramente transportar ele até superfície radicular, onde a disponibilidade do nutriente e o fator transporte depende dos aspectos do solo que afetam a mobilidade de íons (FURTINI NETO et al., 2001).

O pH é um índice que fornece o grau de acidez ou alcalinidade do solo, trata-se de um índice que no caso dos solos é utilizado como indicativo das condições gerais de fertilidade (VILLAR, 2007). De acordo Alvarez, Guimarães e Ribeiro (1999) há duas classificações no que se refere ao pH, a primeira é a classificação química, em que os pontos A, B e D tem uma acidez média, enquanto que o ponto C (um dos pontos de Acari) possui uma alcalinidade elevada e a classificação agrônômica, em que os pontos A e D tem um bom parâmetro de pH, diferente dos pontos B e C que tem pH baixo e alto respectivamente. Vale acrescentar que, por mais que o pH do solo na faixa de 6 e 7 contribuam para uma maior disponibilidade de nutrientes na planta, esse parâmetro varia muito entre os solos (FURTINI NETO et al., 2001).

Furtini Neto (2001) descreveu os parâmetros considerados como referência para análise do solo onde foram comparados com alguns parâmetros dos nossos resultados. Em Jardim do Seridó, o ponto A tem a concentração de magnésio, manganês, ferro e potássio considerada ótima, enquanto que o cálcio e o zinco com uma concentração mediana, e o cobre em concentração muito baixa. Já o ponto D tem um solo com magnésio e potássio em concentrações dentro dos parâmetros considerados bons, o cálcio em concentrações medianas, o zinco está presente em altas concentrações e o cobre está presente em concentrações muito baixas neste solo. Em Acari, no que se refere ao ponto B, o solo possui magnésio, zinco e potássio em uma concentração ótima, considerada “muito boa”, o cálcio mediano, o manganês e o ferro está no solo em concentrações muito altas, o cobre em concentrações muito baixas e o alumínio presente exclusivamente neste ponto. O ponto C tem concentração de magnésio, cálcio e potássio considerada “muito boa”, manganês tem uma concentração mediana, zinco e ferro baixas concentrações, cobre muito baixo e cádmio é exclusivo do ponto.

Considerando que não foi possível estabelecer uma correlação entre características do solo e os demais resultados de transcritos, proteínas e metabólitos, verifica-se apenas um caráter descritivo desses dados.

## 5.2 Análise do perfil metabólico

Os alcaloides de *Erythrina* apresentam uma diversidade estrutural principalmente em relação à sua porção central, permitindo a classificação entre as classes de lactona, dienoide e alcenoide. A diversidade molecular é atribuída aos padrões de substituição de maneira especial em relação à posição dos substituintes clássicos: OH, CH<sub>3</sub>, CH<sub>3</sub>O e resíduos de glicosil. Essas diferenças estruturais discretas não refletem em diferenças significativas no perfil de fragmentação, o que significa que o espectro de massa entre isômeros de posição ou com diferença na posição de um resíduo CH<sub>3</sub>O, por exemplo, pode gerar o mesmo perfil de fragmentação e, conseqüentemente, a elucidação estrutural inequívoca é prejudicada. Além disso, a molécula ionizada, protonada ou na forma de aduto estava ausente em alguns alcaloides e dificultou a identificação, bem como um número escasso de fragmentos gerado em alguns metabólitos. Mesmo com essas limitações, a análise dos dados de MS/MS permitiu sugerir alguns candidatos à identidade molecular dos alcaloides de *Erythrina velutina*.

Mediante a estratégia analítica de HR-LC-MS/MS (*high resolution*), os resultados foram obtidos utilizando o banco de dados espectral GNPS (Rede Natural Molecular Social Global) e uma biblioteca complementar construída manualmente com dados espectrais (MS/MS) de quase todos os metabólitos descritos anteriormente para o gênero *Erythrina*, sendo ambos utilizados para comparação espectral, a fim de sugerir uma identidade química para os alcaloides analisados.

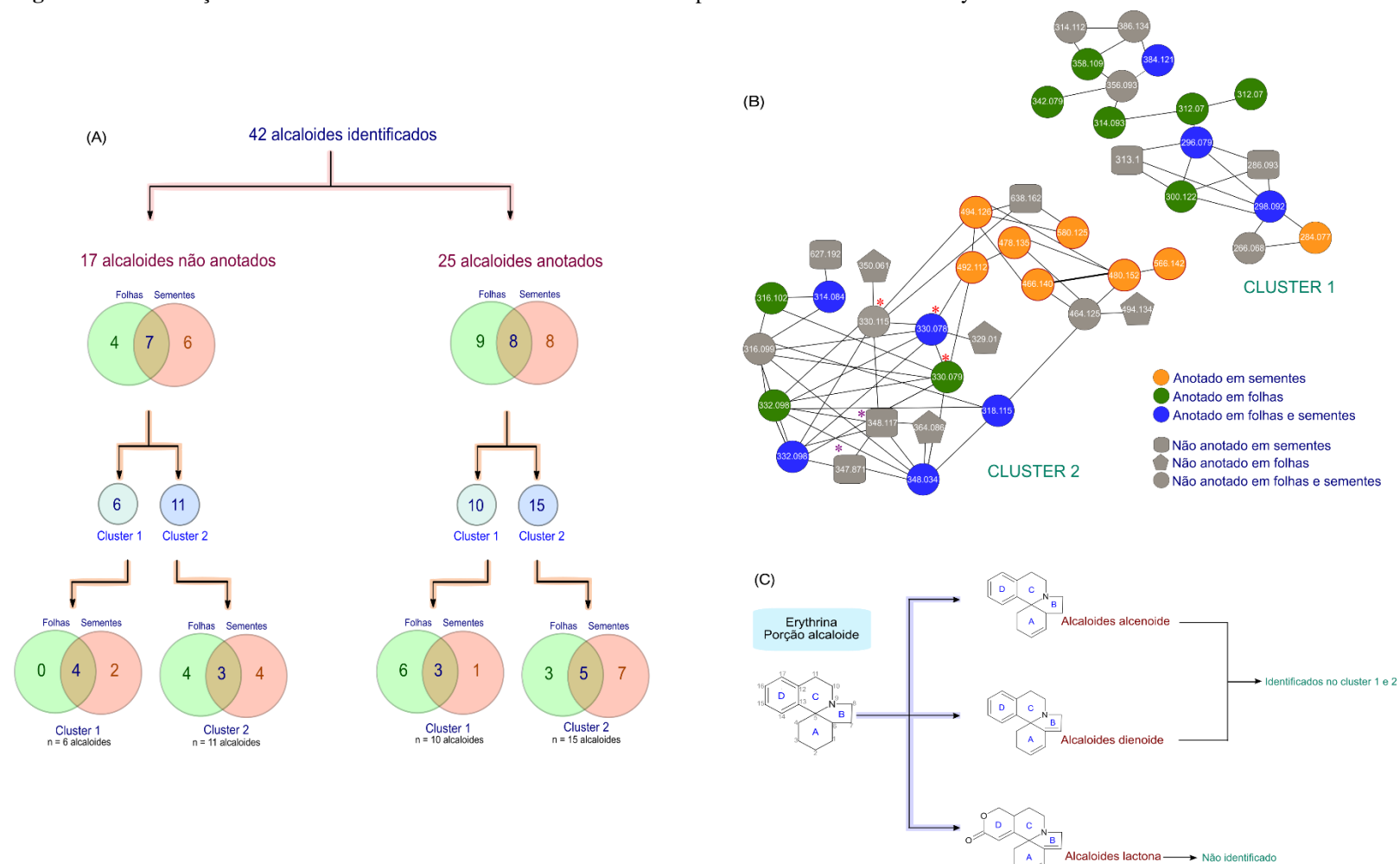
Os resultados revelaram a presença de alcaloides eritrínicos das duas classes relatadas na literatura (dienoide e alcenoide, mas não lactona), tanto em folhas como em sementes por redes moleculares (Figura 7C) compreendendo 42 alcaloides: 25 (59 %) foram anotados e 17 (41 %) foram não anotados. No total dos 25 alcaloides anotados alguns deles possuem mais de uma possibilidade estrutural, porque identificamos os substituintes, mas não o padrão de substituição (detalhes na figura 8 e 9 e Tabela S1 e S2 do anexo I).

As redes moleculares (*molecular networking*) constituem novas abordagens na qual os dados da espectrometria de massas são agrupados baseados no padrão de fragmentação, formando famílias de moléculas e contribuindo na identificação do perfil químico. De acordo com nossos resultados e pela semelhança estrutural, foi gerado dois *clusters*, um representando

os alcaloides principalmente da classe dienoide e outro *cluster* com alcaloides principalmente da classe alcenoide, selecionados e inserido em cada rede de acordo com o padrão de fragmentação (Tabela S8).

Entre os alcaloides anotados propostos, 8 estão em ambos as partes vegetais da planta, 8 são exclusivos de sementes e 9 exclusivos de folhas. A descrição estrutural e a indicação em qual parte da planta (de acordo com  $m/z$  o tempo de retenção – RT) o alcaloide pode ser visualizado na rede molecular e material suplementar. Entre os alcaloides não anotados, 4 são exclusivos de folhas, 6 exclusivos de sementes e 7 estão presente em ambas as estruturas vegetais (figura 7A).

**Figura 7** – Distribuição dos alcaloides e rede Molecular MS/MS obtido a partir de extratos bruto de *Erythrina velutina*



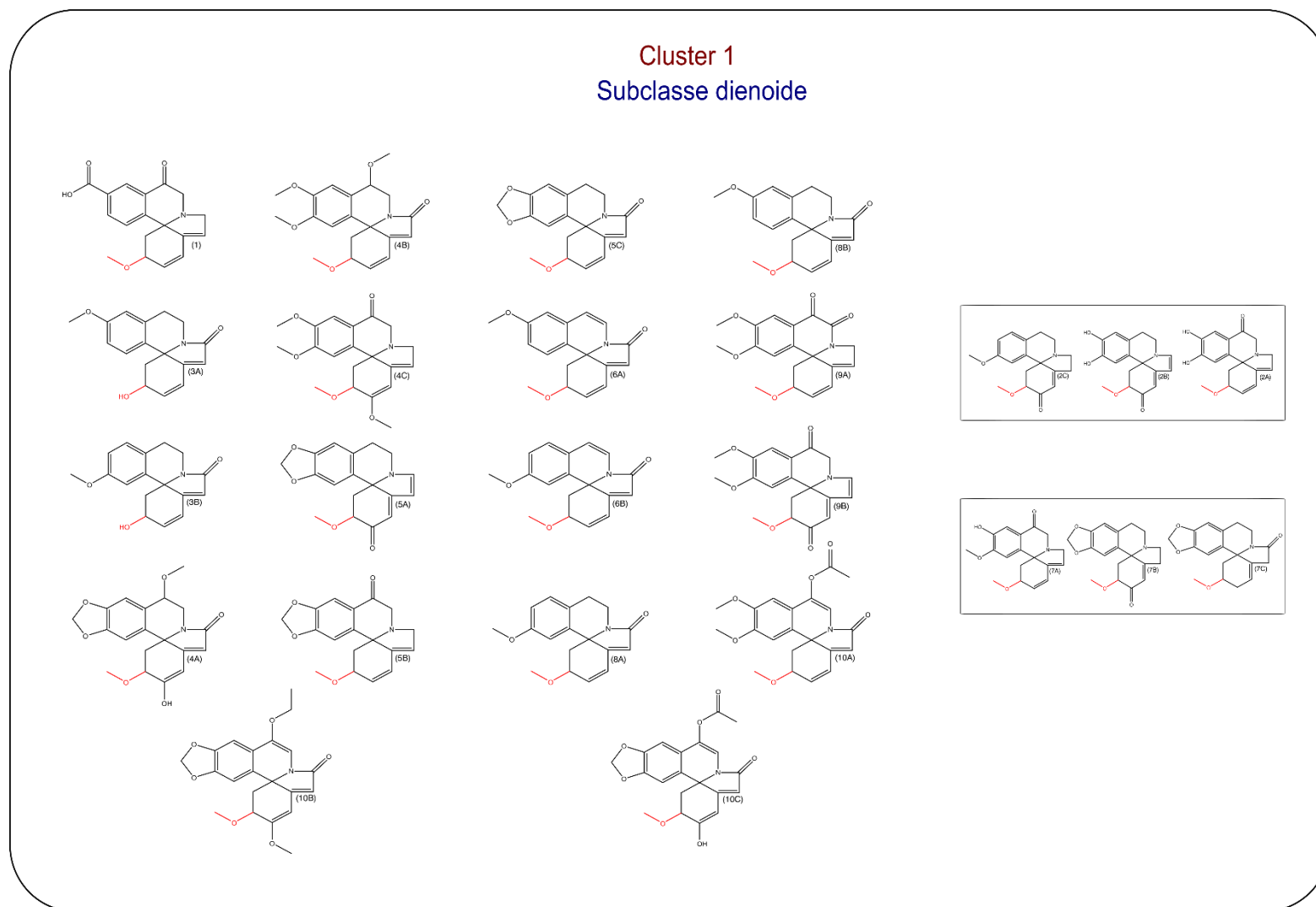
(A) Distribuição dos alcaloides nas estruturas vegetais; (B) Rede de moléculas onde os nós são coloridos de acordo com o tecido da planta para os alcaloides anotados (nós em laranja – alcaloides presentes apenas em sementes; nós em verde - alcaloides presentes apenas em folha e nós em azul – alcaloides presentes em folhas e sementes); nós em cor cinza representa alcaloides detectados, mas não anotados para os alcaloides não anotados: quadrado, pentágono e círculo representam as estruturas vegetais (exclusivamente semente, folhas e folhas e sementes em comum) em que o alcaloide está presente e o número localizado dentro do nó representa a razão m/z detectada. Legenda: \* valores de m/z que foram repetidos no cluster (podem representar isômeros). Fonte: elaborado pelo autor (2020).

Conforme apresentado, foram identificadas 10 estruturas anotadas (1-10 – Tabela S1) com esqueleto químico dos alcalóides principalmente da classe dienoide no cluster 1 e 15 moléculas anotadas (11-25 – Tabela S1) com estrutura principalmente da classe alcenoide no cluster 2.

Algumas características gerais dos alcalóides anotados podem ser destacadas em relação à estrutura e distribuição de órgãos. Alcalóides exclusivos da semente apresentaram substituintes hidroxila em C3, C11 e, em menor grau, em C2, enquanto C3 também poderia apresentar uma O-metilação (ver numeração na Figura 7C). Diversas moléculas apresentaram-se glicosiladas (ver material suplementar do anexo I), no total foram identificadas sete sugestões de alcalóides ligados a carboidratos, foram eles: (11), (12), (15), (19), (20), (21), e (22), e apenas os alcalóides do cluster 2 têm suas formas glicosiladas, curiosamente, esse fato ocorre de forma muito mais abundante em semente que em folhas, e, até onde sabemos, a estrutura proposta é inédita. Os resíduos de glicosil frequentemente estavam ligados em C15 e C16; caso contrário, essas posições poderiam estar envolvidas em uma ponte metilenodioxí (R-O-CH<sub>2</sub>-O-R).

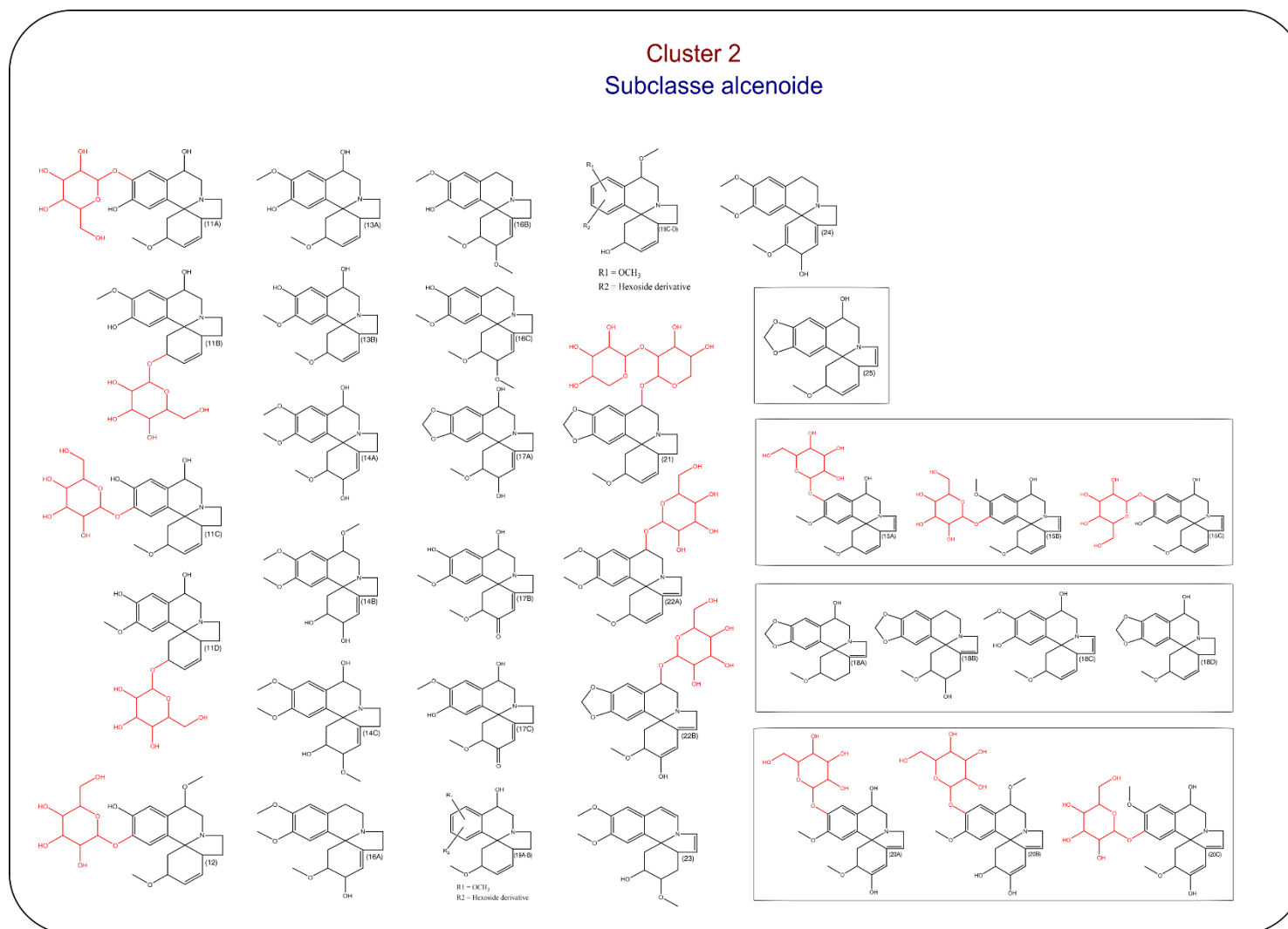
Alcalóides encontrados e anotados apenas nas folhas compartilham a característica de O-metilação em C3, e a forma oxidada (C = O) é predominante em C11 e, em menor extensão, em C2. Nos alcalóides compartilhados por ambos os órgãos, as principais características incluíam a predominância de OH em C2, um padrão de O-metilação em C3, enquanto C15 e C16 estavam principalmente envolvidos em um anel metilenodioxí.

**Figura 8** - Características químicas estruturais dos alcaloides do cluster 1 identificados em amostras de folhas e sementes com análise de LC-MS/MS (modo de ionização positiva para amostras de sementes e folhas de *Erythrina velutina*)



Fonte: elaborado pelo autor (2020).

**Figura 9** - Características químicas estruturais dos alcaloides do cluster 2 identificados em amostras de folhas e sementes com análise de LC-MS/MS (modo de ionização positiva para amostras de sementes e folhas de *Erythrina velutina*)



Fonte: elaborado pelo autor (2020).

A produção de metabólitos secundários é resultado de complexas interações entre biossíntese, transporte, armazenamento e degradação. Alcaloides são frequentemente acumulados em partes específicas de plantas medicinais. Embora se saiba que alguns alcaloides são transportados das partes de origem para essas partes específicas, pouco se sabe sobre como os alcaloides são transportados por meio das membranas e, por fim, se acumulam em organelas como o vacúolo do parte coletor (N. SHITAN; K. YAZAKI, 2007).

Entender o fluxo metabólico é um conhecimento extremamente importante no que diz respeito para incrementar estratégias para aumento de produção de alcaloides. Até o momento não há informações acerca de caminhos ou padrão de translocação de alcaloides eritricos em estruturas vegetais, entretanto, estudos relatam que alcaloides pode ser acumulados em vacúolos, como os alcaloides indólicos terpenóides em *Catharanthus roseus* (CARQUEJEIRO et al., 2013) utilizando um sistema antiportal de prótons específico; alcaloides da espécie *Nicotiana* (MORITA et al 2009), com acumulação induzida por jasmonato 1 (Nt-JAT1), atuando como transportador secundário, e, em destaque alcaloides benzilisoquinolínicos em *Coptis japonica* (OTANI et al., 2005) translocando alcaloides como a beberina mediante o transportador CjMDR1 da raiz para o rizoma, sendo armazenados em vacúolos.

Dessa maneira, acreditamos que após biossíntese tanto em folhas como em sementes, os alcaloides eritricos identificados são translocados entre as estruturas vegetais e podem ter como alvo final da acumulação em estruturas vacuolares, sobretudo os alcaloides identificados como glicosilados, visto que os vacúolos também estão relacionados principalmente com reservas de moléculas polares. Ainda com relação aos alcaloides ligados a porção glicosil, acreditamos que no momento pode se estabelecer duas hipóteses diante os resultados obtidos até aqui: há uma maior disponibilidade de proteínas que realizem reações de glicosilação nas sementes do que em folhas, dado que 100% dos metabólitos glicosilados identificados e anotados estavam presente nessa estrutura (1); ou há disponibilidade enzimática de forma similar de proteínas que incorporam o grupamento glicosil entre folhas e sementes, entretanto, em momentos adversos em que a planta precise como forma de proteção ativar mecanismos de defesa (que pode dar-se por vários estímulos diferentes) e que precisa da translocação dos alcaloides de sementes para folha neste momento pode perder o grupamento glicosil por reações químicas enzimáticas ainda não estabelecidas (2).

Nenhuma proteína transportadora ou que realize reações de glicosilação foram relatadas até o momento para a espécie *Erythrina velutina*. Portanto, o próximo passo consiste em investiga-las para desvendar como se dá o mecanismo de transporte dos alcaloides eritricos entre folhas e sementes, bem como um estudo mais detalhado acerca da localização intracelular

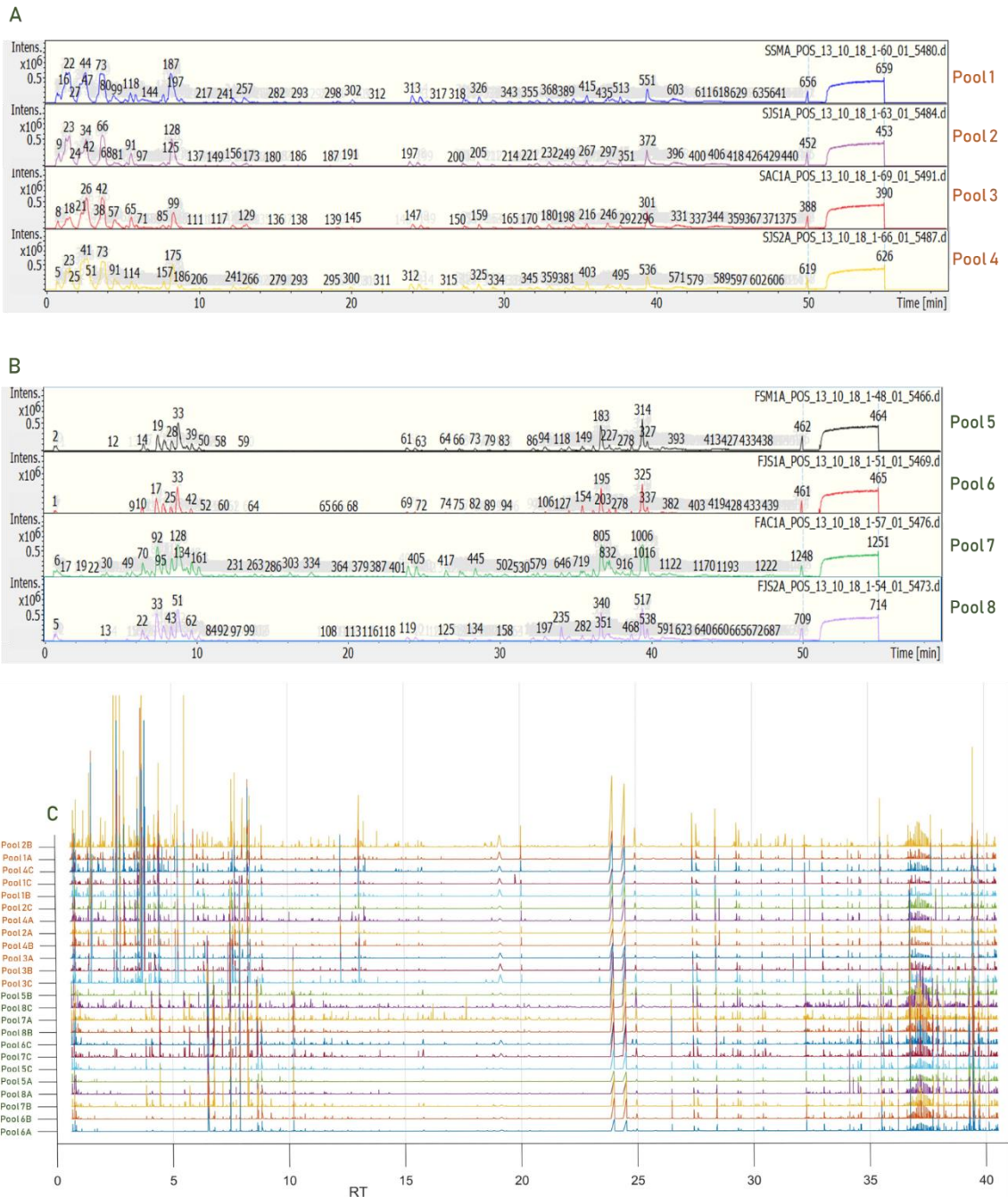
dos metabolitos, identificando marcadores que possam contribuir para o aumento do fluxo de alcaloides.

A análise do cromatograma pela técnica HR-LC-MS/MS revela um perfil de homogeneidade entre as amostras de extratos de sementes (*pool1-pool4* - figura 9A) e entre as amostras dos extratos de folha (*pool5-pool8* – figura 9B), sugerindo o mesmo perfil qualitativo de alcaloides entre as 4 amostras de cada estrutura vegetal e em cada triplicata (figura 9C).

Por exemplo, no cromatograma em todos os *pools* de sementes (1-4) as substâncias majoritárias foram os picos correspondente aos tempo de retenção de 1 a 10 min, destes, identificamos alcaloides da classe dienoide/alcenoide entre 1.38 a 8.25 min, enquanto que os alcaloides da classe lactona em sementes anotados estiveram entre os tempo de retenção que variava de 5.8 a 13,66 min. No que se refere a folhas, a análise do cromatograma revelou em todos os *pools* (5-8) que as substâncias majoritárias foram representadas nos picos correspondente aos tempo de retenção de 5 a 10 min, destes, identificamos alcaloides da classe dienoide/alcenoide entre 2.44 a 8.25 min, enquanto que os alcaloides da classe lactona em folhas anotados estiveram entre os tempo de retenção que variava de 3.92 a 13,66 min.

Desse modo, os dados da espectrometria de massa obtido neste estudo indicam a detecção de 42 alcaloides, os quais em sua maioria foram identificados e anotados. Observou-se também o perfil homogêneo entre as triplicatas das 4 amostras folhas e entre as triplicatas das 4 amostras sementes, os quais continham substancias majoritárias e minoritárias. As etapas seguintes desse estudo baseiam-se em identificar os alcaloides que foram não anotados, bem como determinar estruturas químicas mais precisa com o auxílio de técnicas complementares.

**Figura 10** - Cromatograma dos extratos brutos de folhas e sementes de *Erythrina velutina*



Em A representa o cromatograma dos quatros pools de sementes; em B representa o cromatograma dos 4 pools de folhas; em C – cromatograma do *pool1* ao *pool8*, onde A/B/C representa a triplicada de cada ponto de coleta (*pool*); eixo X representa o tempo de retenção em minutos durante a análise e o eixo Y corresponde a intensidade relativa de cada amostra (em 7A e 7B) e identificação da amostra (em 7C).

Entre um grande número de metabólitos secundários das plantas, os alcaloides são um dos grupos mais importantes devido às suas atividades biológicas marcantes e divergentes e alguns são aplicados para uso clínico. Cerca de 20 % das espécies vegetais acumulam

alcaloides, no entanto, ocorrem em muitas classes de organismos vivos como compostos naturais diversos, abundantes em número e forma (FACCHINI, 2001).

Curiosamente, os alcaloides isolados das folhas de *Cocculus laurifolius*: cocculinina, isococculinina e coccludinona são referidos como alcaloides eritrínicos incomuns da ordem Ranunculales e da família Menispermaceae. Esses alcaloides exibem certa estereoespecificidade para o precursor (*S*)-norprotosinomenina (BHAKUNI; JAIN, 1980; BHAKUNI; SINGH, 1978), demonstrando que Barton et al (1968) não estava totalmente equivocado no que se refere aos percussores dos alcaloides, que embora taxonomicamente distantes, eles compartilham várias características estruturais com os alcalóides *Erythrina* da família Fabaceae, dessa forma, existem várias lacunas a serem preenchidas e, neste estudo, tentamos contribuir para o conhecimento sobre a espiroamina tetracíclica BIA.

Aqui propomos a via de biossíntese de alcaloides encontrados no gênero *Erythrina*, tendo como um dos precursores a noramurina e não (*S*)-norprotosinomenina. Assim, propomos uma segunda rota biossintética para a formação de alcaloides coculínicos, o que nos leva a intensificar a importância da contribuição da quimiotaxonomia e a intensificar a investigação desses alcaloides peculiares por ferramentas moleculares.

Os resultados obtidos com *E. velutina* indicam que folhas e sementes são produtores de alcaloides, no entanto, os mecanismos que regulam o armazenamento deles ainda são desconhecidos.

### 5.3 Análise do transcriptoma

#### 5.3.1 Montagem *de novo* do transcriptoma

Após o sequenciamento das bibliotecas de cDNA, os arquivos *FastQC* foram gerados para realização da análise no Cluster Computacional do Núcleo de processamento de alto desempenho (NPAD) da UFRN. Inicialmente, foi realizada a remoção do RNAr usando o software SortMeRNA (v2.1) para todas as leituras que combinam os fios Forward (R1) e Reverse (R2). Essa etapa foi necessária porque a ausência de RNAr pode melhorar a qualidade das leituras. Montamos o transcriptoma e obtivemos sequências de alta qualidade (*pool1* - 27.465.395; *pool2* - 34.199.767; *pool3* - 20.856.947; *pool4* - 31.730.471; *pool5* - 30.967.679; *pool6* - 28.741.073; *pool7* - 26.473.356; *pool8* - 42.558.990). Em todos os pools, essas leituras têm pares de bases com 20 ou mais escores de qualidade e as leituras têm uma pontuação média de *Phred* acima de 33. Os escores médios de qualidade por sequência e a qualidade média por base podem ser visualizados na tabela 3, onde inclui o número de leituras brutas antes e depois de montar as sequências das amostras de folhas e sementes de *Erythrina velutina*.

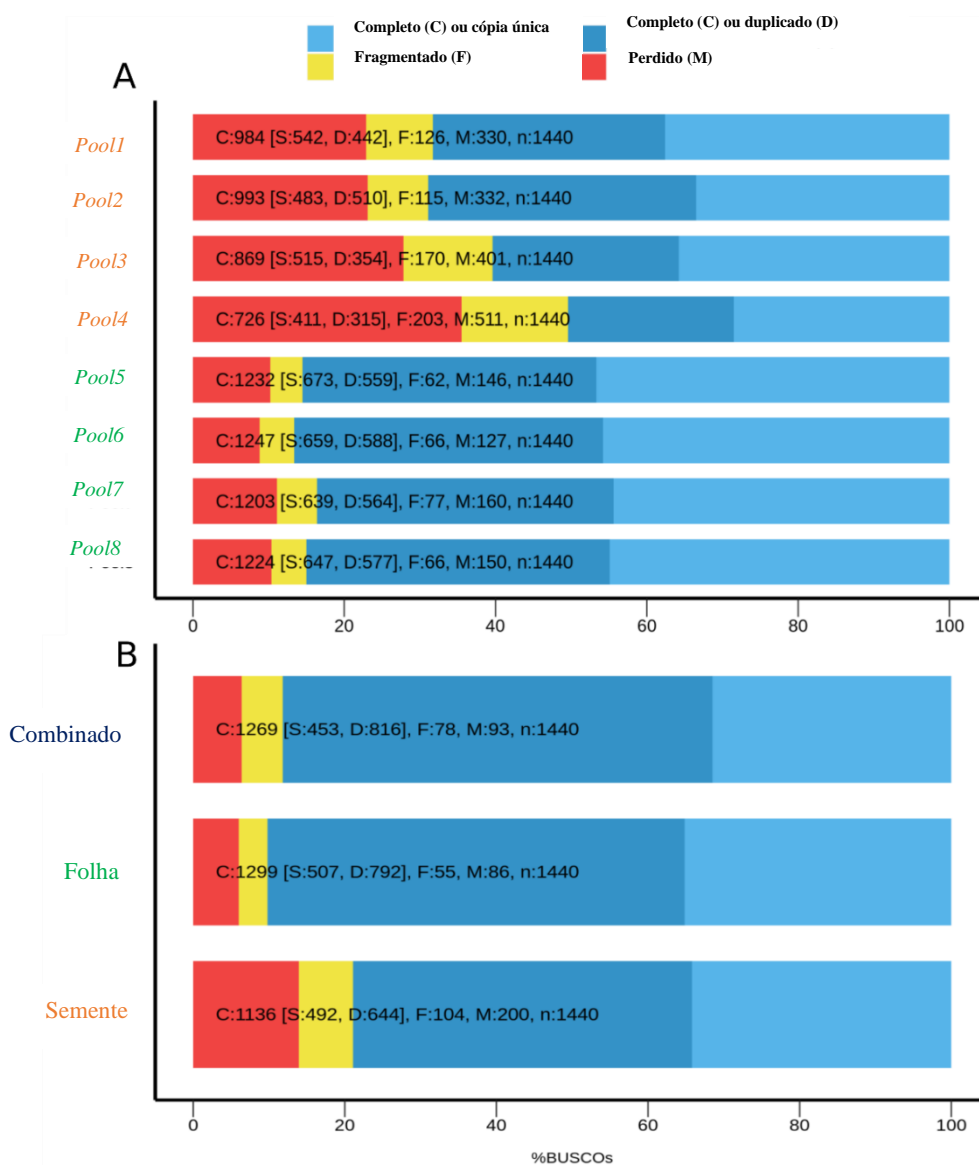
**Tabela 3** - Visão geral da análise e montagem do transcriptoma

Localização	A	B	C	D	A	B	C	D
	<i>Pool 1</i>	<i>Pool 2</i>	<i>Pool 3</i>	<i>Pool 4</i>	<i>Pool 5</i>	<i>Pool 6</i>	<i>Pool 7</i>	<i>Pool 8</i>
<b>Sequencias</b>								
Número de <i>reads</i>	38,517,758	45,536,735	34,385,125	50,105,977	42,873,185	43,766,448	38,583,198	57,391,786
<i>Reads</i> de alta qualidade	27,465,395	34,199,767	20,856,947	31,730,471	30,967,679	28,741,073	26,473,356	42,558,990
Tamanho das leituras (pb)	21-151	21-151	21-151	21-151	21-151	21-151	21-151	21-151
Tamanho médio das leituras (pb)	100	101	101	99	101	102	101	101
<b>Contigs</b>								
Número de <i>contigs</i>	71,747	91,072	68,050	68,804	88,190	93,207	87,819	95,089
Tamanho médio dos <i>contigs</i> (pb)*	1088.18	1102,95	984,88	871,50	1101.2	1096,8	1036,68	1085,99
N50 (pb)**	1606	1700	1535	1370	1683	1697	1596	1648

Cada pool corresponde a um tipo de amostra (sementes ou folhas) de 3 indivíduos agrupados, localizados em um ponto específico de coleta. No ponto A (*pool 1* - sementes / *pool 5* - Folhas); B (*pool 2* - sementes / *pool 6*- folhas); C (*pool 3* - sementes / *pool 7* - folhas); D (*pool 4* - sementes / *pool 8* - folhas). A e B correspondem aos municípios de Acari - RN e C e D aos municípios de Jardim do Seridó – RN. Asterisco (\*): tamanho médio das sequências montadas após sobreposição dos *reads*; (\*\*): mediana do comprimento das montagens após formação dos *contigs*. Fonte: elaborado pelo autor (2020).

Para fornecer medidas quantitativas sobre a avaliação da montagem do transcriptoma foi utilizado o software BUSCOv3. Os resultados dos diferentes pools que construímos a partir de nossos dados usando o software *Trinity* são mostrados na Figura 11. A Figura 11A mostra os resultados da avaliação de bibliotecas individuais em relação à comparação entre nossos transcritos e a figura 11B mostra o resultado das bibliotecas combinadas por partes (sementes e folhas), bem como uma representação do total de transcritos analisados (sementes e folhas como um transcriptoma global).

**Figura 11** - Resultado da avaliação do BUSCOv3



A: Resultados da avaliação BUSCO para bibliotecas individuais de cDNA em relação à comparação entre nossos transcritos em um banco de dados de genes ortólogo de outras espécies vegetais. As amostras nomeadas como pool 1 a pool 4 são de bibliotecas individuais de sementes de *Erythrina velutina*. Amostras nomeadas como pool 5 a pool 8 são das bibliotecas individuais de folhas. B: Resultados da avaliação BUSCO para

bibliotecas combinadas - Transcrições de todas as amostras de *Erythrina velutina* (*pool* 1 a 8) representadas juntas e transcrições representada de forma combinada de folha (*pool*5 a *pool*8 juntas) e sementes (*pool*5 a *pool*8 juntas). Eixo X: representa as bibliotecas de cDNA (individual ou combinada); Eixo Y: representa a porcentagem de sequências completas, fragmentadas e perdidas em nossas análises. Fonte: elaborado pelo autor (2020).

Como resultado, verificou-se a presença de 1.440 ortólogos no conjunto de dados após montagem dos dados de sequenciamento em folhas e sementes de *Erythrina velutina*.

No que se refere a bibliotecas representadas individualmente (figura 11A) de semente, pelo menos 60 % dos ortólogos eram sequências completas, seja de cópia única ou duplicada (*pool*1 – 68 %; *pool*2 – 68 %; *pool*3 – 60 %; *pool*4 – 50 %), as sequências ortólogas fragmentadas representaram 8 %, 8 %, 12 % e 14 % para o *pool*1, *pool*2, *pool*3, e *pool*4 respectivamente, enquanto que as sequências que foram perdidas foram 22 %, 23 %, 28 % e 35 % para o *pool*1, *pool*2, *pool*3, e *pool*4. Ainda a respeito das bibliotecas representadas individualmente e no que trata de amostras de folhas, pelo menos 83 % dos ortólogos eram sequências completas, seja de cópia única ou duplicada (*pool*5 – 85 %; *pool*6 – 86 %; *pool*7 – 83 %; *pool*8 – 85 %), as sequências ortólogas fragmentadas representaram 4 %, 4 %, 5 % e 4 % para o *pool*5, *pool*6, *pool*7, e *pool*8 respectivamente, enquanto que as sequências que foram perdidas foram 10 %, 8 %, 11 % e 10 % para o *pool*5, *pool*6, *pool*7, e *pool*8. Os resultados das bibliotecas representadas de forma combinada (figura 9B) indicam que em folhas (*pool*1 a *pool*4 agregado) 90 % das sequências eram completas, 5 % fragmentada e 6 % das sequências foram perdidas. Já em sementes (*pool*5 a *pool*8 agregado), 79 % das sequências eram completas, 7 % fragmentada e 13 % das sequências foram perdidas. Quando analisamos a representação total dos transcritos (sementes e folhas como um transcriptoma global) 88 % das sequências eram completas, 5 % fragmentada e 6 % das sequências foram perdidas.

De modo geral, obtivemos sequências de altas qualidade tanto em folhas como em sementes (figura 11B), pois grande parte delas eram sequências completas ou de cópia única, e poucas sequências perdidas. Embora note-se uma perda maior de sequências nas sementes, todas as amostras de sementes e folhas (*pool*1-*pool*8) estavam dentro dos critérios de integridade (qualidade) de RNAm e bibliotecas de cDNA.

O software *Bowtie2* também foi utilizado como uma ferramenta para alinhar as leituras de sequenciamento com sequências de referência longas para validar as montagens geradas até o momento (tabela S5 no material suplementar do anexo I). Com

esta etapa, foi possível avaliar a porcentagem de leituras emparelhadas realizadas para formar os *contigs* e, portanto, a montagem final para uma melhor avaliação. Os resultados indicam uma taxa interessante de utilização das leituras limpas e de alta qualidade do alinhamento e montagem.

### **5.3.2 Proposta de via de produção dos alcaloides eritrínicos**

#### **5.3.2.1 Plantas produtoras de BIAS consideradas como referência para este estudo**

O avanço das ferramentas “ômicas” ajudou a investigar como as plantas se comportam ou como respondem em determinadas situações quando estimuladas, ajudam a entender melhor as vias de sinalização e a produção do metabolismo primário e, principalmente, o metabolismo especial ou secundário. No entanto, essas análises em organismos não modelo, como a espécie desse trabalho, se tornam um desafio devido à falta de informações prévias em nível de genes e a nível molecular das espécies do gênero *Erythrina*, além de fatores que podem afetar a montagem *de novo* e a dificuldade da anotação funcional (STEIN, 2001; TAQUARY, 2017). Desde os primeiros genomas tem sido difícil analisar dados separadamente dos organismos, e muitas respostas foram respondidas por meio de análises comparativas e evolutivas (PROSDÓCIMI E MOREIRA, 2015).

A análise comparativa baseada na similaridade de genes e proteínas de plantas produtoras de alcaloides benzilisoquinolínicos (BIAs) bem descrita na literatura e em bancos de dados é uma abordagem poderosa para identificar genes e associá-los a uma função específica (HENINNG et al., 2009), especialmente no que diz respeito à contribuição para a quimiotaxonomia, onde os alcaloides desta classe exercem forte influência, pois além de possuir precursores em comum, compartilham intermediários com esqueleto químico semelhante, o que nos leva a refletir ainda mais sobre a ideia de genes compartilhados ou conservados entre as espécies de uma maneira estabelecida principalmente pela cobertura, identidade e *e-value*.

Possíveis genes e proteínas foram indicados em *Erythrina velutina* como fortes candidatos à participação na produção de alcaloides, baseados principalmente em organismos que consideramos como referência neste estudo que também produzem alcaloides derivados da fenilalanina e tirosina: *Papaver somniferum* e *Eschscholzia*

*californica* (BEAUDOIN; FACCHINI, 2014, 2014; IKEZAWA, IWASA, SATO 2007), identificados no banco de dados *Swissprot* (*BLASTx*).

### 5.3.2.2 Prováveis genes e proteínas envolvidos na biossíntese de alcaloides eritrínicos

As ferramentas modernas para o sequenciamento de RNA de 2ª geração e a identificação de proteínas juntamente com perfil de metabólitos lidam com grandes quantidades de dados e integram procedimentos computacionais com dados analíticos. Tecnologias como o sequenciador *Illumina* e estratégias de LC-MS/MS já foram usadas em plantas não-modelo (EOM; NA, 2019; FU et al., 2019; HAGEL et al., 2015) e são importantes porque podem ajudar na descoberta de genes e metabólitos alvos.

Após o sequenciamento, com objetivo de ter maior confiabilidade dos dados, as oito bibliotecas de cDNA de *Erythrina velutina* foram submetidas a duas análises: a primeira contra o banco *INTERPRO*, uma plataforma de acesso a famílias e domínios proteicos que integra informações de vários bancos de dados. Esse tipo de análise ajudou na identificação de regiões funcionais por domínio que pode ser usada para melhor caracterização funcional/estrutural e isso se torna um fator importante na busca de sequências mais específicas, como a via da biossíntese dos alcaloides.

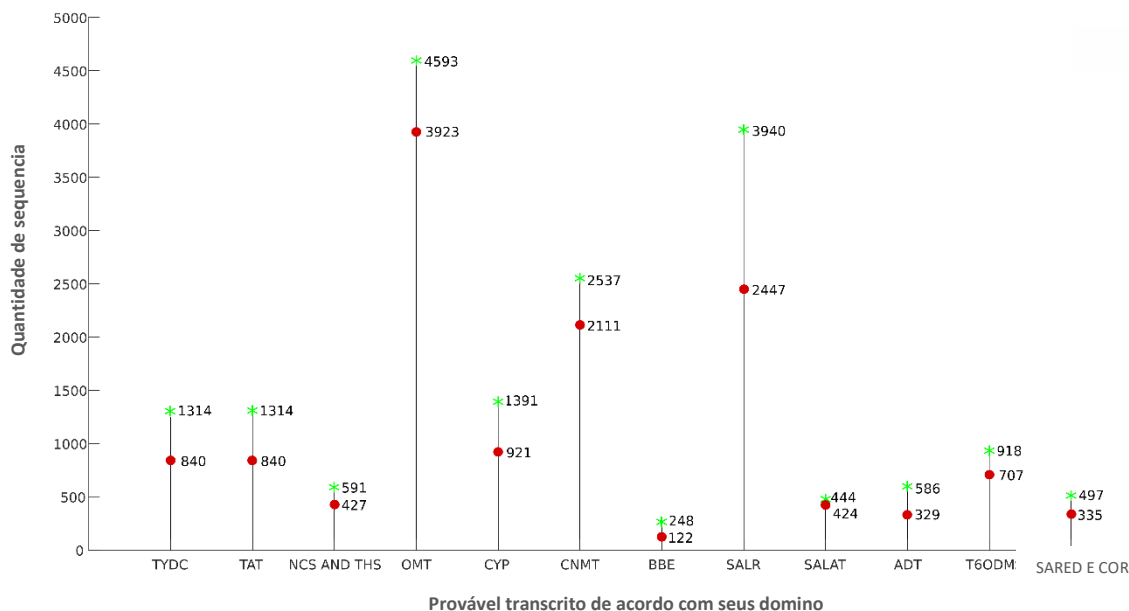
Com relação a essa primeira análise obteve-se 690.879 transcritos no *pool1*; 831.703 transcritos no *pool2*; 625.170 transcritos no *pool3*; 601.449 transcritos no *pool4*; 878.714 transcritos no *pool5*; 879.618 transcritos no *pool6*; 828.667 transcritos no *pool7* e 914.821 transcritos no *pool8*, além de 1.956.400 transcritos diferentes entre os oito pools citados que foram anotados com praticamente todos os tipos de domínios proteicos.

Ainda a respeito, foi realizada uma etapa adicional para seleção de domínios específicos correspondentes aos transcritos candidatos a participarem na via de alcaloides manualmente. Essa busca se deu por meio de pesquisas em banco de dados e informações da literatura. Ao total, foram selecionados 17 transcritos candidatos à rota de biossíntese presentes em outras espécies que produzem alcaloides isoquinolínicos juntamente com todos os seus respectivos domínios funcionais. São eles, respectivamente: **1:** TYROSINE AMINOTRANSFERASE (*TAT*) – AMINOTRANSFERASE, PYRDXLP-DEP\_TRFASE; **2:** TYROSINE/DOPA DECARBOXYLASE (*TYDC*) - AROMATIC\_DEC, PYRDXLP-DEP\_TRFASE; **3:** S-NORCOCLAURINE SYNTHASE (*NCS*) - BET\_V\_I/MLP, START-LIKE\_DOM\_SF; **4:** (RS)-

NORCOCLAURINE 6-O-METHYLTRANSFERASE (*6OMT*) E **5:** (R,S)-RETICULINE 7-O-METHYLTRANSFERASE (*7OMT*) - O-METRFASE\_COMT, PLANT\_METRFASE\_DIMERISATION, SAM-DEPENDENT\_MTASES, WH-LIKE\_DNA-BD\_SF, WH\_DNA-BD\_SF; **6:** (S)-N-METHYLCOCLAURINE 3'-HYDROXYLASE ISOZYME (*CYP80B3*) - CYT\_P450, CYT\_P450\_E\_GRP-I, CYT\_P450\_SF; **7:** 3'-HYDROXY-N-METHYL-(S)-COCLAURINE 4'-O-METHYLTRANSFERASE (*4OMT*) - O-METRFASE\_COMT, PLANT\_METRFASE\_DIMERISATION, SAM-DEPENDENT\_MTASES, WH-LIKE\_DNA-BD\_SF, WH\_DNA-BD\_SF; **8:** NORBELLADINE 4'-O-METHYLTRANSFERASE (*N4OMT*) - SAM-DEPENDENT\_MTASES; **9:** (S)-COCLAURINE N-METHYLTRANSFERASE (*CNMT*) - SAM-DEPENDENT\_MTASES; **10:** RETICULINE OXIDASE (*BBE*) - BBE, FAD-BD\_PCMH, FAD-BD\_PCMH-LIKE\_SF, OXID\_FAD\_BIND\_N, OXY\_OXRDTASE\_FAD\_BS; **11:** SALUTARIDINE SYNTHASE (*CYP719B1*) E **12:** (S)-STYLOPINE SYNTHASE (*CYP719A2*) CYT\_P450, CYT\_P450\_E\_GRP-I, CYT\_P450\_SF; **13:** SALUTARIDINE REDUCTASE (*SALR*) - NAD(P)-BD\_DOM\_SF, SDR\_FAM; **14:** SALUTARIDINOL 7-O-ACETYLTRANSFERASE (*SALAT*) - CAT-LIKE\_DOM\_SF, TRANSFERASE; **15:** THEBAINE 6-O-DEMETHYLASE (*T6ODM*) - DIOX\_N, IPNS-LIKE, OXOGLU/FE-DEP\_DIOXYGENASE; **16:** NADPH-DEPENDENT CODEINONE REDUCTASE (*COR*) - ALDO/KET\_REDUCTASE\_CS, NADP\_OXRDTASE\_DOM; **17:** SANGUINARINE REDUCTASE (*SARED*) - NAD(P)-BD\_DOM, NAD(P)-BD\_DOM\_SF.

Posteriormente, foi feita a busca manual dos domínios selecionados frente ao conjunto de dados do transcriptoma anotado. Foram anotados 13.426 transcritos em sementes que possuíam domínios catalíticos específicos selecionados (domínios citados anteriormente) e 18.373 transcritos anotados em folhas. Os prováveis transcritos correspondentes a metiltransferases (*OMT*), redutases (*SALR*) e citocromo P<sub>450</sub> (*CYP450*) foram os mais abundantes em ambas as partes vegetais, ou seja, foram as classes de domínios funcionais com maior quantidade de transcritos correspondentes. Mais informações estão na figura 12.

**Figura 12** - Quantidade de transcrito baseado em domínio e sua respectiva proteína



Os domínios são uma parte da cadeia polipeptídica que pode de enovelar independentemente para formar uma estrutura compacta e estável. Após anotação de todos os transcritos identificados em nosso conjunto de dados do transcriptoma total foi realizado a busca de transcritos com domínios selecionados e poderia ter envolvimento com a rota de biossíntese. Círculos vermelhos e asterisco verde indicam número de prováveis sequências baseada em domínio e sua respectiva proteína em sementes e folhas, respectivamente; Eixo X: quantidade de transcrito identificado de acordo com o domínio funcional selecionado; eixo Y: provável transcrito identificado. Fonte: elaborado pelo autor (2020).

Com os transcritos anotados com base em seu domínio específicos, realizamos uma segunda análise de identificação que consistia em uma análise local via *BLASTx* usando como parâmetros de no mínimo 50 % de cobertura e  $10^{-10}$  de *e-value* frente ao banco de dados *SWISS-PROT*.

A análise *BLASTx* mostrou que a maioria dos transcritos apresentou similaridade com genes conhecidos no banco de dados e com sequências específicas de espécies consideradas referência deste estudo. Isso foi possível devido a análise adicional mais detalhada dos domínios funcionais, o que foi importante na identificação dos genes que poderiam potencialmente estar envolvidos na biossíntese de BIAs. Como resultado e de acordo com a análise mais detalhada manualmente da sequência, foram selecionados os transcritos: *TYDC*, *CNMT*, *BBE*, *SALR*, *T6ODM* e *COR* com no mínimo 50% de identidade; *NCS*, *CYP719*, *CYP80B*, *4OMT*, *6OMT* e *7OMT* com pelo menos 30% de identidade; *TAT* com 55%; *SALAT* com 40% e *SARED* com 65% (tabela 4).

**Tabela 4** - Transcritos e seus respectivos valores de identidade identificados por BLASTx

Provável transcrito	Identidade (mínimo)
TAT	55 %
TYDC	50 %
NCS	30 %
CYP	30 %
OMT	30 %
CNMT	50 %
BBE	50 %
SALR	50 %
SALAT	40 %
T6ODM	50 %
COR	50 %
SARED	65 %

De acordo com a *International Union of Biochemistry and Molecular Biology* (IUBMB), a classificação por nomenclatura enzimática (código EC - *Enzyme Commission number*) é baseada apenas na reação catalisada distribuídas em seis classes e cada uma das classes está vinculada a uma categoria dividida em subclasses. As seis classes são: oxidorreduções (EC 1); transferases (EC 2); hidrolases (EC 3); liases (EC 4); isomerases (EC 5) e ligases (EC 6). De acordo com nossos achados, identificamos transcritos referentes às classes EC1, EC 2 e EC 4 (tabela 5).

**Tabela 5** - Transcritos identificados de acordo com o nº EC

Nomenclatura enzimática (EC)	Provável transcrito
<b>EC1 - oxidoreductases</b>	<i>COR</i> <i>SALR</i> <i>CYP80B3</i> <i>CYP719B1</i> <i>CYP719A2</i> <i>T6ODM</i> <i>BBE</i> <i>SARED</i>
<b>EC2 - transferases</b>	<i>6-OMT</i> <i>7-OMT</i> <i>4-OMT</i> <i>CNMT</i> <i>TAT</i> <i>SALAT</i>
<b>EC4 - liases</b>	<i>TYDC</i> <i>NCS</i>

Fonte: elaborado pelo autor (2020).

As reações bioquímicas nas quais elétrons são transferidos de uma molécula para outra são catalisadas por uma grande variedade de enzimas denominadas oxidoreductases ou enzimas redox (DIXON AND WEBB, 1979). Para realizar essas reações, as oxidoreductases costumam utilizar cofatores orgânicos, no entanto, há algumas reações em que as enzimas não requerem quaisquer cofatores adicionais e, além disso, é considerada uma das classes enzimáticas bastante importante no que se refere ao setor biotecnológico (CIRINO; ARNOLD, 2002; ELLIS, 2002; IMHOFF et al., 2003; TORRES PAZMIÑO et al., 2010). Para a oxirredução oito transcritos candidatos da rota de biossíntese que tem correspondência dessa classe enzimática, foram eles: *COR* (EC 1.1.1.247) e *SALR* (EC 1.1.1.248), onde segundo a IUBMB os códigos EC 1.1.1 são enzimas de atuação no grupo de doadores CH-OH com  $\text{NAD}^+$  ou  $\text{NADP}^+$  como aceitador. A *CYP80B3* (EC 1.14.14.102), *CYP719B1* (EC 1.14.19.67), *CYP719A2* (EC 1.14.19.64) e a *T6ODM* (1.14.11.3) atuando em doadores emparelhados com incorporação ou redução de oxigênio molecular. A *BBE* (EC 1.21.3.3) agindo sobre a reação  $\text{XH} + \text{YH} = \text{XY}$  com oxigênio como aceitador e a *SARED* (EC 1.3.1.107) atuando no grupo de doadores CH-CH com  $\text{NAD}^+$  ou  $\text{NADP}^+$  como aceitador.

No que se refere à classe EC 2, as transferases são enzimas que catalisam reações de transferência de grupos funcionais e são subdivididas em 10 subcategorias mais amplas (DAVID E NELSON, 2008). Com relação à nomenclatura EC 2 foram encontrados seis possíveis transcritos com correspondência enzimática à aminotransferase, *O*-metiltransferase e *N*-metiltransferase, são elas: *6-OMT* (EC 2.1.1.128), *7-OMT* – (EC 2.1.1.291), *4-OMT* (EC 2.1.1.116) e *CNMT* (EC 2.1.1.140), transferindo grupos O ou N de um carbono. A *TAT* (EC 2.6.1.5), que transfere grupos nitrogenados, e a *SALAT* (EC 2.3.1.150) são aciltransferases envolvidas com grupos de transferência que não sejam grupos amino-acil.

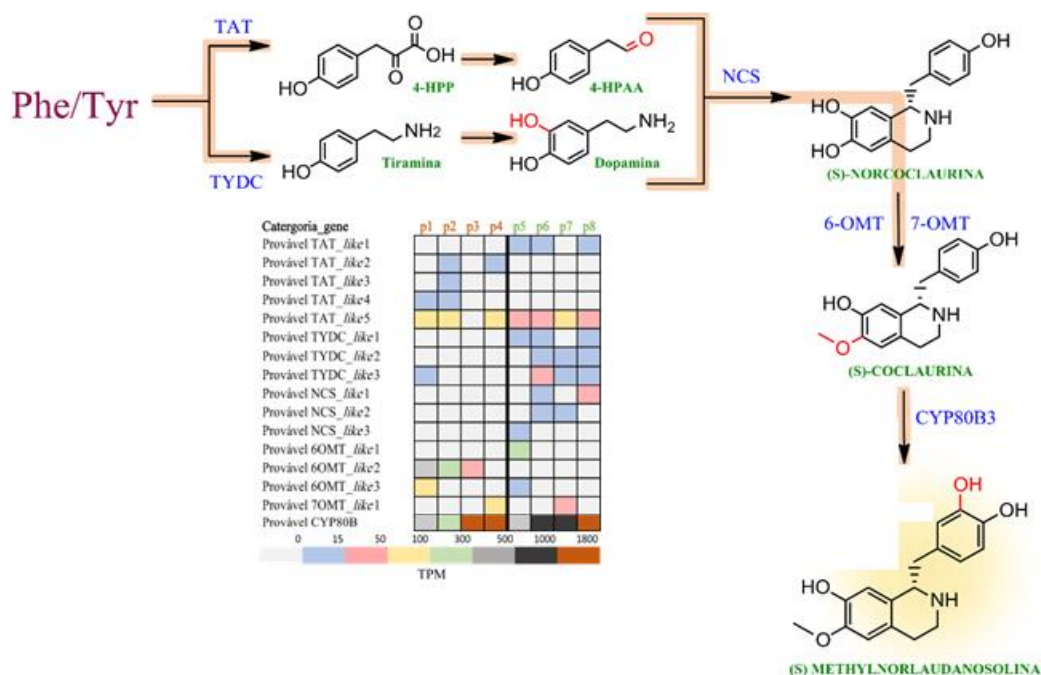
Com relação à terceira e última classe identificada está a nomenclatura EC 4, classe inserida as enzimas liases. Esse grupo de enzimas catalisa a quebra de ligação química ao invés de hidrólise ou oxidação para gerar uma nova ligação dupla ou uma nova estrutura cíclica (SHI et al., 2015). Foram identificados a *TYDC* (EC 4.1.1.25) e a *NCS* (EC 4.2.1.78) nesse grupo.

A via de biossíntese dos alcaloides da classe benzilisoquinolínicos é iniciada pelos aminoácidos aromáticos e tem início com a descarboxilação e transaminação da tirosina pela TIROSINA/DOPA DESCARBOXILASE (*TYDC*) e tirosina AMINOTRANSFERASE (*TAT*), seguido de condensação de dois derivados da L-tirosina: 4-hidroxifenilacetaldeído (4-HPAA) e dopamina, catalisada pela NORCOCLAURINA SINTASE (*NCS*) que produz (*S*)-norcoclaurina (LEE; FACCHINI, 2010; LISCOMBE et al., 2005) (figura 13).

**Figura 13** - Primeiras etapas da biossíntese dos alcaloides eritrínicos

Fenilalanina

Tirosina



Abreviações: *TYDC*, TYROSINE/DOPA DECARBOXYLASE; *TAT*, TYROSINE AMINOTRANSFERASE; *NCS*, NORCOCLAURINE SYNTHASE; *6OMT*, (S)-NORCOCLAURINE 6-O-METHYLTRANSFERASE; *7OMT*, (R,S)-RETICULINE 7-O-METHYLTRANSFERASE; *CYP80B3*, (S)-N-METHYLCOCLAURINE 3'-HYDROXYLASE; Os transcritos considerados possíveis isoformas foram identificadas e nomeadas como "like". P1-p4: transcritos expressos em sementes e p5-p6: transcritos expressos em folhas. Transcritos em azul foram encontrados em nosso transcriptoma. Fonte: elaborado pelo autor (2020).

As sequências dos transcritos correspondentes a essas enzimas envolvidas nos primeiros passos na formação dos alcaloides benzilisoquinolínicos (*TYDC* e *NCS*) foram mais abundantes nas folhas do que nas sementes. Tomando como exemplo, os transcritos que codificam a *NCS* (NORCOCLAURINA SINTASE) estão ausentes nas sementes e presentes nos quatro *pools* de amostras de folhas. As leituras de sequência dos transcritos *TYDC* estão presentes em um *pool* de sementes e em todos os *pools* das folhas. O transcrito *TAT* está presente em 3 *pools* em folhas e sementes, entretanto ligeiramente mais expressa em sementes. O nível de expressão de todas as transcrições candidatas dessa rota é mostrado na Figura 13.

Após essa etapa inicial algumas enzimas já descritas em outras espécies principalmente as que realizam reações de oxidação, redução e transferência do grupamento metil podem atuar nas próximas etapas de formação dos intermediários da via em questão. A primeira reação está envolvida na metilação da (S)-norcoclaurina para formar (S)-coclaurina, em uma etapa conhecida e catalisada pela (R,S)-

NORCOCLAURINA 6-O-METILTRANSFERASE (*6-OMT*) ou (*R,S*)-RETICULINA 7-O-METILTRANSFERASE (*7-OMT*).

No que diz respeito à classe de O-metiltransferases e N-metiltransferases, elas são características onipresentes e recorrentes na biossíntese de muitos metabólitos especializados, especialmente plantas produtoras de alcaloides benzilisoquinólicos (BIAs) (FACCHINI; MORRIS, 2019). O nível de expressão dos transcritos possivelmente associados a *6-OMT* foi maior nas sementes. Esses transcritos estão presentes em pelo menos três *pools* de sementes e ausentes em pelo menos dois nas folhas. Em contraste, a *7-OMT* está presente em apenas um *pool* de amostras em folhas e um de sementes. A *7-OMT* é responsável por catalisar a transferência de um grupo metil para reticulina e formar laudanina, um alcaloide obtido do ópio e caracterizado funcionalmente na biossíntese da BIAs. O silenciamento do gene *7-OMT* causou repressão no conteúdo total de alcaloides no caule, enquanto que a superexpressão resultou em maior acúmulo de morfina (GURKOK et al., 2016).

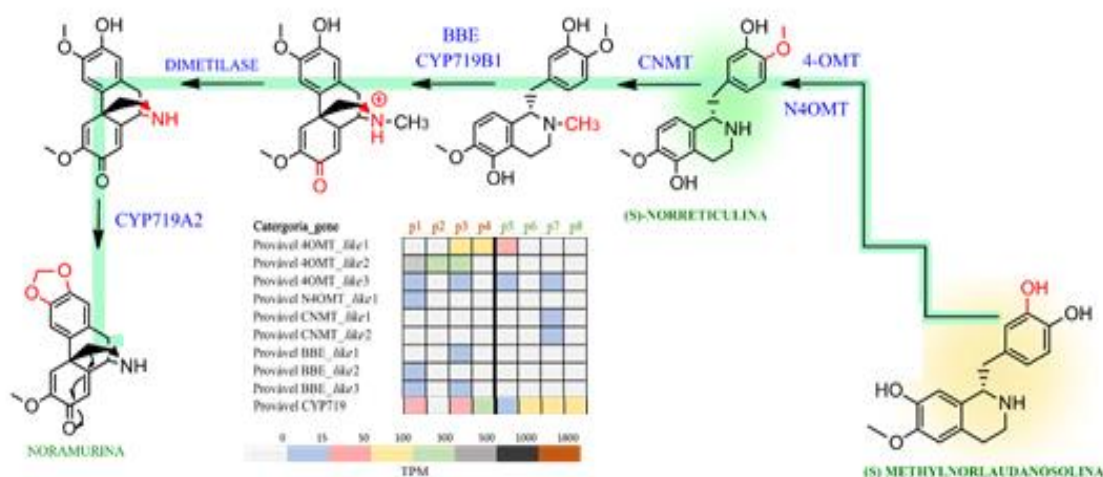
Para originar o intermediário (*S*)-metilnorlaudanolina a possível candidata é a (*S*)-N-METILCOCLAURINA 3'-HIDROXILASE (*CYP80B3*). Esta monooxigenase do citocromo P<sub>450</sub> está presente na papoula do ópio e contribui para o aumento da biossíntese da morfina. Estudos relatam que a superexpressão do cDNA do *CYP80B3* resultou em um aumento na quantidade de alcaloides totais BIAs (FRICK; KRAMELL; KUTCHAN, 2007; ZULAK et al., 2009). Os transcritos que codificam as enzimas do citocromo P<sub>450</sub> são importantes para contribuir para a diversidade estrutural de moléculas, incluindo alcaloides. A literatura mostra que principalmente os membros das subfamílias *CYP719A9*, *CYP719A11*, *CYP719A2*, *CYP719A3*, *CYP719A1* e *CYP80B3* estão envolvidos na biossíntese da porção isoquinolina (FRICK; KRAMELL; KUTCHAN, 2007; IKEZAWA; IWASA; SATO, 2009; HORI et al., 2017). Por ser uma classe abundante que catalisa reações muito específicas em isoformas e subtipos, é apresentado apenas o nível de expressão dos transcritos da *CYP80B* que podem estar potencialmente envolvidos na rota de biossíntese proposta, com base na análise comparativa de similaridade da sequência.

Todos os transcritos que codificam a (*S*)-N-METHYLCOCLAURINE 3'-HYDROXYLASE obtiveram alta expressão em folhas e sementes, indicando que na porção intermediária da rota em questão, possivelmente há uma maior atividade catalítica desta CYP.

A (*S*)-metilnorlaudanosolina é um dos pontos centrais da cascata química dos alcaloides eritrínicos, que pode ser gerada uma bifurcação para (*S*)-norreticulina (i) e norprotosinomenina (ii).

No que diz respeito ao intermediário (*S*)-norreticulina (i), leva à formação de alcaloides comuns de *Erythrina*, ou seja, alcaloides que em seu esqueleto químico contêm uma espiroamina tetracíclica e provêm de espécies do gênero *Erythrina*. Após a formação de (*S*)-metilnorlaudanosolina, esta será metilada pela 4'-O-METILTRANSFERASE (*4-OMT*) para gerar a (*S*)-norreticulina (figura 14). A *4-OMT* é um transcrito relatado na literatura e presente principalmente em espécies produtoras de BIAs, como: *Coptis japonica* (MORISHIGE et al., 2000), *Papaver somniferum* (GURKOK et al., 2016), *Eschscholzia californica* (PURWANTO et al., 2017) e *Narcissus* spp. (KILGORE et al., 2014). O silenciamento dos genes *4-OMT2* causou uma redução em todos os alcaloides medidos nas hastes de *Papaver somniferum* L. (GURKOK et al., 2016), esse silenciamento se deu através da indução por vírus, usando o sistema vetorial baseado em *Tobacco ratle virus* (TRV). Assim como a *6-OMT*, o nível de expressão dos transcritos possivelmente associados a *4-OMT* foi ligeiramente maior nas sementes e presente em pelo menos duas amostras de sementes e uma de folha (figura 14).

**Figura 14** - Etapa intermediária 1 da biossíntese dos alcaloides eritrínicos



Abreviações: 4OMT, 3'-hydroxy-N-methyl-(S)-coclaurine 4'-O-methyltransferase; N4OMT, Norbelladine 4'-O-methyltransferase; CNMT, (S)-coclaurine N-methyltransferase; BBE, berberine bridge enzyme; CYP719B1, Salutaridine synthase; CYP712A2, (S)-stylophine synthase. Fonte: elaborado pelo autor (2020).

Além das O-metiltransferases consideradas compatíveis com as de *Papaver somniferum*, também foi identificada um transcrito presente no primeiro *pool* de sementes, a N4OMT (NORBELLADINA 4'-O-METILTRANSFERASE), um transcrito presente em algumas espécies da família Amaryllidaceae catalisando reações envolvidas na biossíntese da galantamina, um alcaloide de classe das BIAs usado clinicamente para tratar a doença de *Alzheimer* (LI et al., 2019).

A escolha da norreticulina como um dos principais intermediários baseia-se no estudo de Maier et al (1999) com uma nova proposta de que a via da coclaurina-norreticulina é operativa na formação de alcaloides eritrínicos e é o estudo de biossíntese mais aceito atualmente. Segundo o autor, esses intermediários podem originar à noramurina por meio de reações químicas, permitindo assim a formação dos alcaloides. No entanto, a via ainda não foi completamente elucidada, principalmente com relação às reações enzimáticas. Dessa forma, neste trabalho foi explorada a rota de maneira complementar aos estudos de Maier et al (1999) para investigar a produção desses alcaloides pela planta.

A via da reação de formação da noramurina a partir da (S)-norreticulina é necessária principalmente de enzimas que catalisam a formação de ligações C-C e enzimas que formam pontes de metilenodioxí. A primeira reação pode ser catalisada pela enzima ponte de berberina (BBE) ou pela SALUTARIDINE SYNTHASE (CYP719b1) e a segunda reação por monooxigenases dependente de P<sub>450</sub> (CYP719A2 e CYP719A3). Os cDNAs da BBE já foram isolados em alguns estudos e a caracterização cinética relatada

(ALCANTARA et al., 2005; DITTRICH; KUTCHAN, 1991; WINKLER et al., 2009). A *CYP719B1* é um transcrito traduzido em uma enzima altamente estereoquímica e regioseletiva, apenas a (R)-reticulina e (R)-norreticulina resultaram na formação de um produto (salutaridina e norsalutaridina, respectivamente) e sua identificação como um membro da família *CYP719* ocorreu durante uma triagem de citocromos recombinantes P<sub>450</sub> de papoula de ópio expressada funcionalmente em células *Spodoptera frugiperda* Sf9 (GESELL et al., 2009).

Os citocromos P<sub>450</sub> exibem uma notável diversidade de substratos, são altamente específicos, exibindo uma acentuada estereosseletividade (STOUT, 2004). O sequenciamento de várias *CYPs* de diversos organismos identificou 350.000 sequências em fontes públicas e privadas e foi observado que grupos de organismos bem estudados, como vertebrados, são cobertos com maior profundidade via pesquisa *BLAST*, enquanto que organismos como plantas terrestres possuem uma profundidade menor, além disso, essa profundidade dos projetos de sequenciamento determina a cobertura de sequência do P<sub>450</sub> (NELSON, 2018).

Ainda com relação às enzimas pertencentes ao citocromo P<sub>450</sub>, elas podem também formar pontes de metilenodioxí, a literatura mostra que algumas foram clonadas a partir de células cultivadas de *Eschscholzia californica*, mostrando que tanto a *CYP719A2* como a *CYP719A3* têm atividade para catalisar a esse tipo de reação, e, além disso, o tratamento com metil jasmonato resultou na indução coordenada desses genes (IKEZAWA; IWASA; SATO, 2007).

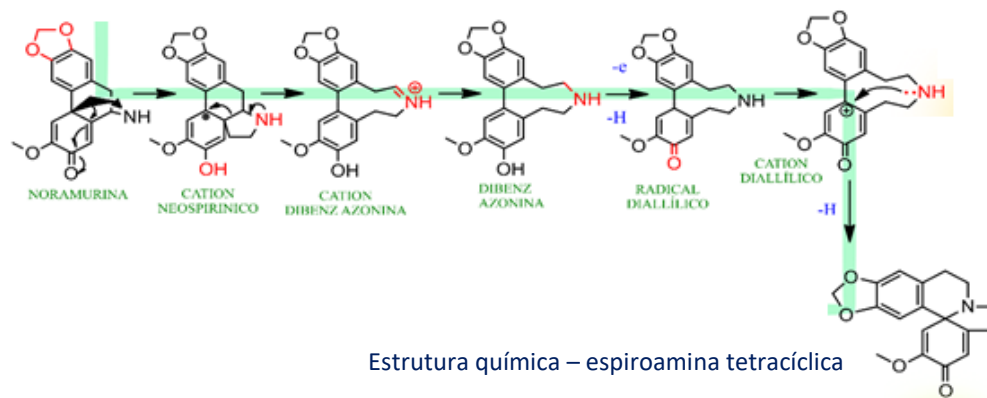
Em geral, os transcritos que codificam enzimas que catalisam reações de oxidação/redução e aqueles pertencentes ao citocromo P<sub>450</sub>, foram as mais abundantes em nossos dados de transcriptoma, propondo assim que eles poderiam participar da via para levar à formação da noramurina (figura 14).

Antes da reação de formação de ligação C-C através da *BEE* e/ou *CYP719B1*, é necessário um ponto de adição de N-metilação (N-CH<sub>3</sub>) para avançar a formação de alcaloides. Em geral, as metiltransferases do metabolismo das BIAs aceitam uma ampla variedade de substratos e alcaloides de várias estruturas, com alguns mostrando mais flexibilidade do que outros em relação à variedade de substratos (OUNAROON et al., 2003). Como possível candidato o transcrito encontrado da S-COCLAURINA N-METILTRANSFERASE (*CNMT*) exclusivos em folhas, poderia usar a norreticulina como substrato, pois o próximo passo que é catalisado pelo *BBE* e/ou *CYP719B1* precisa da porção N-metil para a conversão estereoespecífica. O *CNMT* já foi caracterizado na

biossíntese da Papaverina, um importante alcaloide verdadeiro com núcleo isoquinolínico derivado do aminoácido tirosina (PATHAK et al., 2013).

Os mecanismos de reação a partir da noramurina para formação do cátion dialílico são semelhantes ao mecanismo biossintético proposto por Maier et al (1999) e Barton et al (1968) (figura 15).

**Figura 15** - Etapa intermediária 2 da biossíntese dos alcaloides eritrínicos

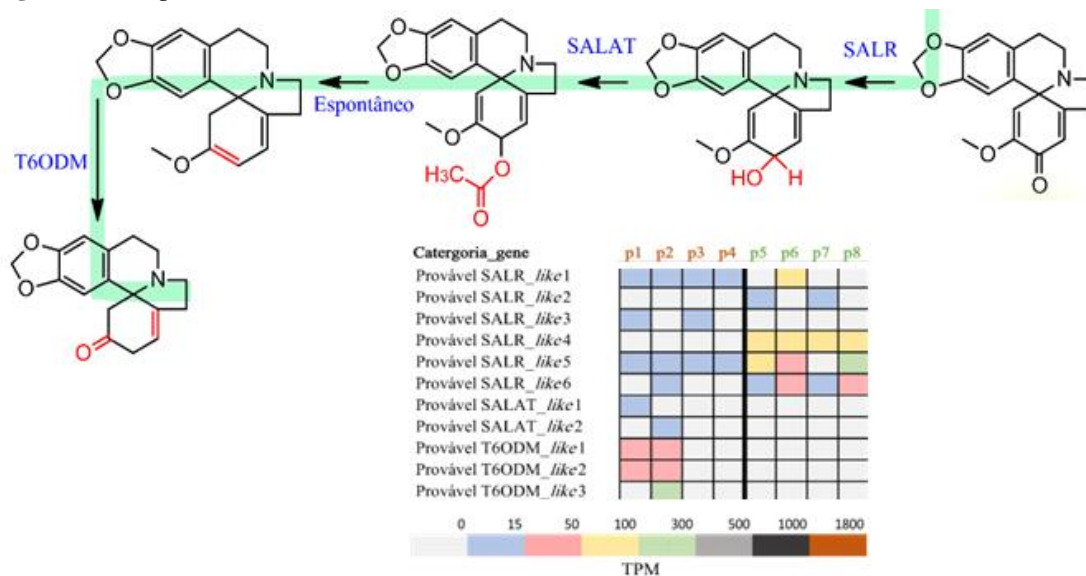


Fonte: Adaptado de Maier et al (1999)

Dessa forma, com o esqueleto básico eritrínico já formado (espiroamina tetracíclica) na etapa seguinte a formação do cátion dialílico, em nossa via de biossíntese propusemos que a SALUTARIDINA REDUTASE (SALR) (figura 16), membro da família redutase de cadeia curta poderia fazer uma reação de redução usando o NADPH como coenzima no anel A do esqueleto químico básico eritrínico.

Essa enzima está envolvida principalmente na formação dos alcaloides morfina e codeína (GERARDY; ZENK, 1993; HIGASHI; KUTCHAN; SMITH, 2011). A SALR está presente em todos os *pools* de folhas e sementes de *Erythrina velutina*.

**Figura 16** - Etapa intermediária 3 da biossíntese dos alcaloides eritrínicos



Abreviações: *SALR*, SALUTARIDINE REDUCTASE; *SALAT*, SALUTARIDINOL 7-O-ACETYLTRANSFERASE; *T6ODM*, THEBAINE 6-O-DEMETHYLASE. Fonte: elaborado pelo autor (2020).

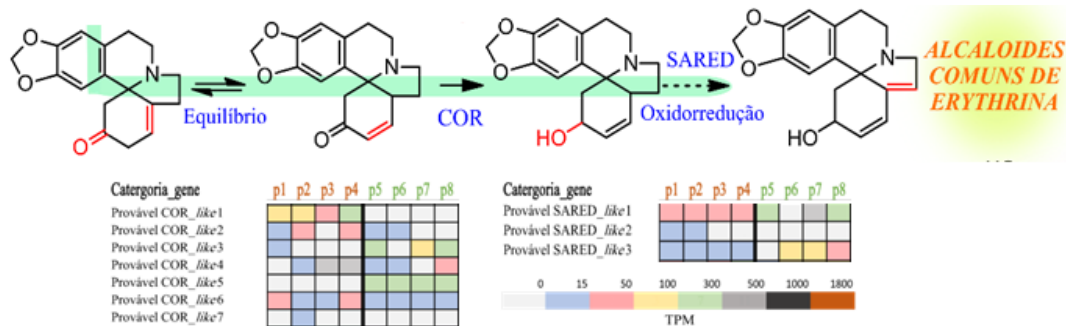
O próximo passo proposto da reação é catalisado por SALUTARIDINOL 7-O-ACETILTRANSFERASE (*SALAT*), que acetila especificamente o epímero 7 do salutaridinol em salutaridinol-7-O-acetato na formação de alcaloides benzilisoquinolina, que, dependendo do pH, o grupo acetil introduzido pode ser eliminado espontaneamente, levando à formação de uma ponte de óxido entre C4 e C5 para produzir a tebaína o primeiro alcaloide pentacíclico antes da formação de morfina (FACCHINI; DE LUCA, 2008). Em nossa proposta, haveria inicialmente uma acetilação no anel A, seguida de desacetilação espontânea. O *SALAT* foi encontrado apenas em dados de sementes em *Erythrina velutina*.

Em seguida, o membro da enzima THEBAINA 6-O-DIMETILASE (*T6ODM*) da família das dioxigenase dependente de Fe (II)/2-oxoglutarato (Fe<sup>2+</sup>/2OG) envolvida na formação de alcaloides da morfina (HAGEL; FACCHINI, 2010) poderia catalisar a reação no anel A do esqueleto químico para continuidade no fluxo de produção dos alcaloides. Os transcritos *T6ODM* foram únicos na semente, presentes em dois pools de amostras.

Os estágios finais da biossíntese consistem em uma reação de equilíbrio entre dois intermediários eritrínicos com o ambiente químico semelhante ao que ocorre entre os intermediários da neopinona/codeinona como demonstrado por Larkin et al. (2007) em *P. somniferum*, e reações de oxidoredução para gerar alcaloides mais específicos. Para

essa última reação citada, os possíveis candidatos incluem a CODEINONA REDUTASE (*COR*) e a SANGUINARINA REDUTASE (*SARED*) (figura 17).

**Figura 17** - Etapa final da biossíntese dos alcaloides eritrínicos

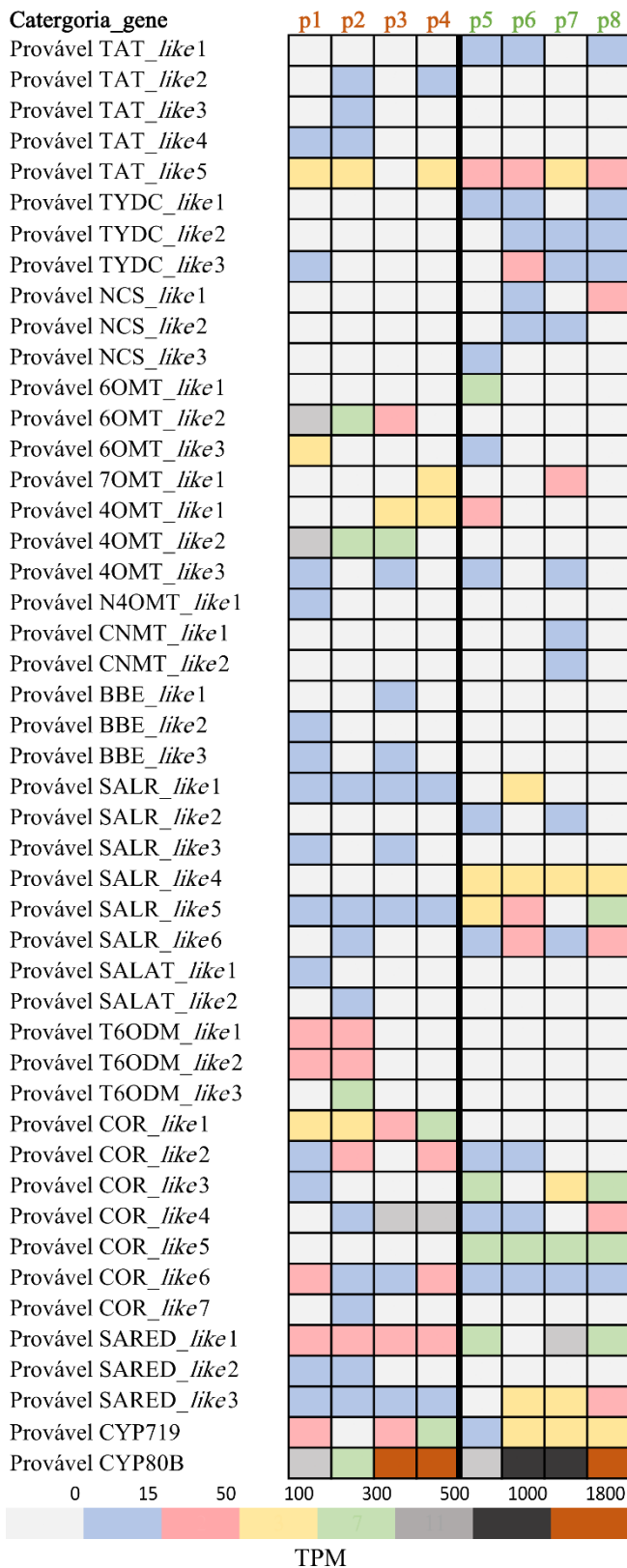


Abreviações: *COR*, NADPH-DEPENDENT CODEINONE REDUCTASE; *SARED*, SANGUINARINE REDUCTASE. Setas pontilhadas indicam etapas ainda não relatadas. Fonte: elaborado pelo autor (2020).

A *COR* que já possui cDNA identificado em papoula e expressão funcional em *Escherichia coli* (UNTERLINER; LENZ; KUTCHAN, 1999) faria uma reação inicial de redução, enquanto que a SANGUINARINA REDUTASE (*SARED*), uma enzima que catalisa reações de redução, está envolvida na desintoxicação de fitoalexinas produzidas por plantas, usaria NAD(P) ou NAD como doador de hidrogênio. Essa proteína já foi isolada a partir de culturas de células de *Eschscholzia californica* (VOGEL et al., 2010).

Esses possíveis candidatos realizariam reações de oxidação e redução que levariam à estrutura final dos alcaloides comuns de *Erythrina*, uma vez que a diversidade estrutural dos alcaloides identificados neste estudo é enorme e, na maioria das vezes, essas diferenças estruturais se devem principalmente à oxidação e redução nas estruturas. A *SARED* e *COR* estão presentes em todos os conjuntos de amostras de folhas e sementes (figura 17). A rota de biossíntese proposta completa, bem como as respectivas expressões dos transcritos identificado neste trabalho estão na figura 18 e 19 respectivamente.

**Figura 18** - Mapa de calor dos transcritos identificados a participarem da via de biossíntese dos alcaloides eritrínicos

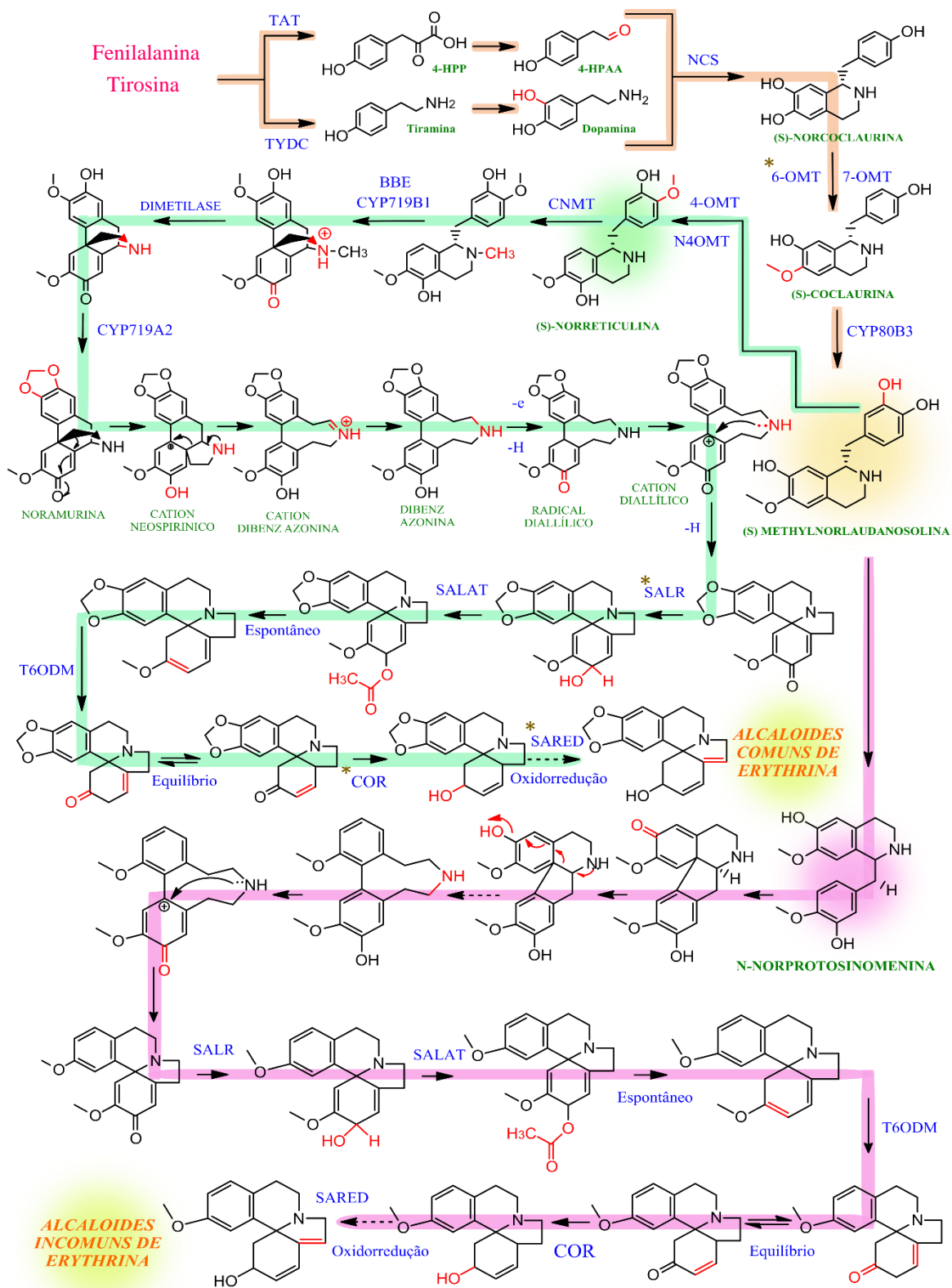


O perfil de expressão dos 17 candidatos envolvidos na via de biossíntese dos alcaloides de *Erythrina*. Os valores de expressão foram identificados em TPM (transcrito por milhão). As contagens de leitura dos transcritos expressos foram agrupadas com base na identificação mais detalhada através do alinhamento. Os transcritos considerados possíveis isoformas foram nomeadas como "*like*". P1-p4: transcritos em sementes e p5-p6: transcritos em folhas. Fonte: elaborado pelo autor (2020).

No que diz respeito à norprotosinomenina (segunda abordagem da rota), este intermediário leva à formação de alcaloides eritrínicos incomuns, isto é, alcaloides que em seu esqueleto químico contêm uma espiroamina tetracíclica, mas é originalmente da espécie *Cocculus laurifolius*. Estudos sobre a biossíntese desses alcaloides são extremamente escassos, o que representa um grande desafio para este trabalho, que mesmo com todas as limitações, foi possível propor um caminho para a formação desses intrigantes alcaloides.

O foco principal são os estudos mais aceitos na literatura sobre a biossíntese de alcaloides da coculina (BHAKUNI; JAIN, 1980; BHAKUNI; SINGH, 1978), cuja norprotosinomenina foi um intermediário que apresentou forte incorporação específica para alcaloides coculínicos, enquanto a norreticulina apresentou baixa incorporação, o que indica que seu envolvimento não foi significativo na produção desses alcaloides. O caminho começa com uma série de rearranjos químicos para formar o cátion dialílico da norprotosinomenina. Os próximos passos da rota biossintética são semelhantes aos descritos para os alcaloides comuns de *Erythrina* (Figura 19).

Figura 19 - Proposta da rota de biossíntese para alcaloides da *Erythrina*



Os alcaloides eritrínicos são do tipo isoquinolina, formando um componente do precursor da fenilalanina. O caminho começa com ramos iniciados com 4-HPP e tirosina, que por ação enzimática formam 4-HPAA e dopamina, respectivamente, produzindo intermediário (*s*)-norcoclaurina e (*s*)-coclaurina. A (*S*)-metilnorlaudanosolina é um dos pontos cruciais na cascata química dos alcaloides eritrínicos, pois parte dela pode gerar uma bifurcação para (*S*)-norreticulina e norprotosinomenina. Caminhos em verde levam à formação de alcaloides comuns da *Erythrina*; Caminhos cor de rosa levam à formação de alcaloides incomuns; Setas pontilhadas indicam reações ainda não estabelecidas totalmente; Asterisco em amarelo indicam os prováveis transcritos que tiveram correspondência proteica. Abreviações: *TYDC*, TYROSINE/DOPA DECARBOXYLASE; *TAT*, TYROSINE AMINOTRANSFERASE; *NCS*, NORCOCLAURINE SYNTHASE; *6OMT*, (*S*)-NORCOCLAURINE 6-O-METHYLTRANSFERASE; *7OMT*, (R,S)-RETICULINE 7-O-METHYLTRANSFERASE; *CYP80B3*, (*S*)-N-METHYLCOCLAURINE 3'-HYDROXYLASE; *4OMT*, 3'-HYDROXY-N-METHYL-(*S*)-COCLAURINE 4'-O-METHYLTRANSFERASE; *N4OMT*, NORBELLADINE 4'-O-METHYLTRANSFERASE; *CNMT*, (*S*)-COCLAURINE N-METHYLTRANSFERASE; *BBE*, BERBERINE BRIDGE ENZYME; *CYP719B1*, SALUTARIDINE SYNTHASE; *CYP712A2*, (*S*)-STYLOPINE SYNTHASE; *SALR*, SALUTARIDINE REDUCTASE; *SALAT*, SALUTARIDINOL 7-O-ACETYLTRANSFERASE; *T6ODM*, THEBAINE 6-O-DEMETHYLASE; *COR*, NADPH-DEPENDENT CODEINONE REDUCTASE; *SARED*, SANGUINARINE REDUCTASE. Transcritos em azul foram encontrados em nosso transcriptoma; Setas pontilhadas indicam etapas ainda não relatadas. Fonte: elaborado pelo autor (2020).

De forma geral, os transcritos que participam da porção inicial da via foram mais abundantes nas folhas, enquanto os transcritos envolvidos nos estágios intermediário e final foram semelhantes nas duas partes vegetais, sendo ligeiramente maiores nas folhas, principalmente nos transcritos que realizam oxidação e redução. Curiosamente, esses resultados demonstram características importantes quando combinados com o conjunto de dados do metaboloma, pois os resultados indicam que tanto as folhas como as sementes são produtores de alcaloides, com a folha produzindo mais alcaloides exclusivos do que as sementes.

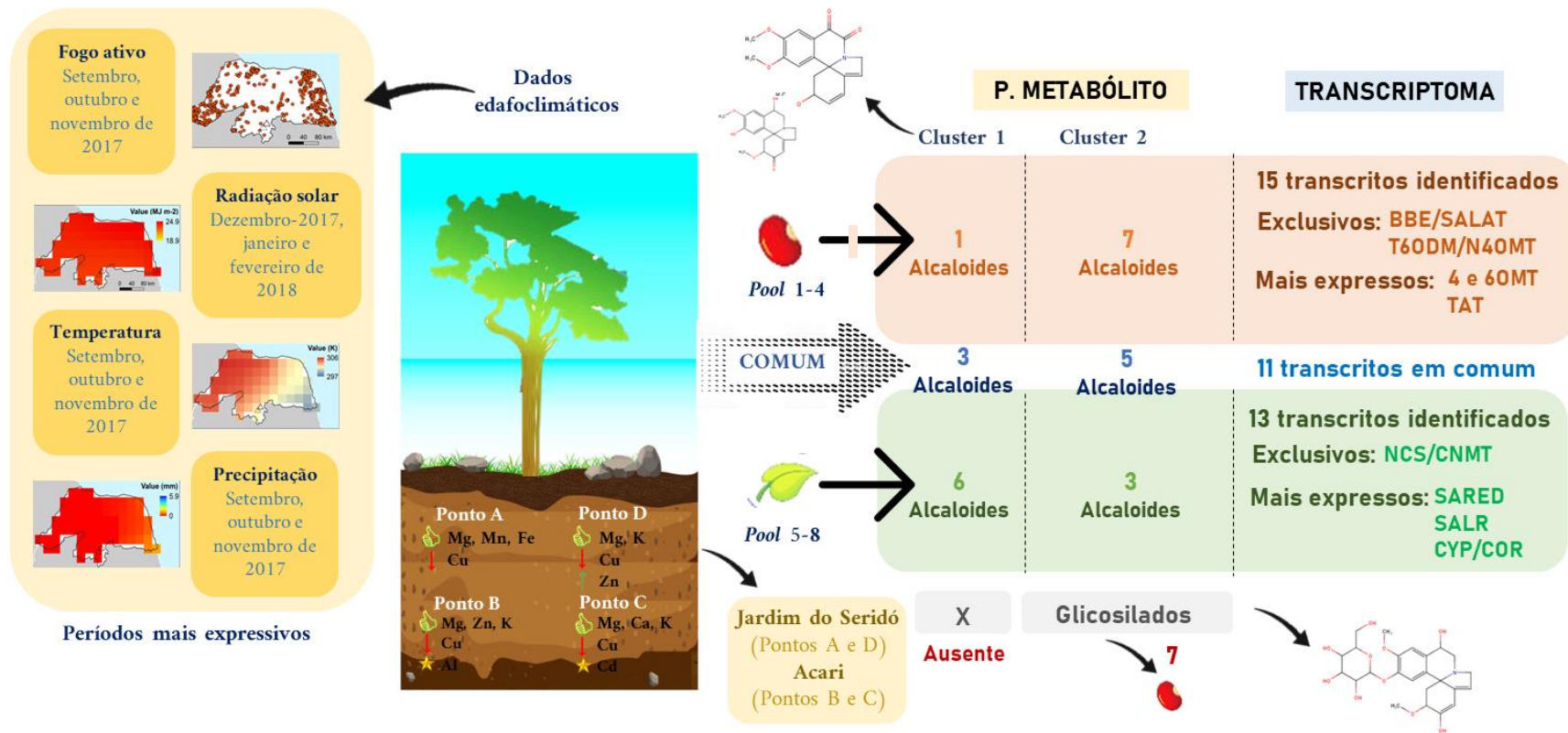
Entre o cenário de biodiversidade de produtos naturais com atividade farmacológica, os alcaloides ocupam lugar de destaque. Os alcaloides bioativos de origem vegetal têm sua importância tradicionalmente conhecida no contexto das Ciências Farmacêuticas e, concomitantemente, espécies do gênero *Erythrina* também têm seu uso popular tradicionalmente relatado por um longo tempo. No entanto, até o momento, não há estudo moleculares para nenhuma espécie do gênero. Este trabalho, através da tecnologia de sequenciamento de próxima geração, é o primeiro e mais abrangente estudo sobre o conjunto de genes de *Erythrina velutina* até o momento.

Em geral, através deste estudo foi discutida a relação entre a expressão dos transcritos, bem como o acúmulo alcaloides em folhas e sementes. Essa diferença de acúmulo associada à expressão de transcritos e proteínas permite inferir que o armazenamento e o transporte também são fundamentais para aumentar a produção de metabólitos em folhas e sementes, o que estimula a continuidade desse conhecimento,

explorando e identificando os transcritos envolvidos na biossíntese, bem como proteínas transportadores, proteínas que realizam reações de glicosilação e confirmando-as funcionalmente.

Nos últimos anos, esforços têm sido feitos em pesquisas relacionadas à biossíntese e aplicação biotecnológica de alcaloides benzilisoquinolínicos, no entanto, resultados que podem confirmar esse tema ainda são escassos. Acredita-se que os resultados discutidos aqui contribuirão como uma ferramenta de informação para a continuidade da investigação da biossíntese e para uma identidade molecular e biotecnológica para aplicação futura, a fim de otimizar o aumento da produção de alcaloides de *Erythrina*. Compreender vias biossintéticas é crucial para a produção de metabólitos. Uma representação geral de todos os resultados desse trabalho pode ser encontrada na figura 20.

Figura 20 - Resumo dos resultados obtidos



Na imagem, os períodos correspondentes aos fatores edafoclimáticos mencionados são os mais expressivos durante o tempo de análise. No que se refere a análise do solo, setas vermelhas e verdes indicam nutrientes em concentração baixas e altas respectivamente; Estrelas indicam nutrientes exclusivos no ponto. No que se refere a análises “ômicas”, retângulos laranja e verde indicam metabólitos, transcritos e proteínas exclusivos de sementes e folhas, enquanto que retângulo e palavras em azul indicam metabólitos, transcritos e proteínas comum a folhas e sementes. Alcaloides glicosilados foram identificados apenas na classe dienoide/alcenoide. Fonte: elaborado pelo autor (2020).

## 6 CONCLUSÃO

De acordo com o presente estudo foi possível a realização do transcriptoma, proteoma e perfil metabólito direcionado para a biossíntese de alcaloides, o qual identificou os possíveis transcritos e proteínas envolvidos na produção desses importantes metabólitos.

No que se refere ao estudo do transcriptoma, os transcritos candidatos em nossa proposta que participam da porção inicial foram mais abundantes em folhas do que em sementes, com exceção de *TAT* que foi mais abundante em sementes. Os transcritos que participam da porção intermediária estão em ambas as partes, porém e mais expressos em semente, exceto *CNMT* e *BBE* que se apresentaram como transcritos exclusivos de folhas e semente, respectivamente. Os transcritos participantes da porção final da rota foram muito abundantes em ambas as partes, com exceção da *T6ODM* que estava presente exclusivamente em semente.

Para uma melhor investigação de funcionalidade e de forma complementar ao conjunto de transcritos identificados, foi anotado 25 alcaloides eritrínicos, das classes dienoide e alcenoide, presentes tanto em folhas como em sementes, dos quais 8 são exclusivos de sementes e 9 exclusivos de folhas.

Em resumo, os resultados desse estudo contribuíram para compreender como a espécie *Erythrina velutina* se comporta no que se refere a produção de alcaloides frente às condições naturais no bioma da Caatinga. Acredita-se que essa espécie neste importante cenário com vários fatores bióticos e abióticos impostos pode produzir moléculas sem precedentes, além de este estudo servir como base para futuros estudos de melhoramento molecular para maior obtenção desses metabólitos de interesse com a integração das análises “ômicas”, de forma abranger áreas da engenharia metabólica e biologia sintética desenhando circuitos biológicos por meio do redirecionamento e construção da rota metabólicas, em especial de alcaloides.



## REFERÊNCIA BIBLIOGRÁFICA

ABDELNUR, P. V. **Metabolômica e Espectrometria de Massas**. Circular Técnica, 10, p. 1–4, 2011.

ALCANTARA, J. et al. Sanguinarine biosynthesis is associated with the endoplasmic reticulum in cultured opium poppy cells after elicitor treatment. **Plant Physiology**, v. 138, n. 1, p. 173–183, 2005.

ALMEIDA, Wamberto Alristenio Moreira. **Design e síntese de derivados da retronecina e avaliação da atividade in vivo no modelo de *Caenorhabditis elegans* para a doença de alzheimer**. 2018. 98 f. Dissertação (Mestrado) - Curso de Farmácia, Universidade Federal do Rio Grande do Norte, Natal, 2018.

ANHESINE, Naira Buzzo. **Perfil cromatográfico (fingerprinting) por LC-DAD-MS de espécies do gênero *Erythrina*: desenvolvimento e validação de método para controle de qualidade de drogas vegetais**. 2018. 177 f. Tese (Doutorado) - Curso de Instituto de Química, Universidade Estadual Paulista Júlio de Mesquita Filho, Araraquara, 2018.

BORGES, Warley de Sousa; BERLINCK, Roberto G. S.; SCOTTI, Marcus T.; VIEIRA, Paulo Cezar. A Química de Produtos Naturais do Brasil no Século XXI. **Química Nova**, p.106-110, 3 jul. 2017.

DUARTE, Danilo Freire. Uma breve história do ópio e dos opióides. **Revista Brasileira de Anestesiologia**, v. 55, n. 1, p.135-146, fev. 2005.

LIN, Shuhai et al. LC/MS-based non-targeted metabolomics for the investigation of general toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in C57BL/6J and DBA/2J mice. **International Journal of Mass Spectrometry**, Hong Kong, v. 301, n. 1-3, p.29-36, mar. 2011.

PLUSKAL, Tomáš et al. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. **BMC Bioinformatics**, Okinawa, v. 11, n. 1, p.1-11, 23 jul. 2010.

FURTINI NETO, Antônio Eduardo; VALE, Fabiano Ribeiro do; RESENDE, Alvaro Vilela de. **Fertilidades do solo**. 2001. 261 f. Especialização - Curso de Fertilidade do Solo e Nutrição de Plantas, Universidade Federal de Lavras, Lavras, 2001.

AMER, M. E.; SHAMMA, M.; FREYER, A. J. The Tetracyclic *Erythrina* Alkaloids. **Journal of Natural Products**, v. 54, n. 2, p. 329–363, mar. 1991.

AMIRKIA, V.; HEINRICH, M. Alkaloids as drug leads - A predictive structural and biodiversity-based analysis. **Phytochemistry Letters**, v. 10, p. 48–53, 2014.

ANSORGE, W. J. Next-generation DNA sequencing techniques. **New Biotechnology**, v. 25, n. 4, p. 195–203, 2009.

AVERY, O. T.; MACLEOD, C. M.; MC CARTY, M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types: Induction or

transformatioby A Desoxyribonucleic acid fraction isolated from pneumococcus type III. **Journal of Experimental Medicine**, v. 179, n. 2, p. 378–384, 1994.

BARBOSA, E. B. et al. Proteômica: metodologias e aplicações no estudo de doenças humanas. **Revista da Associação Médica Brasileira**, v. 58, n. 3, p. 366–375, maio 2012.

BARREIRO, E. J.; BOLZANI, V. D. S. Biodiversidade: fonte potencial para a descoberta de fármacos. **Química Nova**, v. 32, n. 3, p. 679–688, 2009.

BARTON, D. H. R. et al. Phenol oxidation and biosynthesis. Part XVIII. The structure and biosynthesis of Erythrina alkaloids. **Journal of the Chemical Society Organic**, v. 5, p. 1529, 1968.

BEAUDOIN, Guillaume A. W.; FACCHINI, Peter J.. Benzylisoquinoline alkaloid biosynthesis in opium poppy. **Planta**, [s.l.], v. 240, n. 1, p.19-32, 2014.

BEZERRA JÚNIOR, José Gilberto Olimpio; SILVA, Nubelia Moreira da. Caracterização geoambiental da microrregião do seridó oriental do rio grande do norte. **Holos**, Natal, v. 2, n. 23, p.78-91, abr. 2007;

BHAKUNI, D. S.; JAIN, S. Late stages in the biosynthesis of abnormal Erythrina alkaloids. **Tetrahedron**, v. 36, n. 14, p. 2153–2156, jan. 1980.

BHAKUNI, D. S.; SINGH, A. N. Biosynthesis of the abnormal Erythrina alkaloids, cocculidine and cocculine. **Journal of the Chemical Society**, Perkin Transactions 1, v. 1, n. 6, p. 618, 1978.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: A flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 2014.

BRADFORD, M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. **Analytical Biochemistry**, v. 72, n. 1-2, p.248-254, 7 maio 1976.

BRAZ FILHO, R. Contribuição da fitoquímica para o desenvolvimento de um país emergente. **Química Nova**, v. 33, n. 1, p. 229–239, 2010.

BROADLEY, M. et al. Function of Nutrients. In: Marschner's Mineral Nutrition of Higher Plants. **Elsevier**, p. 191–248, 2012.

CARQUEJEIRO, I.; NORONHA, H.; DUARTE, P.; GEROS, H.; SOTTOMAYOR, M.. Vacuolar Transport of the Medicinal Alkaloids from *Catharanthus roseus* Is Mediated by a Proton-Driven Antiport. **Plant Physiology**, v. 162, n. 3, p.1486-1496, 17 maio 2013.

CARVALHO, A. C. C. S. et al. Evidence of the mechanism of action of *Erythrina velutina* Willd (Fabaceae) leaves aqueous extract. **Journal of Ethnopharmacology**, v. 122, n. 2, p. 374–378, mar. 2009.

CARVALHO, P. E. R. **Mulungu (*Erythrina velutina*)**. **Circular Técnica** - Empresa Brasileira de Pesquisa Agropecuária, v. 3, n. 160, p. 8, 2008.

- CELIK, I. et al. Development of genomic simple sequence repeat markers in opium poppy by next-generation sequencing. **Molecular Breeding**, v. 34, n. 2, p. 323–334, 2014.
- CIRINO, P. C.; ARNOLD, F. H. Protein engineering of oxygenases for biocatalysis. **Current Opinion in Chemical Biology**, v. 6, n. 2, p. 130–135, 2002.
- CONSORTIUM, The Uniprot. UniProt: a worldwide hub of protein knowledge. **Nucleic Acids Research**, Cambridge, v. 47, n. 1, p.506-515, 5 nov. 2018. Oxford University Press (OUP).
- COSTA, Valerio et al. Uncovering the Complexity of Transcriptomes with RNA-Seq. **Journal of Biomedicine and Biotechnology**, v. 2010, p.1-19, 2010.
- DANTAS, M. C. et al. Central nervous system effects of the crude extract of *Erythrina velutina* on rodents. **Journal of Ethnopharmacology**, v. 94, n. 1, p. 129–133, 2004.
- DAVID, L.; NELSON, M. M. C. Lehninger principles of biochemistry. 5th ed. New York: W. H. **Freeman**, 2008. 1158 p.
- DE VOS, R. C. H. et al. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. **Nature Protocols**, v. 2, n. 4, p. 778–791, 2007.
- DECKER, M. W. et al. Diversity of Neuronal nicotinic acetylcholine receptors: lessons from behavior and implications for CNS therapeutics. **Life Sciences**, v. 56, n. 8, p. 26, 1995.
- DEVI, M. R.; MANOHARAN, A. Characteristics of pharmacognostical significance of *Erythrina*. **Journal of Chemical and Pharmaceutical Research**, v. 3, n. 6, p. 707–714, 2011.
- DEWICK, Paul M. The Shikimate Pathway: Aromatic Amino Acids and Phenylpropanoids. **Medicinal Natural Products**, v. 1, p.137-186, 2002.
- DITTRICH, H.; KUTCHAN, T. M. Molecular cloning, expression, and induction of berberine bridge enzyme, an enzyme essential to the formation of benzophenanthridine alkaloids in the response of plants to pathogenic attack. *Proceedings of the National Academy of Sciences of the United States of America*, v. 88, n. 22, p. 9969–9973, 1991.
- DIXON, M.; WEBB, E.c.. *Enzymes*. New York: **Academic Press**, 1979. 86 p.
- DUDLEY, E. et al. Targeted metabolomics and mass spectrometry. 1. ed. **Elsevier Inc**, 2010. v. 80
- ELLIS, E. M. Microbial aldo-keto reductases. **FEMS Microbiology Letters**, v. 216, n. 2, p. 123–131, 2002.
- EMBRAPA. Manual de Métodos de Análise de Solo. Rio de Janeiro: **Embrapa-cnps**, 1997. 212 p.
- EOM, S. H.; NA, J. K. Leaf transcriptome data of two tropical medicinal plants: *Sterculia lanceolata* and *Clausena excavata*. **Data in Brief**, v. 25, p. 104297, 2019.

FACCHINI, P. J. ALKALOID BIOSYNTHESIS IN PLANTS : Biochemistry, Cell Biology, Molecular Regulation, and Metabolic Engineering Applications. **Annual Review of Plant Physiology and Plant Molecular Biology**, v. 52, n. 1, p. 29–66, 2001.

FACCHINI, P. J.; DE LUCA, V. Opium poppy and Madagascar periwinkle: Model non-model systems to investigate alkaloid biosynthesis in plants. **Plant Journal**, v. 54, n. 4, p. 763–784, 2008.

FACCHINI, P. J.; MORRIS, J. S. Molecular Origins of Functional Diversity in Benzyloisoquinoline Alkaloid Methyltransferases. **Frontiers in Plant Science**, v. 10, 2019.

FAGGION, S. A. et al. Anticonvulsant profile of the alkaloids (+)-erythravine and (+)-11- $\alpha$ -hydroxy-erythravine isolated from the flowers of *Erythrina mulungu* Mart ex Benth (Leguminosae-Papilionaceae). **Epilepsy and Behavior**, v. 20, n. 3, p. 441–446, 2011.

FAHMY, Nouran M. et al. Alkaloids of genus *Erythrina*: An updated review. **Natural Product Research**, Cairo, p.1-22, 21 jun. 2019;

FARROW, S. C.; HAGEL, J. M.; FACCHINI, P. J. Transcript and metabolite profiling in cell cultures of 18 plant species that produce benzyloisoquinoline alkaloids. **Phytochemistry**, v. 77, p. 79–88, 2012.

FIEHN, O. et al. Metabolite profiling for plant functional genomics. **Nature Biotechnology**, v. 18, n. 11, p. 1157–1161, 2000.

FIEHN, O. Metabolomics - The link between genotypes and phenotypes. **Plant Molecular Biology**, v. 48, n. 1–2, p. 155–171, 2002.

FIETTO, J. L. R.; MACIEL, T. E. F. **Sequenciando genomas**. In: Ciências genômicas: fundamentos e aplicações. [s.l: s.n.]. p. 27–64, 2015.

FRANK, A. M. A ranking-based scoring function for peptide-spectrum matches. **Journal of Proteome Research**, v. 8, n. 5, p. 2241–2252, 2009.

FRENCH, K. E.; HARVEY, J.; MCCULLAGH, J. S. O. Targeted and Untargeted Metabolic Profiling of Wild Grassland Plants identifies Antibiotic and Anthelmintic Compounds Targeting Pathogen Physiology, Metabolism and Reproduction. **Scientific Reports**, v. 8, n. 1, p. 1–10, 2018.

FRICK, S.; KRAMELL, R.; KUTCHAN, T. M. Metabolic engineering with a morphine biosynthetic P450 in opium poppy surpasses breeding. **Metabolic Engineering**, v. 9, n. 2, p. 169–176, 2007.

FU, S. et al. De novo transcriptome analysis of tibetan medicinal plant *Dysphania schraderiana*. **Genetics and Molecular Biology**, v. 42, n. 2, p. 480–487, 2019.

FURTINO NETO, Antônio Eduardo; VALE, Fabiano Ribeiro do; RESENDE, Alvaro Vilela de. **Fertilidade do solo**. 2001. 261 f. TCC (especialização) - Curso de Fertilidade do Solo e Nutrição de Plantas, Universidade Federal de Lavras, Lavras, 2001;

- GARBER, M. et al. Computational methods for transcriptome annotation and quantification using RNA-seq. **Nature methods**, v. 8, n. 6, p. 469–77, 2011.
- GERARDY, R.; ZENK, M. H. Purification and characterization of salutaridine: NADPH 7-oxidoreductase from *Papaver somniferum*. **Phytochemistry**, v. 34, n. 1, p. 125–132, 1993.
- GESELL, A. et al. CYP719B1 is salutaridine synthase, the C-C phenol-coupling enzyme of morphine biosynthesis in opium poppy. **Journal of Biological Chemistry**, v. 284, n. 36, p. 24432–24442, 2009.
- GLEDHILL, David. **The names of plants**. Cambridge: Cambridge University Press, 2008. 436 p.
- GROTHER, T.; LENZ, R.; KUTCHAN, T. M. Molecular Characterization of the Salutaridinol 7-O-Acetyltransferase Involved in Morphine Biosynthesis in Opium Poppy *Papaver somniferum*. **Journal of Biological Chemistry**, v. 276, n. 33, p. 30717–30723, 2001.
- GURKOK, T. et al. Functional Characterization of 4'OMT and 7OMT Genes in BIA Biosynthesis. **Frontiers in Plant Science**, v. 7, n. February, p. 1–11, 2016.
- HAGEL, J. M. et al. Transcriptome analysis of 20 taxonomically related benzyloisoquinoline alkaloid-producing plants. **BMC plant biology**, v. 15, n. 1, p. 227, 2015.
- HAGEL, J. M.; FACCHINI, P. J. Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. **Nature Chemical Biology**, v. 6, n. 4, p. 273–275, 2010.
- HALL, R. D. Plant metabolomics: From holistic hope, to hype, to hot topic. **New Phytologist**, v. 169, n. 3, p. 453–468, 2006.
- HELDT, H.-W.; PIECHULLA, B. **Plant Biochemistry**. Translation of the 4th German edition. [s.l.: s.n.].
- HENINNG, Fernando Augusto et al. Predição in silico de marcadores microssatélites relacionados ao tegumento de sementes de soja. **Revista Brasileira de Sementes**, Pelotas, v. 31, n. 4, p.49-56, 2009;
- HIGASHI, Y.; KUTCHAN, T. M.; SMITH, T. J. Atomic structure of salutaridine reductase from the opium poppy (*Papaver somniferum*). **Journal of Biological Chemistry**, v. 286, n. 8, p. 6532–6541, 2011.
- HILDONEN, S.; HALVORSEN, T. G.; REUBSAET, L. Why less is more when generating tryptic peptides in bottom-up proteomics. **Proteomics**, v. 14, n. 17–18, p. 2031–2041, 2014.
- HORI, Kentaro et al. Mining of the Uncharacterized Cytochrome P450 Genes Involved in Alkaloid Biosynthesis in California Poppy Using a Draft Genome Sequence. **Plant and Cell Physiology**, v. 59, n. 2, p.222-233, 29 dez. 2017.

- HUSSAIN, M. M. et al. Constituents of Erythrina - a Potential Source of Secondary Metabolites: A Review. **Bangladesh Pharmaceutical Journal**, v. 19, n. 2, p. 237–253, 2016.
- HUTCHISON, C. A. DNA sequencing: Bench to bedside and beyond. **Nucleic Acids Research**, v. 35, n. 18, p. 6227–6237, 2007.
- IKEZAWA, N.; IWASA, K.; SATO, F. CYP719A subfamily of cytochrome P450 oxygenases and isoquinoline alkaloid biosynthesis in *Eschscholzia californica*. **Plant Cell Reports**, v. 28, n. 1, p. 123–133, 2009.
- IKEZAWA, N.; IWASA, K.; SATO, F. Molecular cloning and characterization of methylenedioxy bridge-forming enzymes involved in stylophine biosynthesis in *Eschscholzia californica*. **FEBS Journal**, v. 274, n. 4, p. 1019–1035, 2007.
- IMHOFF, R. D. et al. General base catalysis in the urate oxidase reaction: Evidence for a novel Thr-Lys catalytic diad. **Biochemistry**, v. 42, n. 14, p. 4094–4100, 2003.
- International Union of Biochemistry and Molecular Biology. 2016. Disponível em: <https://www.qmul.ac.uk/sbcs/iubmb/>. Acesso em: 21 de novembro 2019.
- KANG, H. et al. Glaucine inhibits breast cancer cell migration and invasion by inhibiting MMP-9 gene expression through the suppression of NF- $\kappa$ B activation. **Molecular and Cellular Biochemistry**, v. 403, n. 1–2, p. 85–94, 2015.
- KATAJAMAA, Mikko; ORESIC, Matej. Processing methods for differential analysis of LC/MS profile data. **Bmc Bioinformatics**, v. 6, n. 1, p.179-190, 2005.
- KHAN, M. A. et al. Transcriptome profiling of faba bean (*Vicia faba* L.) drought-tolerant variety hassawi-2 under drought stress using RNA sequencing. **Electronic Journal of Biotechnology**, v. 39, p. 15–29, 2019.
- KILGORE, Matthew B. et al. Cloning and Characterization of a Norbelladine 4'-O-Methyltransferase Involved in the Biosynthesis of the Alzheimer's Drug Galanthamine in *Narcissus* sp. aff. *pseudonarcissus*. **Plos One**, United States of America, v. 9, n. 7, p.1-11, 25 jul. 2014.
- KIRCHER, M.; STENZEL, U.; KELSO, J. Improved base calling for the Illumina Genome Analyzer using machine learning strategies. **Genome Biology**, v. 10, n. 8, p. 1–9, 2009.
- KUMAR, Ravi et al. A High-Throughput Method for Illumina RNA-Seq Library Preparation. **Frontiers In Plant Science**, v. 3, p.1-11, ago. 2012.
- LANÇAS, F. M. A cromatografia líquida moderna e a espectrometria de massas: Finalmente “compatíveis”? **Scientia Chromatographica**, v. 5, n. 1, p. 27–46, 2009.
- LANGMEAD, B.; SALZBERG, S. L. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, v. 9, n. 4, p. 357–359, 2012.
- LARKIN, Philip J. et al. Increasing morphinan alkaloid production by over-expressing codeinone reductase in transgenic *Papaver somniferum*. **Plant Biotechnology Journal**, Canberra, v. 5, n. 1, p.26-37, jan. 2007.

- LEE, E. J.; FACCHINI, P. Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family. **Plant Cell**, v. 22, n. 10, p. 3489–3503, 2010.
- LEI, Z.; HUHMANN, D. V.; SUMNER, L. W. Mass spectrometry strategies in metabolomics. **Journal of Biological Chemistry**, v. 286, n. 29, p. 25435–25442, 2011.
- LI, W. et al. The versatile O-methyltransferase LrOMT catalyzes multiple O-methylation reactions in amaryllidaceae alkaloids biosynthesis. **International Journal of Biological Macromolecules**, v. 141, p. 680–692, 2019.
- LIMA, Diogo Borges. **Método computacional para identificação de peptídeos marcados com fenil-isotiocianato e analisados por cromatografia líquida acoplada a espectrometria de massa em tandem**. 2013. 61 f. Tese (Doutorado) - Curso de Engenharia de Sistemas e Computação., Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013.
- LIN, S. et al. LC/MS-based non-targeted metabolomics for the investigation of general toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in C57BL/6J and DBA/2J mice. **International Journal of Mass Spectrometry**, v. 301, n. 1–3, p. 29–36, 2011.
- LINDON, J. C.; NICHOLSON, J. K. Analytical technologies for metabonomics and metabolomics, and multi-omic information recovery. **TrAC - Trends in Analytical Chemistry**, v. 27, n. 3, p. 194–204, 2008.
- LISCOMBE, D. K. et al. Evidence for the monophyletic evolution of benzyloisoquinoline alkaloid biosynthesis in angiosperms. **Phytochemistry**, v. 66, n. 20, p. 2501–2520, 2005.
- LISCOMBE, D. K.; FACCHINI, P. J. Molecular cloning and characterization of tetrahydroprotoberberine cis-N-methyltransferase, an enzyme involved in alkaloid biosynthesis in opium poppy. **Journal of Biological Chemistry**, v. 282, n. 20, p. 14741–14751, 2007.
- LIU, W. et al. From <sup>1</sup>H NMR-based non-targeted to LC–MS-based targeted metabolomics strategy for in-depth chemome comparisons among four Cistanche species. **Journal of Pharmaceutical and Biomedical Analysis**, v. 162, p. 16–27, 2019.
- LISEC, Jan; SCHAUER, Nicolas; KOPKA, Joachim; WILLMITZER, Lothar; FERNIE, Alisdair R. Erratum: Corrigendum. **Nature Protocols**, v. 10, n. 9, p.1457-1457, 2015.
- LÓPEZ, S. et al. Acetylcholinesterase inhibitory activity of some Amaryllidaceae alkaloids and Narcissus extracts. **Life Sciences**, v. 71, n. 21, p. 2521–2529, out. 2002.
- LPWG, T. L. P. W. et al. Legume phylogeny and classification in the 21st century: Progress, prospects and lessons for other species-rich clades. **Taxon**, v. 62, n. 2, p. 217–248, 20 abr. 2013.
- LPWG, T. L. P. W. G. et al. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. **Taxon**, v. 66, p. 44–77, 2017.
- MAIER, U. H. et al. Biosynthesis of Erythrina alkaloids in Erythrina crista-galli. **Phytochemistry**, v. 52, n. 3, p. 373–382, 1999.

- MARCHIORO, M. et al. Anti-nociceptive activity of the aqueous extract of *Erythrina velutina* leaves. **Fitoterapia**, v. 76, n. 7–8, p. 637–642, 2005.
- MARDIS, E. R. Next-Generation DNA Sequencing Methods. **Annual Review of Genomics and Human Genetics**, v. 9, n. 1, p. 387–402, set. 2008.
- MELANI, R. D. **Estratégias proteômicas “BOTTOM-UP ” E “ TOP-DOWN ” aplicadas à genômica**. Tese (doutorado) – Bioquímica, Universidade Federal do Rio de Janeiro, Natal, 2018; p. 1–202, 2016.
- MOHAMMED, Magdy M.d. et al. Anti-HIV-1 and cytotoxicity of the alkaloids of *Erythrina abyssinica* Lam. growing in Sudan. **Natural Product Research**, Cairo, v. 26, n. 17, p.1565-1575, set. 2012;
- MORITA, M.; SHITAN, N.; SAWADA, K.; VAN MONTAGU, M. C. E.; INZE, D.; RISCHER, H.; GOOSSENS, A.; OKSMAN-CALDENTEY, K.-m.; MORIYAMA, Y.; YAZAKI, K.. Vacuolar transport of nicotine is mediated by a multidrug and toxic compound extrusion (MATE) transporter in *Nicotiana tabacum*. **Proceedings of The National Academy Of Sciences**, v. 106, n. 7, p.2447-2452, 23 jan. 2009.
- MORISHIGE, Takashi et al. Molecular Characterization of the S-Adenosyl-l-methionine: 3'-Hydroxy-N-methylcoclaurine 4'-O-Methyltransferase Involved in Isoquinoline Alkaloid Biosynthesis in *Coptis japonica*. **Journal of Biological Chemistry**, Kyoto, v. 275, n. 30, p.23398-23405, 12 maio 2000.
- N. SHITAN; K. YAZAKI. Accumulation and Membrane Transport of Plant Alkaloids. **Current Pharmaceutical Biotechnology**, v. 8, n. 4, p. 244–252, 2007.
- NARCROSS, L. et al. Microbial Factories for the Production of Benzylisoquinoline Alkaloids. **Trends in Biotechnology**, v. 34, n. 3, p. 228–241, 2016.
- NELSON, David R.. Cytochrome P450 diversity in the tree of life. *Biochimica Et Biophysica Acta (bba) - Proteins and Proteomics*, Memphis, v. 1866, n. 1, p.141-154, jan. 2018;
- NESATYY, Victor J.; SUTER, Marc J.-f.. Analysis of environmental stress response on the proteome level. **Mass Spectrometry Reviews**, Switzerland, v. 27, n. 6, p.556-574, nov. 2008.
- NOMURA, T.; KUTCHAN, T. M. Three new O-methyltransferases are sufficient for all O-methylation reactions of ipecac alkaloid biosynthesis in root culture of *Psychotria ipecacuanha*. **Journal of Biological Chemistry**, v. 285, n. 10, p. 7722–7738, 2010.
- NYRÉN, P.; PETTERSSON, B.; UHLÉN, M. Solid phase DNA minisequencing by an enzymatic luminometric inorganic pyrophosphate detection assay **Analytical Biochemistry**, 1993.
- OLIVEIRA, J. H. G. DE; IWAZAKI, M. C.; OLIVEIRA, D. M. T. Análises ontogenéticas do pericarpo de espécies de *Mimosa* Linnaeus (Fabaceae, Mimosoideae). **Hoehnea**, v. 41, n. 4, p. 483–497, 2014.
- OTANI, Mihoko; SHITAN, Nobukazu; SAKAI, Kyoko; MARTINOIA, Enrico; SATO, Fumihiko; YAZAKI, Kazufumi. Characterization of Vacuolar Transport of the

Endogenous Alkaloid Berberine in *Coptis japonica*. **Plant Physiology**, v. 138, n. 4, p.1939-1946, 15 jul. 2005.

OTAVIO APARECIDO FLAUSINO et al. Effects of Erythrinian Alkaloids Isolated from *Erythrina mulungu* (Papilionaceae) in Mice Submitted to Animal Models of Anxiety. **Biological and Pharmaceutical Bulletin**, v. 30, p. 375–378, 2007.

OUNAROON, A. et al. (R,S)-Reticuline 7-O-methyltransferase and (R,S)-norcoclaurine 6-O-methyltransferase of *Papaver somniferum* - cDNA cloning and characterization of methyl transfer enzymes of alkaloid biosynthesis in opium poppy. **Plant Journal**, v. 36, n. 6, p. 808–819, 2003.

OZAWA, M. et al. Bioorganic & Medicinal Chemistry Letters TRAIL-enhancing activity of Erythrinan alkaloids from *Erythrina velutina*. **Bioorganic and Medicinal Chemistry Letters**, v. 19, n. 1, p. 234–236, 2009.

PAGLIOSA, L. B. et al. Effect of isoquinoline alkaloids from two *Hippeastrum* species on in vitro acetylcholinesterase activity. **Phytomedicine**, v. 17, n. 8–9, p. 698–701, 2010.

PARSONS, A. F.; PALFRAMAN, M. J. Chapter 1 *Erythrina* and Related Alkaloids. In: *Alkaloids: Chemistry and Physiology*. [s.l: s.n.]. v. 18p. 1–98.

PATHAK, S. et al. Comparative Transcriptome Analysis Using High Papaverine Mutant of *Papaver somniferum* Reveals Pathway and Uncharacterized Steps of Papaverine Biosynthesis. **PLOS ONE**, v. 8, n. 5, 2013.

PATRO, R. et al. Salmon provides fast and bias-aware quantification of transcript expression. **Nature Methods**, v. 14, n. 4, p. 417–419, 2017.

PENG, L. et al. Antibacterial activity and mechanism of berberine against *Streptococcus agalactiae*. **International Journal of Clinical and Experimental Pathology**, v. 8, n. 5, p. 5217–5223, 2015.

PERGANDE, M. R.; COLOGNA, S. M. Isoelectric point separations of peptides and proteins. **Proteomes**, v. 5, n. 1, 2017.

PINTO, A. C. et al. Produtos naturais: Atualidade, desafios e perspectivas. **Quimica Nova**, v. 25, n. SUPPL. 1, p. 45–61, 2002.

PRADO, D.E. 2003. **As caatingas da América do Sul**. Pp. 3-73. In: I.R. LEAL, M. TABARELLI, J.M.C. SILVA (eds.). *Ecologia e conservação da Caatinga*. Recife, Editora Universitária, Universidade Federal de Pernambuco.

PROSDÓCIMI, Francisco; MOREIRA, Leandro Marcio. **Genômica comparativa**. São Paulo: Cubo, 2016.

PURWANTO, Ratmoyo et al. Unraveling Additional O-Methylation Steps in Benzyloisoquinoline Alkaloid Biosynthesis in California Poppy (*Eschscholzia californica*). **Plant and Cell Physiology**, Kyoto, v. 58, n. 9, p.1528-1540, 11 jul. 2017.

RAJBIR KAUR; SAROJ ARORA. Alkaloids-Important Therapeutic Secondary Metabolites of Plant. **Journal of Critical Reviews**, v. 2, n. 3, p. 1–8, 2015.

- RAMBO, D. F. et al. The genus *Erythrina* L.: A review on its alkaloids, preclinical, and clinical studies. **Phytotherapy Research**, v. 33, p. 1–19, 2019.
- RANDHIR, R.; LIN, Y. T.; SHETTY, K. Stimulation of phenolics, antioxidant and antimicrobial activities in dark germinated mung bean sprouts in response to peptide and phytochemical elicitors. **Process Biochemistry**, v. 39, n. 5, p. 637–646, 2004.
- RAUPP, I. M. et al. Anxiolytic-like effect of chronic treatment with *Erythrina velutina* extract in the elevated plus-maze test. **Journal of Ethnopharmacology**, v. 118, n. 2, p. 295–299, 2008.
- RODRIGUES, D. R. et al. Phenotypic, genetic and symbiotic characterization of *Erythrina velutina* rhizobia from Caatinga dry forest. **Brazilian Journal of Microbiology**, v. 49, n. 3, p. 503–512, 2018.
- RODRIGUES, F. T. S. et al. Effects of standard ethanolic extract from *Erythrina velutina* in acute cerebral ischemia in mice. **Biomedicine and Pharmacotherapy**, v. 96, n. November, p. 1230–1239, 2017.
- SALVATO, Fernanda; CARVALHO, Mayra Costa da Cruz Gallo de. **Métodos e estratégias em proteômica e suas aplicações na área vegetal**. Ciência Rural, Santa Maria, v. 40, n. 3, p.727-734, 2010.
- SALZBERG, S. L.; POP, M. Bioinformatics challenges of new sequencing technology. **Trends in Genetics**, v. 24, n. 3, p. 142–149, 2008.
- SANTOS ROSA, D. et al. Erysothrine, an alkaloid extracted from flowers of *Erythrina mulungu* Mart. ex Benth: Evaluating its anticonvulsant and anxiolytic potential. **Epilepsy and Behavior**, v. 23, n. 3, p. 205–212, mar. 2012.
- SCHWARZBACH, E. et al. Gregor J. Mendel - Genetics founding father. **Czech Journal of Genetics and Plant Breeding**, v. 50, n. 2, p. 43–51, 2014.
- SETTI-PERDIGÃO, P. et al. *Erythrina* mulungu Alkaloids Are Potent Inhibitors of Neuronal Nicotinic Receptor Currents in Mammalian Cells. **PLOS ONE**, v. 8, n. 12, p. e82726, 13 dez. 2013.
- SEVERIN, A. J. et al. RNA-Seq Atlas of *Glycine max*: A guide to the soybean transcriptome. **BMC Plant Biology**, v. 10, n. 2007, 2010.
- SHI, J. et al. Enzymatic conversion of carbon dioxide. **Chemical Society Reviews**, v. 44, n. 17, p. 5981–6000, 2015.
- SHTARK, Oksana Y.; BORISOV, Alexey Y.; ZHUKOV, Vladimir A.. Intimate Associations of Beneficial Soil Microbes with Host Plants. In: DIXON, Geoff R; TILSTON, Emma L. **Soil Microbiology and Sustainable Crop Production**. Russia: Springer, 2010. p. 119-196;
- SILVA, M. R. DA. Rosalind Franklin e seu papel na construção do modelo da dupla hélice do DNA Rosalind Franklin e seu papel na construção do modelo da dupla-hélice do DNA. **Filosofia e História da Biologia**, v. 2, p. 297–310, 2007.

- SIMÃO, Felipe A. et al. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. **Bioinformatics**, v. 31, n. 19, p. 3210-3212, 2015.
- SOHRABI, S. M.; ISMAILI, A.; NAZARIAN-FIROUZABADI, F. Simultaneous over-expression and silencing of some benzylisoquinoline alkaloid biosynthetic genes in opium poppy. **Industrial Crops and Products**, v. 123, n. 1, p. 581–590, 2018.
- SOUZA, Luciani de Liz; RHODEN, Sandro Augusto; PAMPHILE, João Alencar. The importance of omics as a tool for the study of exploration of microorganisms: prospects and challenges. **Revista Uningá Review**, v. 18, n. 2, p.16-21, 2014.
- SOUZA, Luciani de Liz; RHODEN, Sandro; PAMPHILE, João. A importância das ômicas como ferramentas para o estudo da prospecção de microrganismos: perspectivas e desafios. **Revista Uningá Review**, Maringá, v. 18, n. 2, p.1-5, abr. 2014;
- STEIN, Lincoln. Genome annotation: from sequence to biology. **Nature Reviews Genetics**, Bungtown, v. 2, n. 7, p.493-503, jul. 2001;
- STOUT, C.d.. Cytochrome P450 Conformational Diversity. **Structure**, v. 12, n. 11, p.1921-1922, nov. 2004.
- TAN, Qing-wei et al. A New Erythrinan Alkaloid Glycoside from the Seeds of *Erythrina crista-galli*. **Molecules**, Fujian, v. 22, n. 9, p.1558-1564, 16 set. 2017;
- T AQUARY, Adriana Maria Antunes. **Tamanho, montagem de novo e anotação do genoma de *Dipteryx alata* (leguminosa)**. 2017. 137 f. Tese (Doutorado) - Curso de Ciências Biológicas, Universidade Federal de Goiás, Goiania, 2017;
- TORRES PAZMIÑO, D. E. et al. Monooxygenases as biocatalysts: Classification, mechanistic aspects and biotechnological applications. **Journal of Biotechnology**, v. 146, n. 1–2, p. 9–24, 2010.
- UNTERLINNER, B.; LENZ, R.; KUTCHAN, T. M. Molecular cloning and functional expression of codeinone reductase: The penultimate enzyme in morphine biosynthesis in the opium poppy *Papaver somniferum*. **Plant Journal**, v. 18, n. 5, p. 465–475, 1999.
- VALLEDOR, L.; JORRÍN, J. Back to the basics: Maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. **Journal of Proteomics**, v. 74, n. 1, p. 1–18, 2011.
- VASCONCELOS, S. M. M. et al. Anticonvulsant activity of hydroalcoholic extracts from *Erythrina velutina* and *Erythrina mulungu*. **Journal of Ethnopharmacology**, v. 110, n. 2, p. 271–274, 2007.
- VATANSEVER, R.; OZYIGIT, I. I.; FILIZ, E. Essential and Beneficial Trace Elements in Plants, and Their Transport in Roots: a Review. **Applied Biochemistry and Biotechnology**, v. 181, n. 1, p. 464–482, 29 jan. 2017;
- VILLAS-BÔAS, S. G. et al. Mass spectrometry in metabolome analysis. **Mass Spectrometry Reviews**, v. 24, n. 5, p. 613–646, 2005.

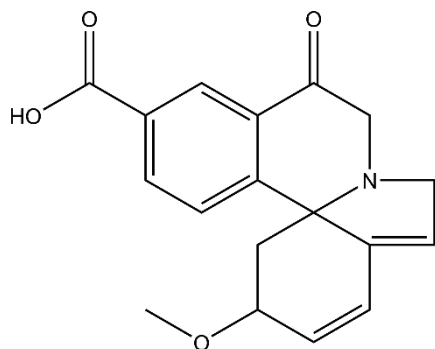
- VOGEL, M. et al. Structure and mechanism of sanguinarine reductase, an enzyme of alkaloid detoxification. **Journal of Biological Chemistry**, v. 285, n. 24, p. 18397–18406, 2010.
- WANG, Z.; GERSTEIN, M.; SNYDER, M. Nihms229948. **Nature Reviews Genetics**, v. 10, n. 1, p. 57–63, 2009.
- WATERHOUSE, Robert M. et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. **Molecular biology and evolution**, v. 35, n. 3, p. 543-548, 2017.
- WILKINS, M. R. et al. Current challenges and future applications for protein maps and post-translational vector maps in proteome projects. **Electrophoresis**, v. 17, n. 5, p. 830–838, 1996.
- WINK, M. Evolution of secondary metabolites in legumes (Fabaceae). **South African Journal of Botany**, v. 89, p. 164–175, 2013.
- WINKLER, A. et al. Berberine bridge enzyme catalyzes the six electron oxidation of (S)-reticuline to dehydroscoulerine. **Phytochemistry**, v. 70, n. 9, p. 1092–1097, 2009.
- WISHART, D. S. et al. HMDB: the Human Metabolome Database. **Nucleic Acids Research**, v. 35, n. Database, p. 521–526, 3 jan. 2007.
- WU, Jing et al. A new erythrinan N-oxide alkaloid from *Erythrina stricta*. **Natural Product Research**, Kunming, v. 33, n. 14, p.2004-2010, 11 jun. 2018;
- XIMENES, N. C. et al. Ethanolic extract of *Erythrina velutina* Willd ameliorate schizophrenia-like behavior induced by ketamine in mice. **Journal of Complementary and Integrative Medicine**, v. 16, n. 2, p. 1–8, 2018.
- YOBI, A. et al. Comparative metabolic profiling between desiccation-sensitive and desiccation-tolerant species of Selaginella reveals insights into the resurrection trait. **Plant Journal**, v. 72, n. 6, p. 983–999, 2012.
- YOBI, A. et al. Metabolomic profiling in selaginella lepidophylla at various hydration states provides new insights into the mechanistic basis of desiccation tolerance. **Molecular Plant**, v. 6, n. 2, p. 369–385, 2013.
- ZHANG, Y. et al. Enzymatic biosynthesis of benzylisoquinoline alkaloid glycosides via promiscuous glycosyltransferases from *Carthamus tinctorius*. **Chinese Chemical Letters**, v. 30, n. 2, p. 443–446, 2019.
- ZHAO, Hao-en et al. Alkaloids from flowers of *Erythrina corallodendron*. **Natural Product Research**, Helfel, v. 33, n. 9, p.1298-1303, 6 jun. 2018;
- ZHERNAKOV, A. et al. Massive Analysis of cDNA Ends (MACE) for transcript-based marker design in pea (*Pisum sativum* L.). **Genomics Data**, v. 11, p. 75–76, 2017.
- ZULAK, K. G. et al. Plant defense responses in opium poppy cell cultures revealed by liquid chromatography-tandem mass spectrometry proteomics. **Molecular and Cellular Proteomics**, v. 8, n. 1, p. 86–98, 2009.

**ANEXO I***Material Suplementar*

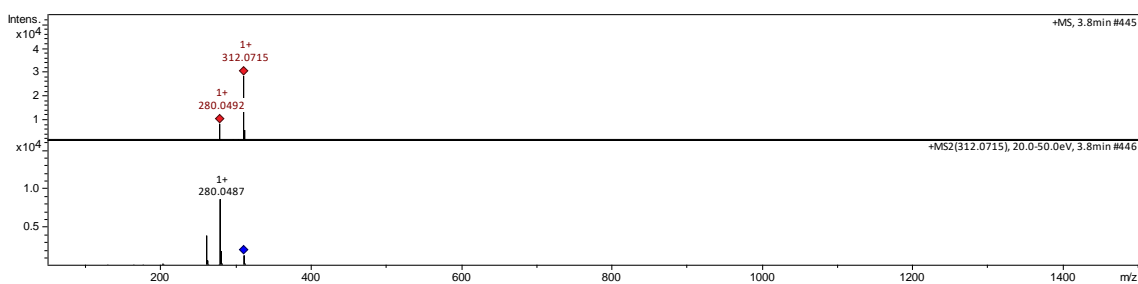
**Tabela S1.** Alcaloides de *Erythrina velutina* anotados pela estratégia HR-LC-MS/MS**N° Structure and chemical name (Cluster 1) and mass spectra (MS and MS/MS)**

**1** RT - 3.8 min  
 $m/z$  312.0715  $[M+H]^+$

A)



11-methoxy-5-oxo-2H,4H,5H,10H,11H-indolo[7a,1-a]isoquinoline-7-carboxylic acid

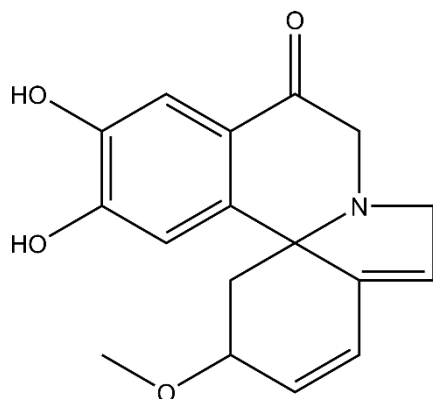


Pool \ Fragments	312.0738	280.0508	262.0411	204.0464
p7(a)	•	•	•	•
p7(b)	•	•	•	•
p7(c)	•	•	•	-

2

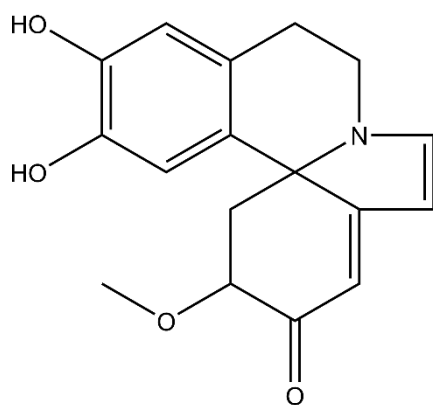
RT - 4.4 min  
 $m/z$  300.1101  $[M+H]^+$ 

A)

7,8-dihydroxy-11-methoxy-  
2H,4H,5H,10H,11H-indolo[7a,1-a]isoquinolin-  
5-one

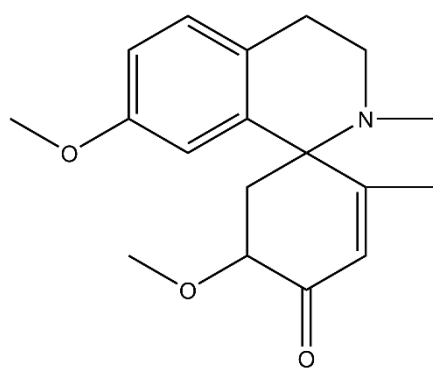
(11-Oxoerysopine)

B)

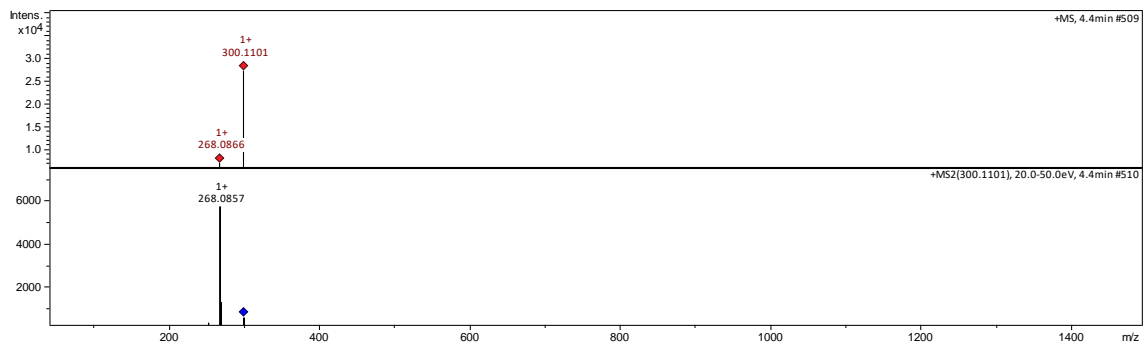
7,8-dihydroxy-11-methoxy-  
4H,5H,10H,11H,12H-indolo[7a,1-  
a]isoquinolin-12-one

(2,3-Demethoxyerythratidinone)

C)

8,11-dimethoxy-1H,2H,4H,5H,10H,11H,12H-  
indolo[7a,1-a]isoquinolin-12-one

(Coccudinone)

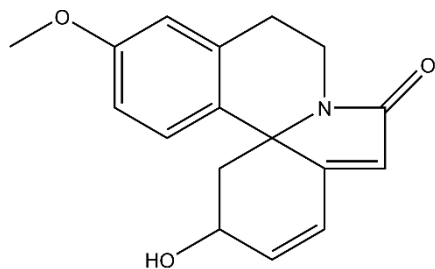


Pool	Fragments	
	300.1099	268.0893
p7(a)	●	●
p7(b)	●	●
p7(c)	●	●
p6(c)	●	●
p8(b)	●	●

3

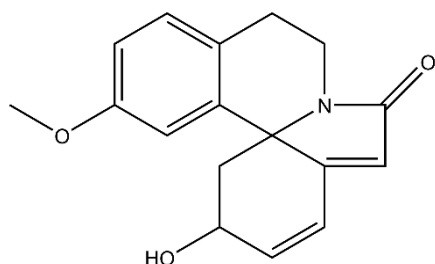
RT - 5.1 min  
 $m/z$  284.0787  $[M+H]^+$

A)

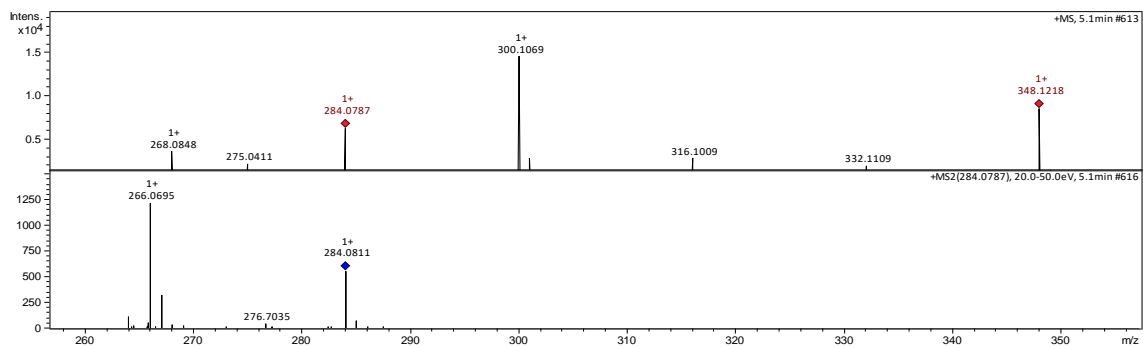


11-hydroxy-7-methoxy-2H,4H,5H,10H,11H-  
 indolo[7a,1-a]isoquinolin-2-one

B)



11-hydroxy-8-methoxy-2H,4H,5H,10H,11H-  
 indolo[7a,1-a]isoquinolin-2-one

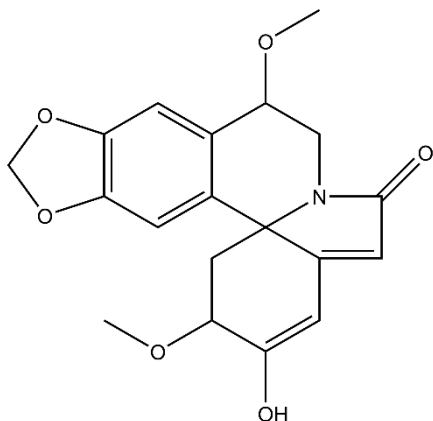


Pool	Fragments	
	284.0796	266.0715
p3(b)	●	●
p3(c)	●	●
p2(a)	●	●
p2(b)	●	●
p2(c)	●	●
p4(a)	●	●
p4(b)	●	●
p4(c)	●	●
p1(a)	●	●

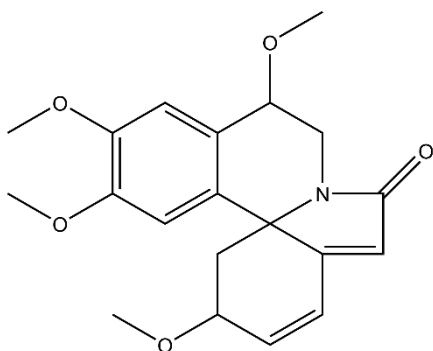
4

RT - 5.5 min  
 $m/z$  358.1098  $[M+H]^+$ 

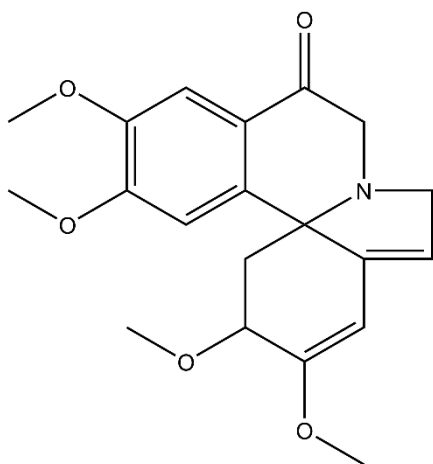
A)

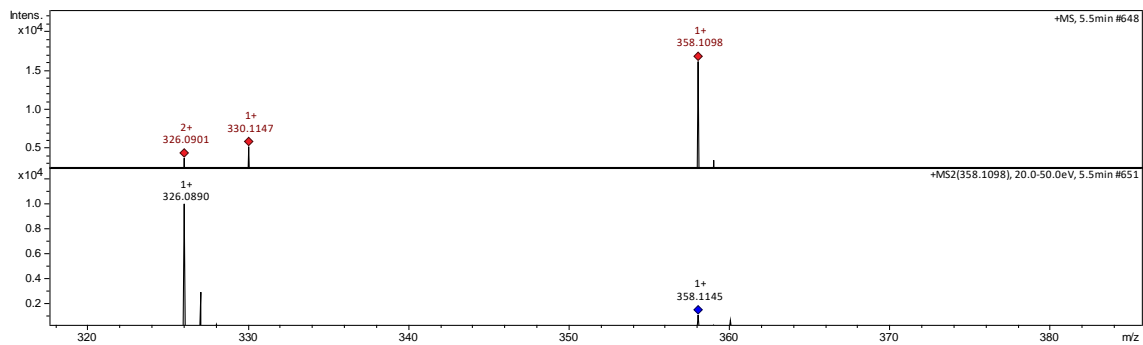
18-hydroxy-11,19-dimethoxy-5,7-dioxo-  
13-azapentacyclo[11.7.0.0<sup>1,16</sup>.0<sup>2,10</sup>.0<sup>4,8</sup>]icosa-  
2,4(8),9,15,17-pentaen-14-one

B)

5,7,8,10-tetramethoxy-  
2H,4H,5H,10H,11H-indolo[7a,1-a]isoquinolin-  
2-one

C)

7,8,11,12-tetramethoxy-2H,4H,5H,10H,11H-  
indolo[7a,1-a]isoquinolin-5-one

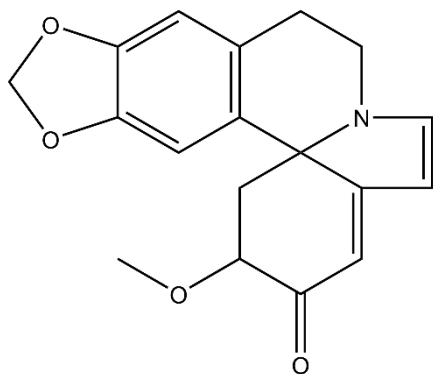


Pool	Fragments	358.1139	326.0874	267.0807
	p7(a)		•	•
P6(c)		•	•	•

5

RT - 3.8 min  
 $m/z$  312.0738 [M+H]<sup>+</sup>

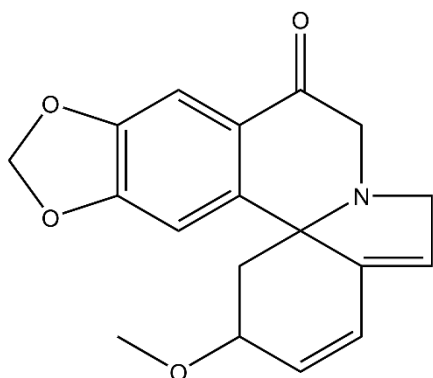
A)



19-methoxy-5,7-dioxo-13-azapentacyclo[11.7.0.0<sup>1</sup>,<sup>16</sup>.0<sup>2</sup>,<sup>10</sup>.0<sup>4</sup>,<sup>8</sup>]icosa-2(10),3,8,14,16-pentaen-18-one

(3-Oxo-Erythraline)

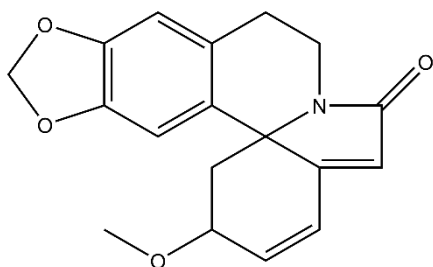
B)



19-methoxy-5,7-dioxo-13-azapentacyclo[11.7.0.0<sup>1</sup>,<sup>16</sup>.0<sup>2</sup>,<sup>10</sup>.0<sup>4</sup>,<sup>8</sup>]icosa-2,4(8),9,15,17-pentaen-11-one

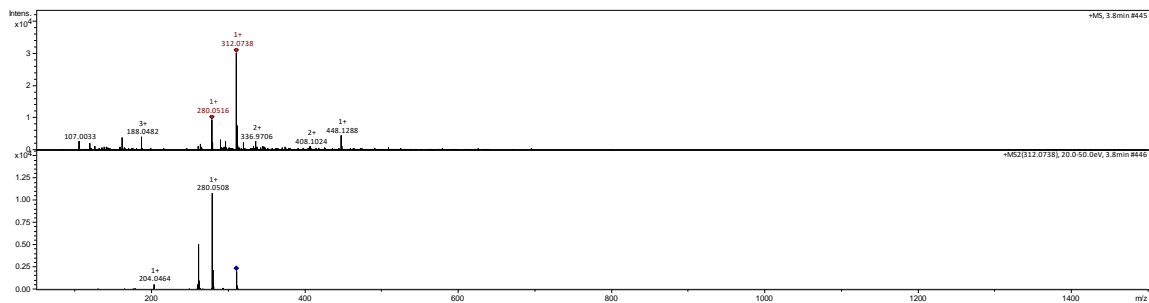
(11-Oxo-Erythraline)

C)



19-methoxy-5,7-dioxo-13-azapentacyclo[11.7.0.0<sup>1</sup>,<sup>16</sup>.0<sup>2</sup>,<sup>10</sup>.0<sup>4</sup>,<sup>8</sup>]icosa-2,4(8),9,15,17-pentaen-14-one

(8-Oxo-Erythraline)

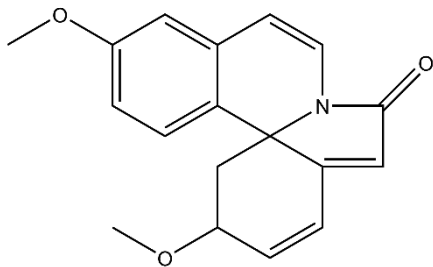


Pool \ Fragments	312.0755	280.0507	262.0424
p7(a)	●	●	●
p7(b)	●	●	●
p7(c)	●	●	●
p6(b)	●	●	●

6

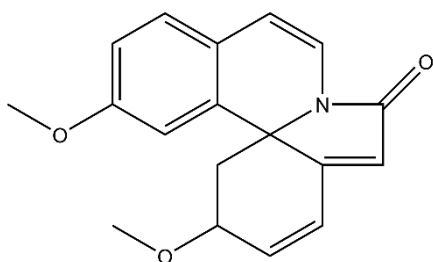
 RT - 6.3 min  
 $m/z$  296.0805  $[M+H]^+$ 

A)

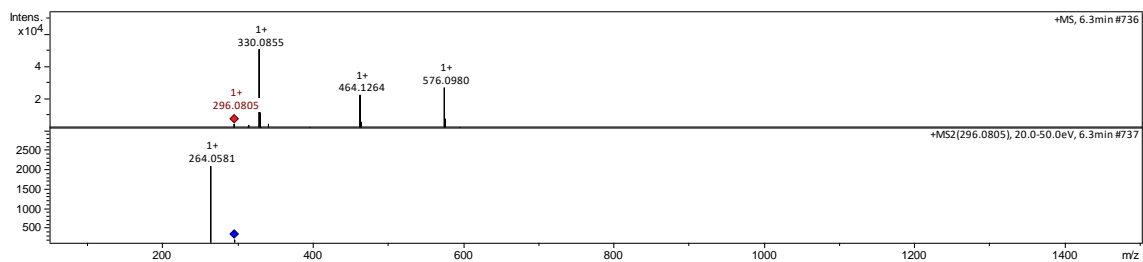


7,11-dimethoxy-2H,10H,11H-indolo[7a,1-a]isoquinolin-2-one

B)



8,11-dimethoxy-2H,10H,11H-indolo[7a,1-a]isoquinolin-2-one

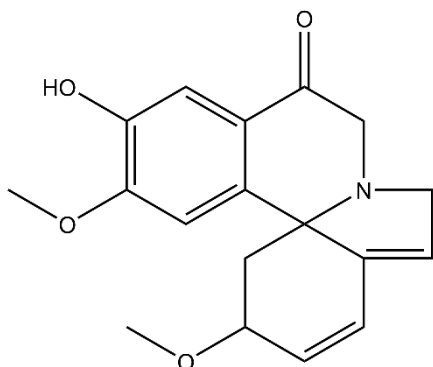


pool	Fragments	
	296.0769	264.0813
p7(a)	●	●
p7(b)	●	●
p7(c)	●	●
p2(b)	●	●

7

RT - 6.7 min  
 $m/z$  314.0894  $[M+H]^+$

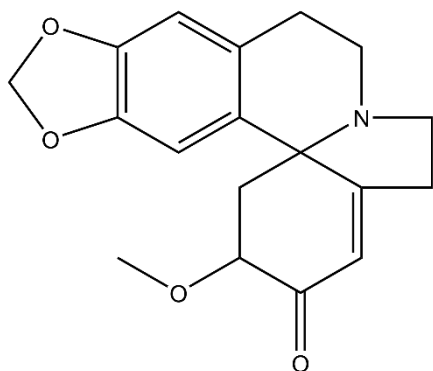
A)



7-hydroxy-8,11-dimethoxy-  
2H,4H,5H,10H,11H-indolo[7a,1-a]isoquinolin-  
5-one

(11-Oxoerysodine)

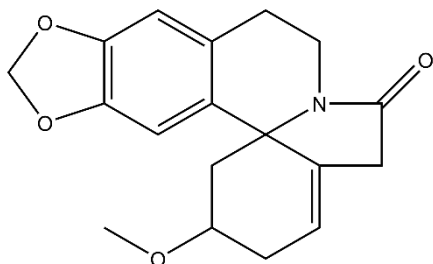
B)



19-methoxy-5,7-dioxa-13-  
azapentacyclo[11.7.0.0<sup>1,16</sup>.0<sup>2,10</sup>.0<sup>4,8</sup>]icosa-  
2,4(8),9,16-tetraen-18-one

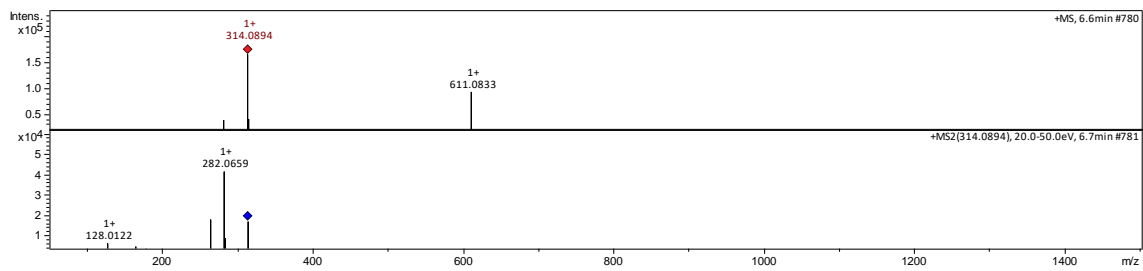
(Erythratinone)

C)



19-methoxy-5,7-dioxa-13-  
azapentacyclo[11.7.0.0<sup>1,16</sup>.0<sup>2,10</sup>.0<sup>4,8</sup>]icosa-  
2,4(8),9,16-tetraen-14-one

(8-Oxoerysodine)

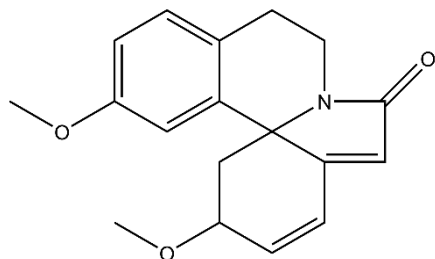


Fragments Pool	314.0874	282.0638	264.0592	165.0381	128.0122
p7(a)	●	●	●	-	-
p7(b)	●	●	●	●	●
p7(c)	●	●	●	●	●

8

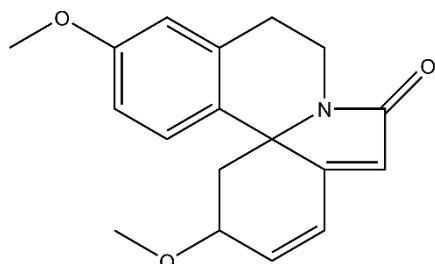
RT - 8.2 min  
 $m/z$  298.0966  $[M+H]^+$ 

A)

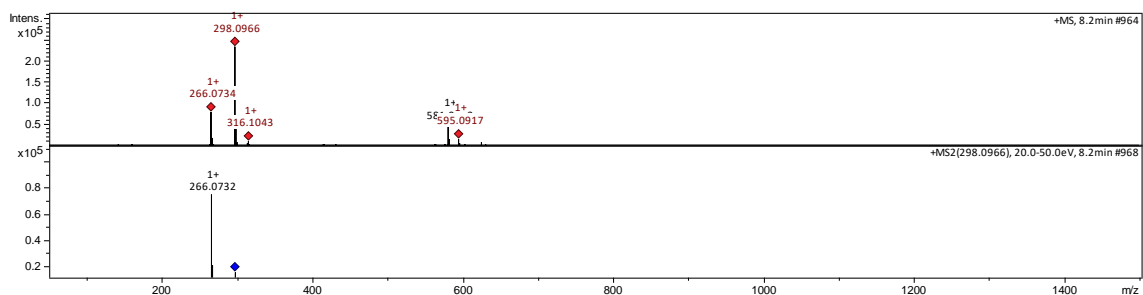
8,11-dimethoxy-2H,4H,5H,10H,11H-  
indolo[7a,1-a]isoquinolin-2-one

(Coccolinine)

B)

7,11-dimethoxy-2H,4H,5H,10H,11H-  
indolo[7a,1-a]isoquinolin-2-one

(Isococcolinine)

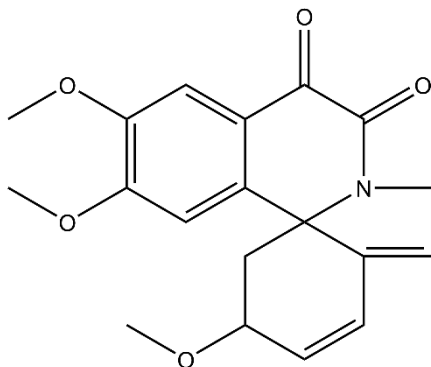


Pool	Fragments	
	298.0941	266.0721
p7(a)	•	•
p7(b)	•	•
p7(c)	•	•
p8(b)	•	•
p8(c)	•	•
p5(b)	•	•
p5(c)	•	•
p3(a)	•	•

9

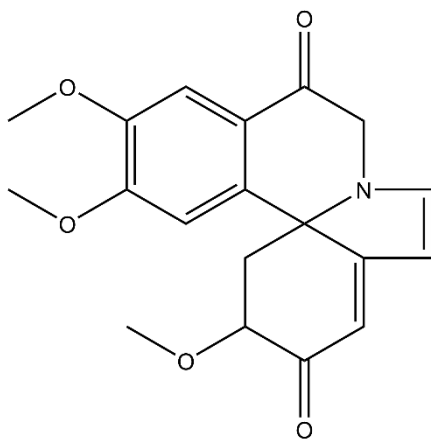
RT - 9.9 min  
 $m/z$  342.0816  $[M+H]^+$

A)

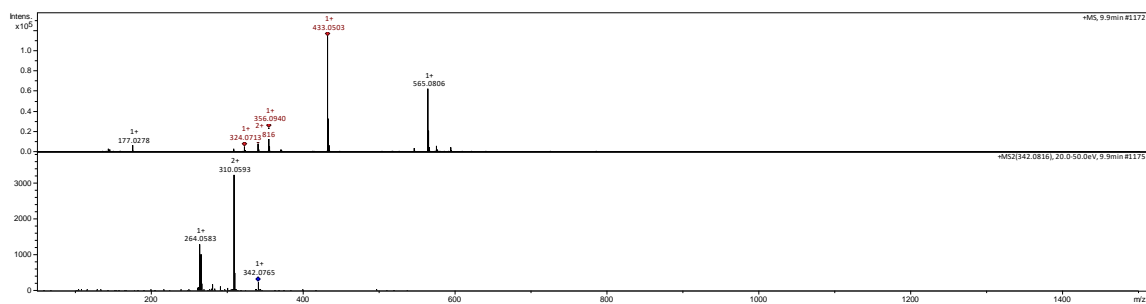


7,8,11-trimethoxy-2H,4H,5H,10H,11H-  
 indolo[7a,1-a]isoquinoline-4,5-dione  
 (10,11-Dioxoerysotrine)

B)



7,8,11-trimethoxy-4H,5H,10H,11H,12H-  
 indolo[7a,1-a]isoquinoline-5,12-dione

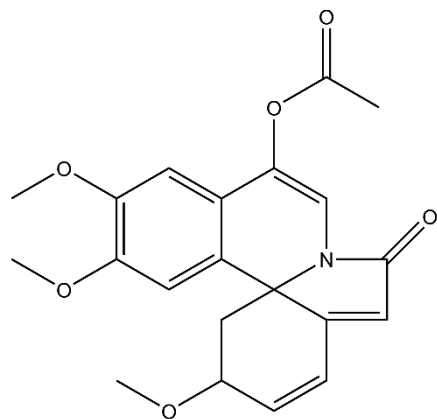


	Pool	342.0765	310.0593	264.0683
Fragments				
p7(a)		•	•	•

10

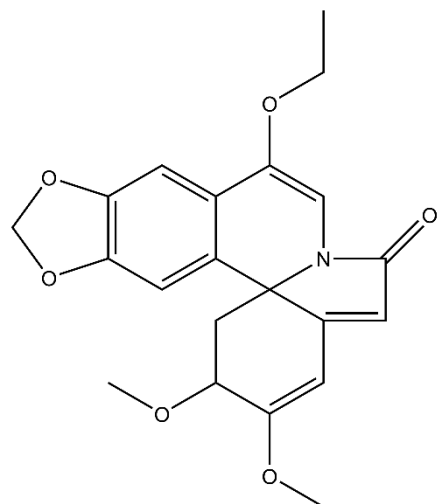
 RT - 13.7 min  
 $m/z$  384.1203  $[M+H]^+$ 

A)



19-methoxy-14-oxo-5,7-dioxa-13-azapentacyclo[11.7.0.0<sup>1</sup>,<sup>16</sup>.0<sup>2</sup>,<sup>10</sup>.0<sup>4</sup>,<sup>8</sup>]icosa-2,4(8),9,11,15,17-hexaen-11-yl acetate

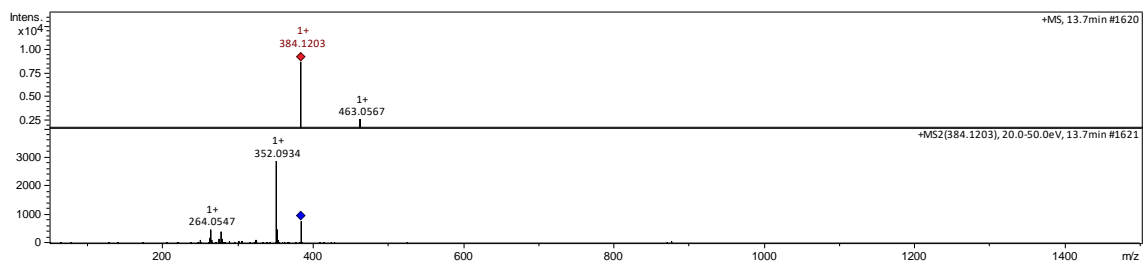
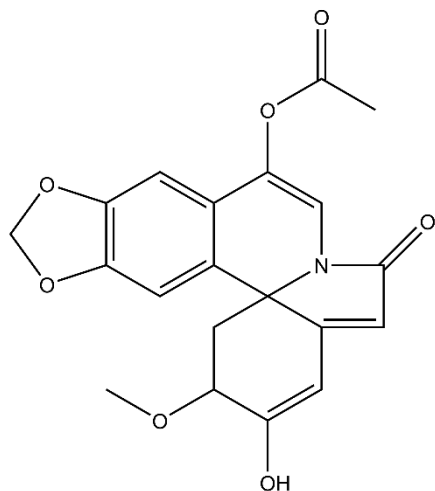
B)



11-ethoxy-18,19-dimethoxy-5,7-dioxa-13-azapentacyclo[11.7.0.0<sup>1</sup>,<sup>16</sup>.0<sup>2</sup>,<sup>10</sup>.0<sup>4</sup>,<sup>8</sup>]icosa-2,4(8),9,11,15,17-hexaen-14-one

C)

18-hydroxy-19-methoxy-14-oxo-5,7-dioxa-13-azapentacyclo[11.7.0.0<sup>1</sup>,<sup>16</sup>.0<sup>2</sup>,<sup>10</sup>.0<sup>4</sup>,<sup>8</sup>]icosa-2,4(8),9,11,15,17-hexaen-11-yl acetate



Pool	Fragments			
	384.1285	352.1007	278.0747	264.0569
p7(a)	•	•	•	•
p2(a)	•	•	•	•
p4(a)	•	•	•	•
p1(a)	•	•	•	•
p1(b)	•	•	•	•

---

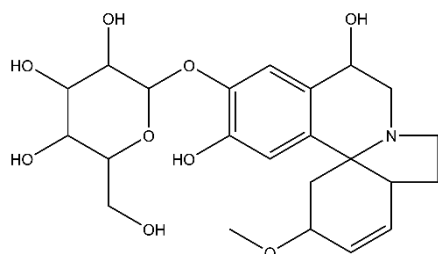
**N° Structure and chemical name (Cluster 2) and mass spectra (MS and MS/MS)**


---

**11**
 RT - 1.3 min  
 $m/z$  466.1401  $[M+H]^+$ 

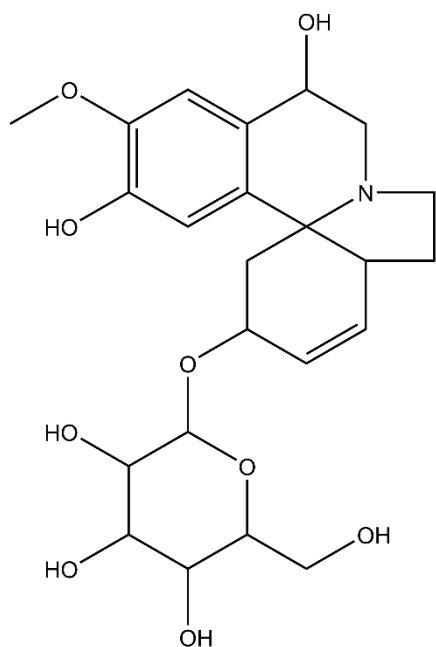

---

A)



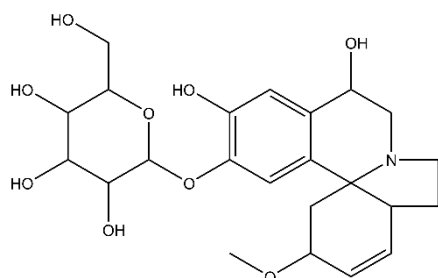
2-({5,8-dihydroxy-11-methoxy-1H,2H,4H,5H,10H,11H,13aH-indolo[7a,1-a]isoquinolin-7-yl}oxy)-6-(hydroxymethyl)oxane-3,4,5-triol

B)



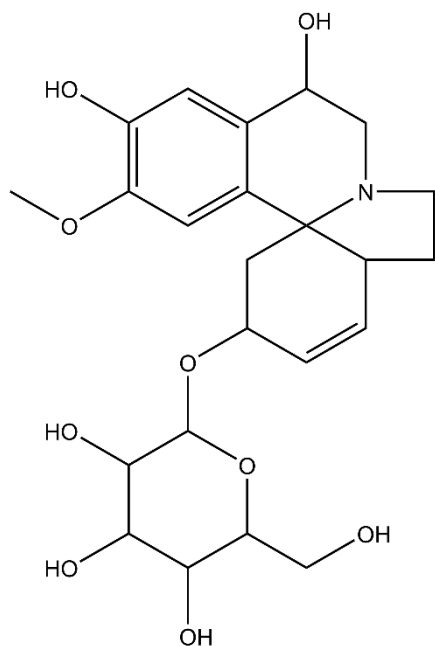
2-({5,8-dihydroxy-7-methoxy-1H,2H,4H,5H,10H,11H,13aH-indolo[7a,1-a]isoquinolin-11-yl}oxy)-6-(hydroxymethyl)oxane-3,4,5-triol

C)

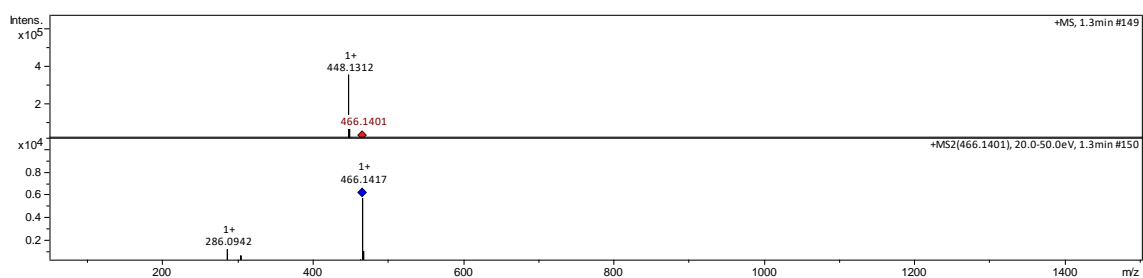


2-({5,7-dihydroxy-11-methoxy-1H,2H,4H,5H,10H,11H,13aH-indolo[7a,1-a]isoquinolin-8-yl}oxy)-6-(hydroxymethyl)oxane-3,4,5-triol

D)



2-({5,7-dihydroxy-8-methoxy-1H,2H,4H,5H,10H,11H,13aH-indolo[7a,1-a]isoquinolin-11-yl}oxy)-6-(hydroxymethyl)oxane-3,4,5-triol

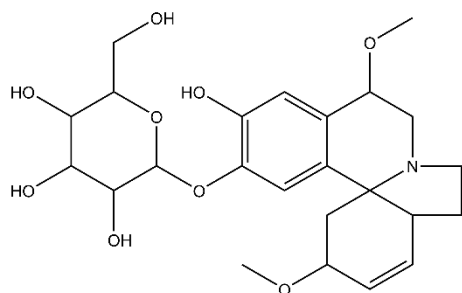


pool	Fragments		
	466.1381	304.104	286.0921
p3(b)	●	●	●
p3(c)	●	●	●
p2(a)	●	●	●
p2(b)	●	●	●
p2(c)	●	●	●
p4(a)	●	●	●
p4(b)	●	●	●
p4(c)	●	●	●
p1(a)	●	●	●

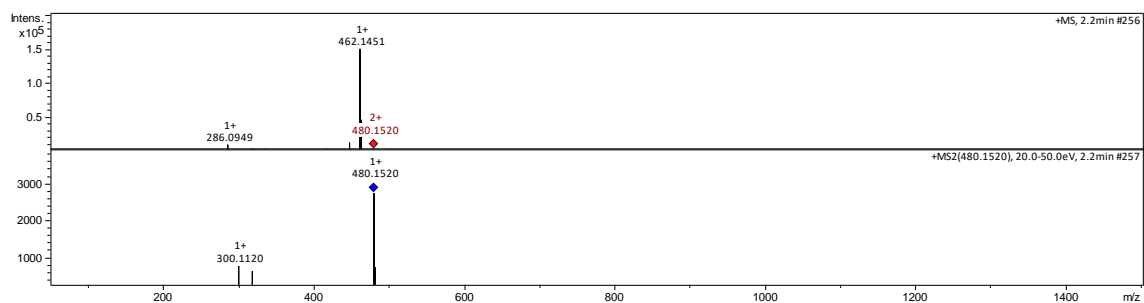
12

RT - 2.2 min  
 $m/z$  480.1520  $[M+H]^+$

A)



2-((7-hydroxy-5,11-dimethoxy-1H,2H,4H,5H,10H,11H,13aH-indolo[7<sup>a</sup>,1-a]isoquinolin-8-yl)oxy)-6-(hydroxymethyl)oxane-3,4,5-triol

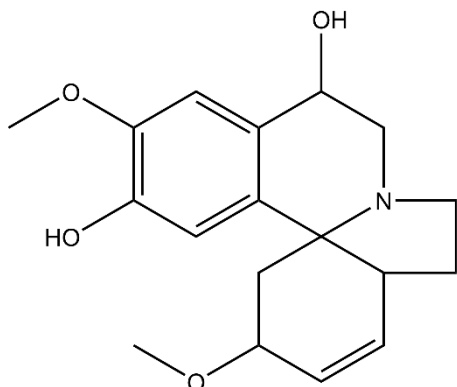


pool	Fragment	480.1492	318.1136	300.1065
	p3(a)		•	•
p3(b)		•	•	•
p3(c)		•	•	•

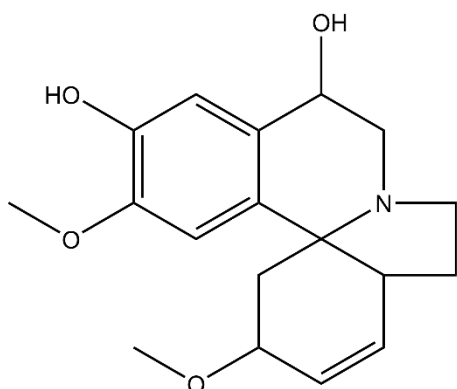
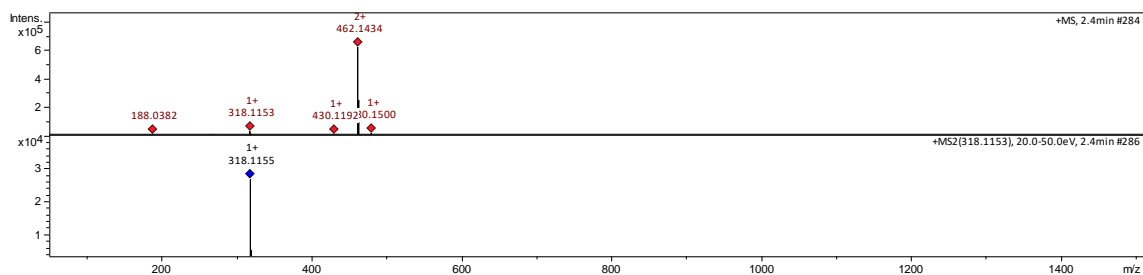
13

RT - 2.4 min  
 $m/z$  318.1153  $[M+H]^+$ 

A)

7,11-dimethoxy-  
1H,2H,4H,5H,10H,11H,13aH-indolo[7a,1-  
a]isoquinoline-5,8-diol

B)

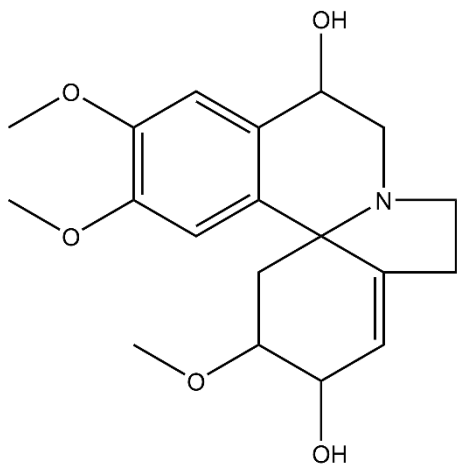
8,11-dimethoxy-  
1H,2H,4H,5H,10H,11H,13aH-indolo[7a,1-  
a]isoquinoline-5,7-diol

pool	Fragments	
	318.1161	300.1108
p7(a)	●	●
p8(c)	●	●
p3(a)	●	-
p3(b)	●	-

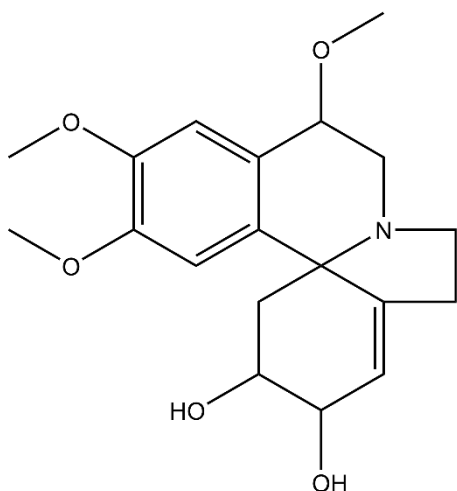
14

RT - 3.4 min  
 $m/z$  348.0348 [M+H]<sup>+</sup>

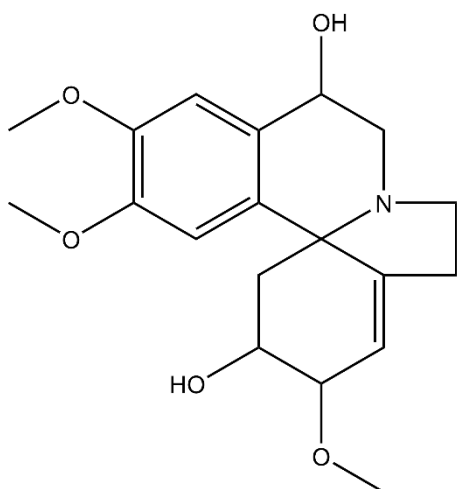
A)

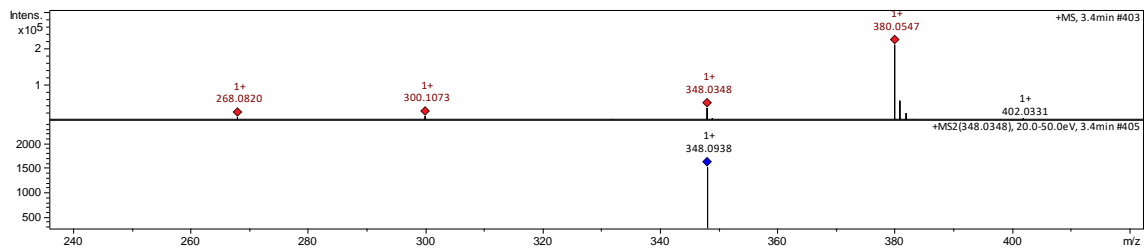
7,8,11-trimethoxy-  
1H,2H,4H,5H,10H,11H,12H-indolo[7a,1-  
a]isoquinoline-5,12-diol(11-Hydroxyerythratidine or 11-  
Hydroxyepierythratidine)

B)

5,7,8-trimethoxy-  
1H,2H,4H,5H,10H,11H,12H-indolo[7a,1-  
a]isoquinoline-11,12-diol

C)

7,8,12-trimethoxy-  
1H,2H,4H,5H,10H,11H,12H-indolo[7a,1-  
a]isoquinoline-5,11-diol

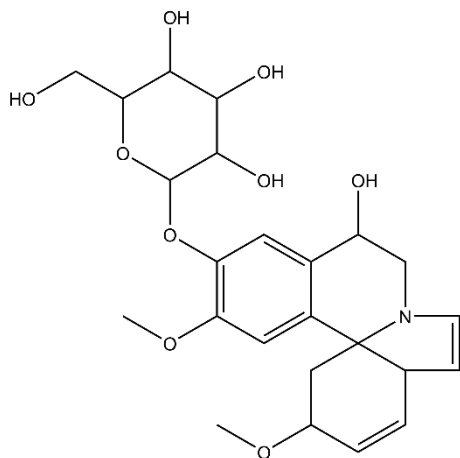


pool \ Fragments	348.0890	330.0759
p8(b)	●	●
p8(c)	●	-
p5(b)	●	●
p5(c)	●	-
p3(b)	●	-
p2(b)	●	●
p4(a)	●	-
p4(b)	●	-

15

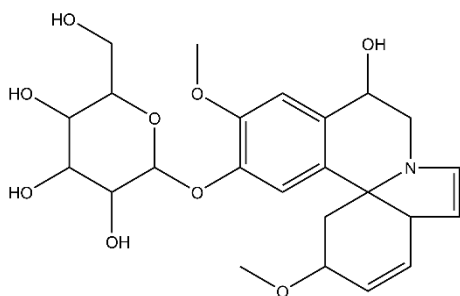
 RT - 3.3 min  
 $m/z$  478.1358  $[M+H]^+$ 

A)

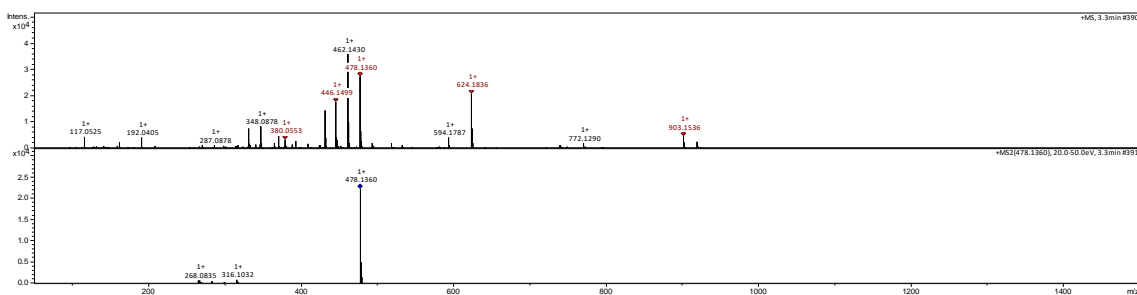


2-({5-hydroxy-8,11-dimethoxy-4H,5H,10H,11H,13aH-indolo[7a,1-a]isoquinolin-7-yl}oxy)-6-(hydroxymethyl)oxane-3,4,5-triol

B)



2-({5-hydroxy-7,11-dimethoxy-4H,5H,10H,11H,13aH-indolo[7a,1-a]isoquinolin-8-yl}oxy)-6-(hydroxymethyl)oxane-3,4,5-triol

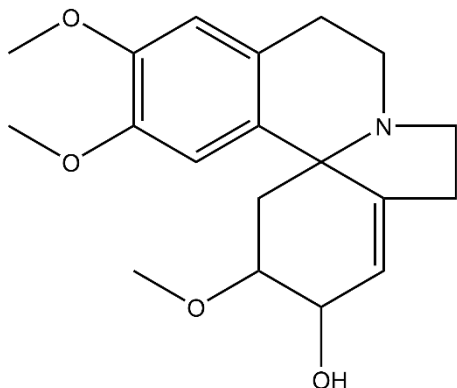


pool	Fragments	478.1324	316.0959	268.0777
	p3(a)	●	●	●
	p3(b)	●	-	●
	p3(c)	●	●	●
	p2(a)	●	●	-
	p2(b)	●	●	-
	p2(c)	●	●	●

16

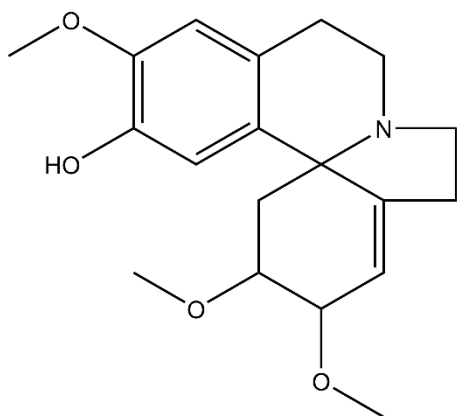
RT - 3.2 min  
m/z 332.0980 [M+H]<sup>+</sup>

A)



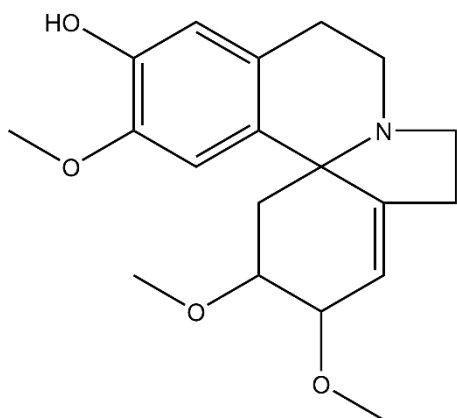
7,8,11-trimethoxy-  
1H,2H,4H,5H,10H,11H,12H-indolo[7<sup>a</sup>,1-  
a]isoquinolin-12-ol  
(Erythratidine)

B)

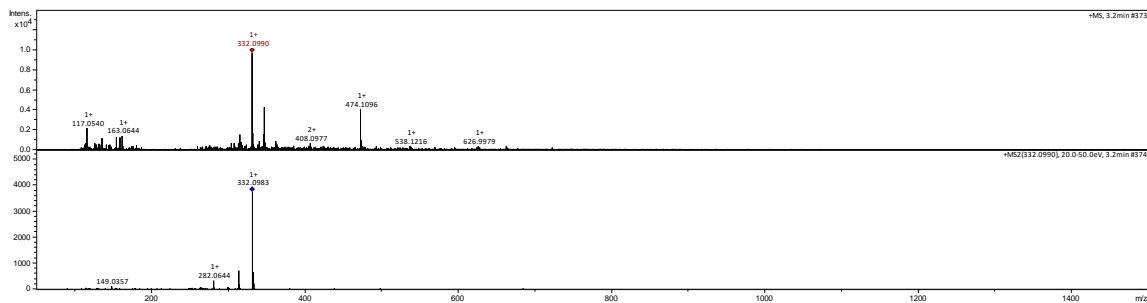


7,11,12-trimethoxy-  
1H,2H,4H,5H,10H,11H,12H-indolo[7<sup>a</sup>,1-  
a]isoquinolin-8-ol

C)



8,11,12-trimethoxy-  
1H,2H,4H,5H,10H,11H,12H-indolo[7<sup>a</sup>,1-  
a]isoquinolin-7-ol

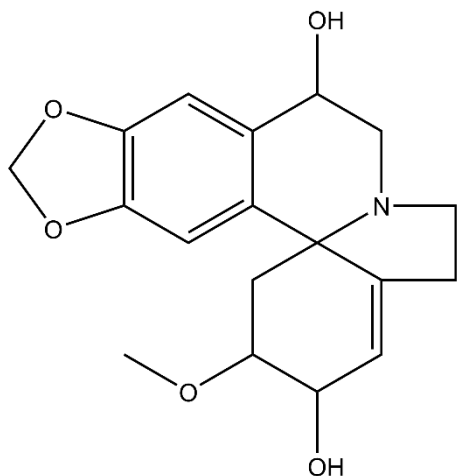


pool	Fragments	332.0985	314.0956	282.0757
	p7(a)	●	●	●
	p7(b)	●	●	●
	p6(b)	●	-	-
	p6(c)	●	-	-
	p8(b)	●	●	-
	p8(c)	●	●	●
	p1(c)	●	-	-

17

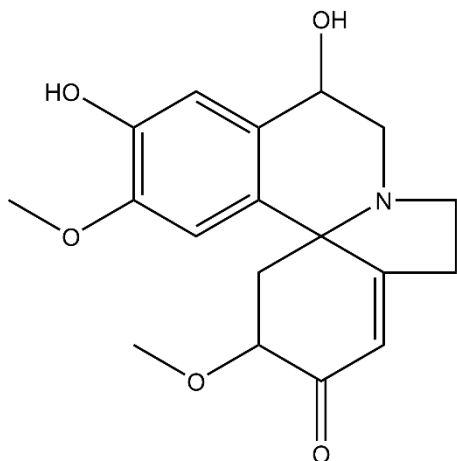
RT - 4.1 min  
m/z 332.0982 [M+H]<sup>+</sup>

A)

19-methoxy-5,7-dioxa-13-azapentacyclo[11.7.0.0<sup>1</sup>,<sup>16</sup>.0<sup>2</sup>,<sup>10</sup>.0<sup>4</sup>,<sup>8</sup>]icosa-2,4(8),9,16-tetraene-11,18-diol

(11-Hydroxyerythratine or 11-Hydroxyepierythratine)

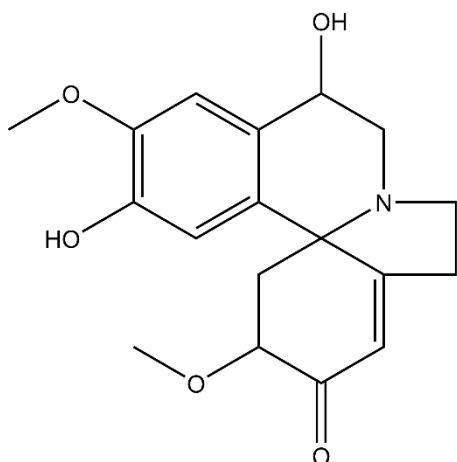
B)



5,7-dihydroxy-8,11-dimethoxy-1H,2H,4H,5H,10H,11H,12H-indolo[7a,1-a]isoquinolin-12-one

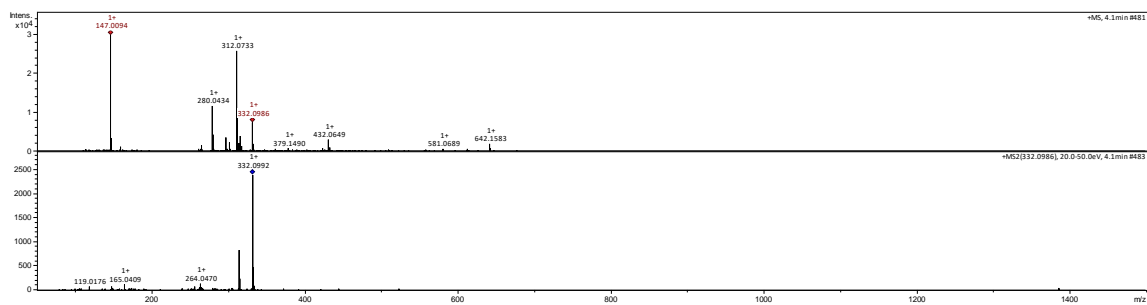
(11-Hydroxyerysotinine)

C)



5,8-dihydroxy-7,11-dimethoxy-1H,2H,4H,5H,10H,11H,12H-indolo[7a,1-a]isoquinolin-12-one

(11-Hydroxyerysosalvinone)

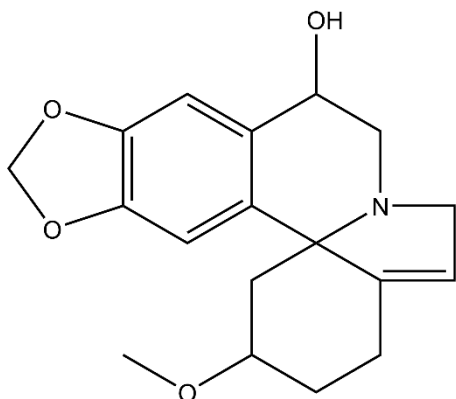


pool	Fragments	332.0971	314.0914	282.0621
	p7(a)	●	●	●
	p7(b)	●	●	●
	p7(c)	●	●	●

18

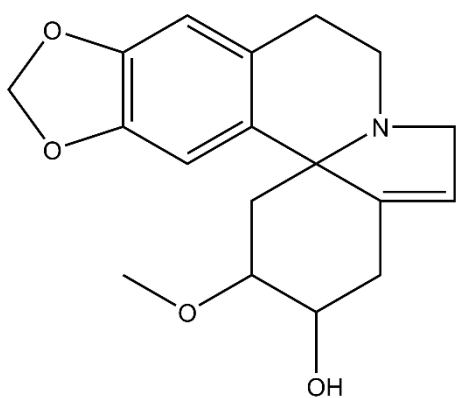
RT - 5.7 min  
 $m/z$  316.1044  $[M+H]^+$

A)



19-methoxy-5,7-dioxo-13-  
 azapentacyclo[11.7.0.0<sup>1</sup>,16.0<sup>2</sup>,10.0<sup>4</sup>,8]icosa-  
 2,4(8),9,15-tetraen-11-ol

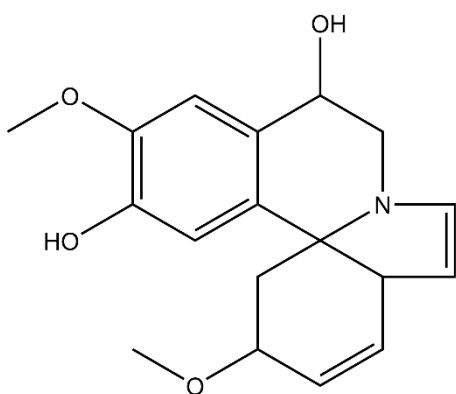
B)



19-methoxy-5,7-dioxo-13-  
 azapentacyclo[11.7.0.0<sup>1</sup>,16.0<sup>2</sup>,10.0<sup>4</sup>,8]icosa-  
 2,4(8),9,15-tetraen-18-ol

(Erythratine or Epierythratine)

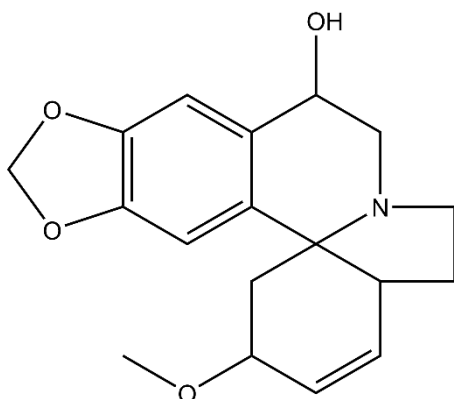
C)



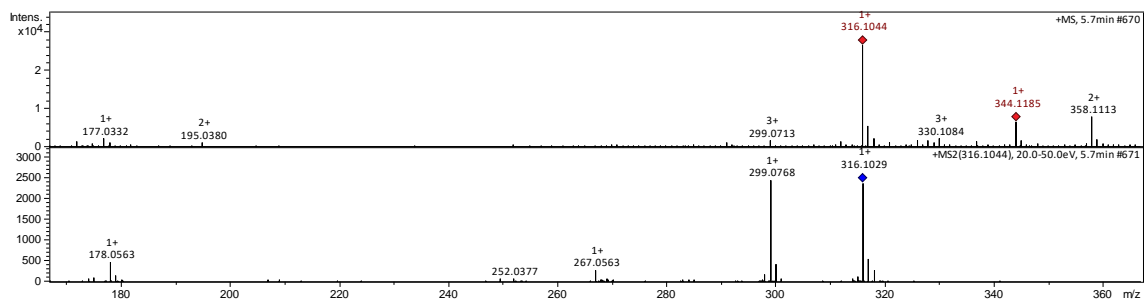
7,11-dimethoxy-4H,5H,10H,11H,13aH-  
 indolo[7<sup>a</sup>,1-a]isoquinoline-5,8-diol

(11-Hydroxyerysovin (alpha or beta))

D)



19-methoxy-5,7-dioxa-13-  
azapentacyclo[11.7.0.0<sup>1</sup>,16.0<sup>2</sup>,10.0<sup>4</sup>,8]icosa-  
2,4(8),9,17-tetraen-11-ol

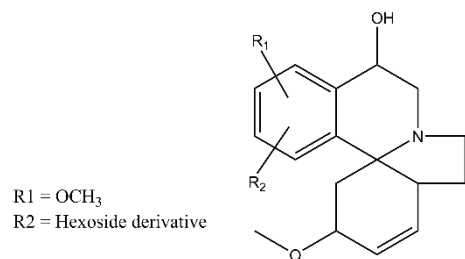


pool	Fragments	m/z				
		316.1048	299.0792	267.0565	178.0547	163.0290
p7(a)		●	●	●	-	●
p7(b)		●	●	●	●	-
p7(c)		●	●	●	●	●

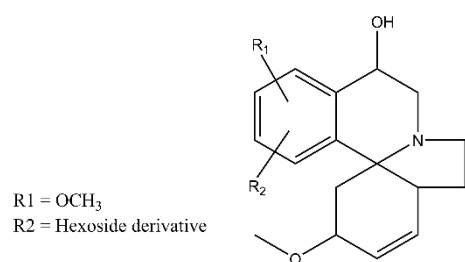
19

RT - 5.1 min  
 $m/z$  566.1426  $[M+H]^+$ 

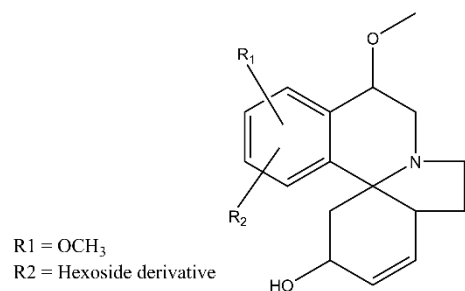
A)



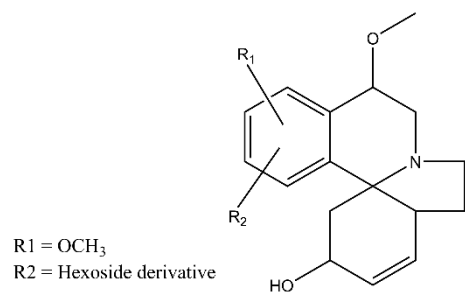
B)

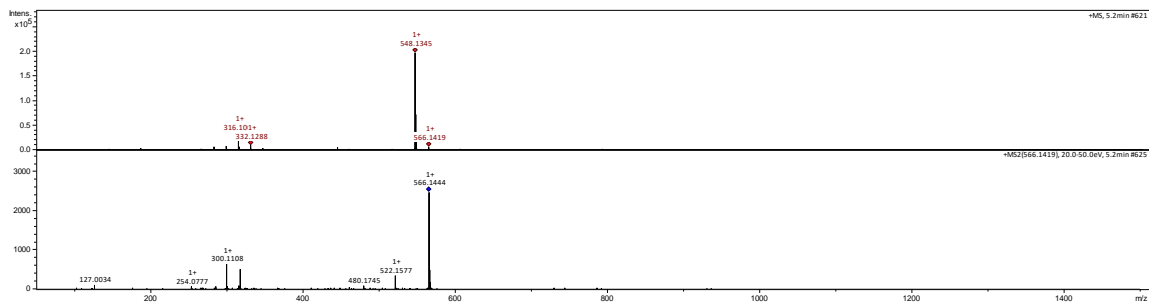


C)



D)



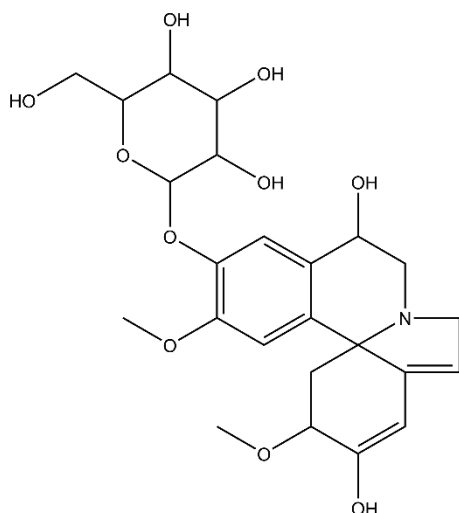


pool	Fragments	566.1417	522.1626	318.1154	300.1049
	p3(b)		●	●	●
p3(c)		●	●	●	-
p2(b)		●	●	●	●

20

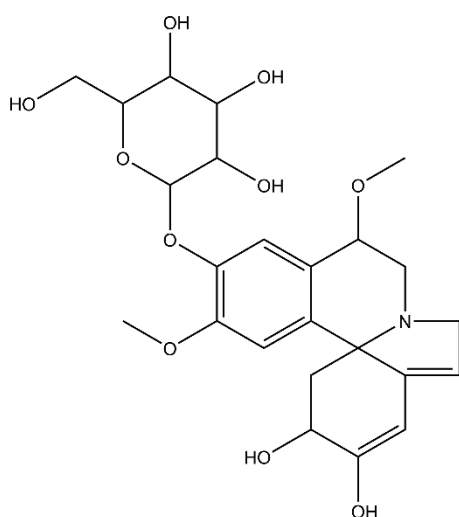
RT - 5.6 min  
 $m/z$  494.1265  $[M+H]^+$ 

A)



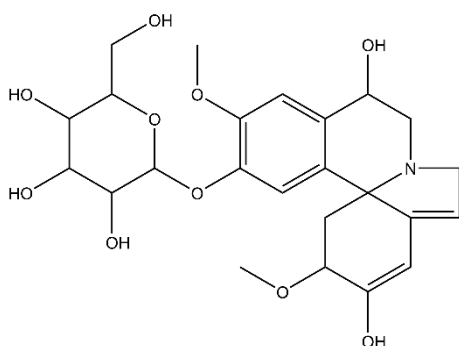
2-({5,12-dihydroxy-8,11-dimethoxy-2H,4H,5H,10H,11H-indolo[7a,1-a]isoquinolin-7-yl}oxy)-6-(hydroxymethyl)oxane-3,4,5-triol

B)

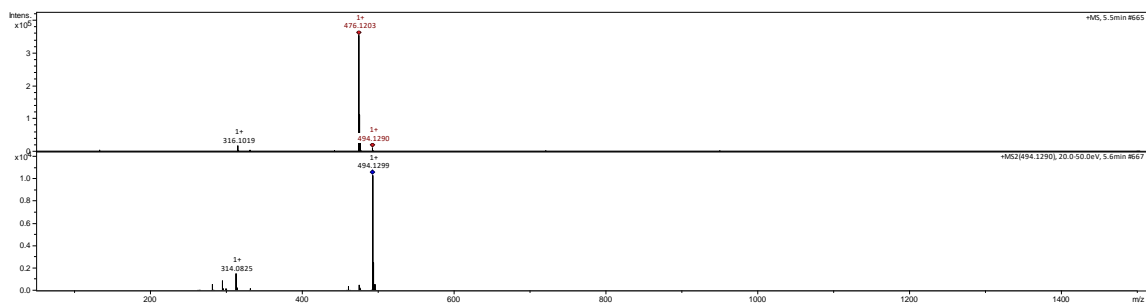


2-({11,12-dihydroxy-5,8-dimethoxy-2H,4H,5H,10H,11H-indolo[7a,1-a]isoquinolin-7-yl}oxy)-6-(hydroxymethyl)oxane-3,4,5-triol

C)



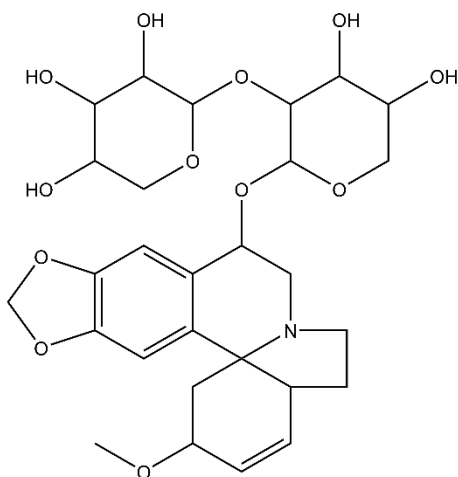
2-({5,12-dihydroxy-7,11-dimethoxy-2H,4H,5H,10H,11H-indolo[7a,1-a]isoquinolin-8-yl}oxy)-6-(hydroxymethyl)oxane-3,4,5-triol



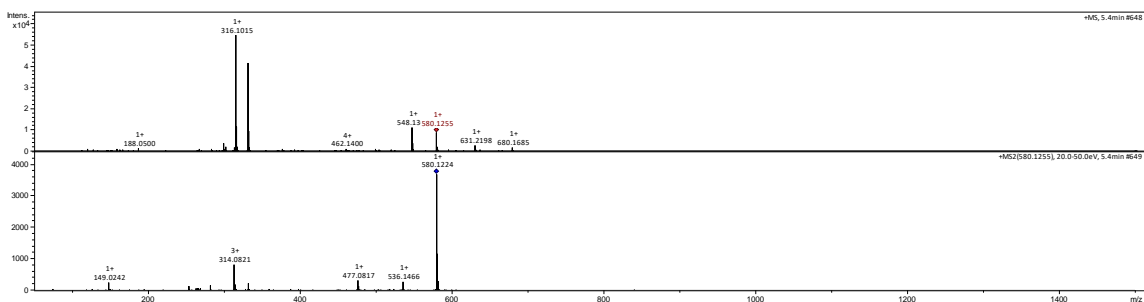
pool	Fragments	494.1261	476.1184	314.0807
	p3(a)	●	-	●
	p3(b)	●	-	●
	p3(c)	●	●	●
	p2(a)	●	-	-
	p2(b)	●	●	-
	p2(c)	●	-	●

21

RT - 5.4 min  
 $m/z$  580.1255  $[M+H]^+$



2-{{4,5-dihydroxy-2-({19-methoxy-5,7-dioxa-13-azapentacyclo[11.7.0.0<sup>1</sup>,<sup>16</sup>.0<sup>2</sup>,<sup>10</sup>.0<sup>4</sup>,<sup>8</sup>]icosa-2,4(8),9,17-tetraen-11-yl}oxy)oxan-3-yl}oxy}oxane-3,4,5-triol

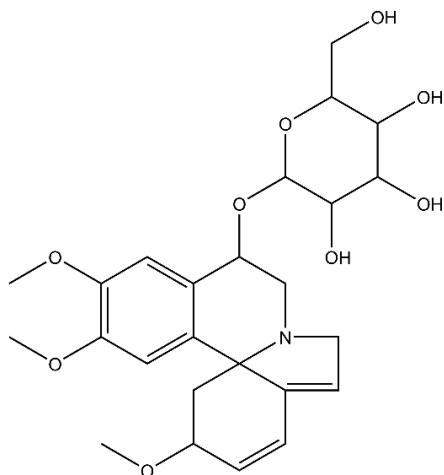


pool	Fragments	580.1201	536.1500	314.0841	149.0211
	p3(b)	●	●	●	●
	p3(c)	●	●	●	●
	p2(b)	●	●	●	●
	p2(c)	●	-	●	-
	p4(c)	●	●	●	●

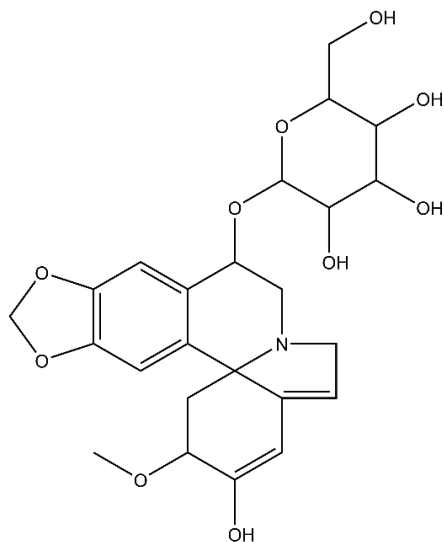
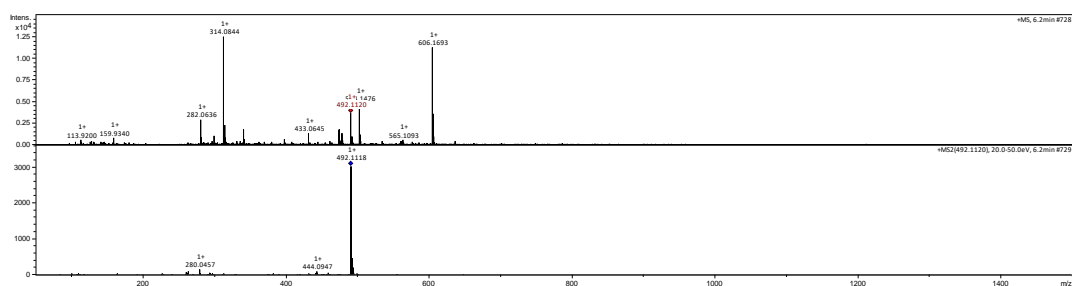
22

 RT - 6.2 min  
 $m/z$  492.1120  $[M+H]^+$ 

A)


 2-(hydroxymethyl)-6-({7,8,11-trimethoxy-  
 2H,4H,5H,10H,11H-indolo[7a,1-  
 a]isoquinolin-5-yl}oxy)oxane-3,4,5-triol

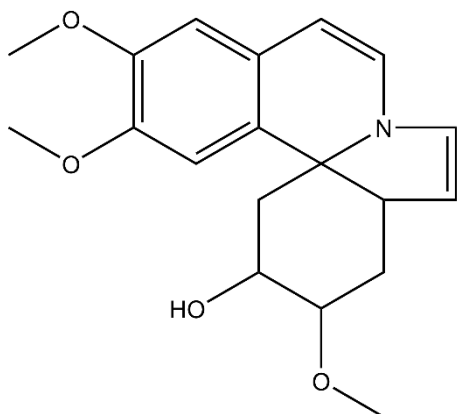
B)


 2-({18-hydroxy-19-methoxy-5,7-dioxo-13-  
 azapentacyclo[11.7.0.0<sup>1</sup>,16.0<sup>2</sup>,10.0<sup>4</sup>,8]  
 icosa-2,4(8),9,15,17-pentaen-11-yl}oxy)-  
 6-(hydroxymethyl)oxane-3,4,5-triol


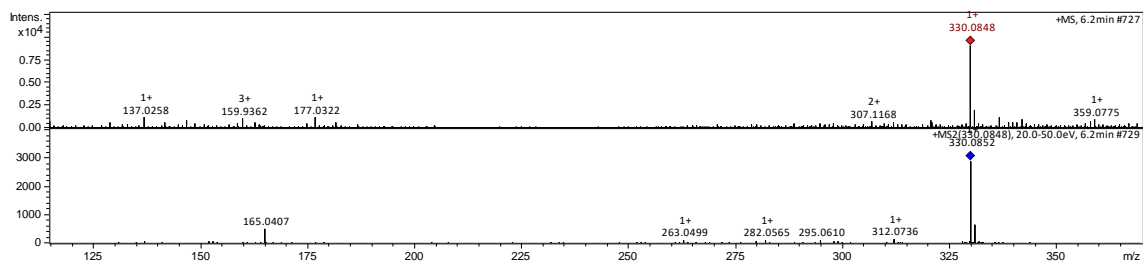
	Fragments	492.1118	280.0457	264.0554
pool				
	p3(a)	●	●	●
	p2(a)	●	●	-
	p2(b)	●	-	●
	p4(a)	●	●	-
	p1(a)	●	-	●

23

RT - 6.2 min  
 $m/z$  330.0852  $[M+H]^+$

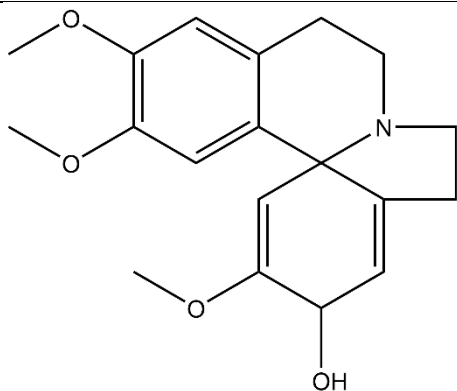
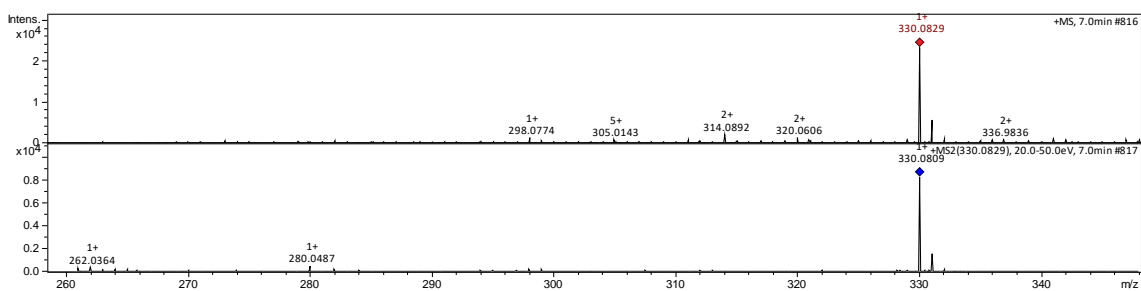


7,8,12-trimethoxy-10H,11H,12H,13H,13aH-  
 indolo[7a,1-a]isoquinolin-11-ol



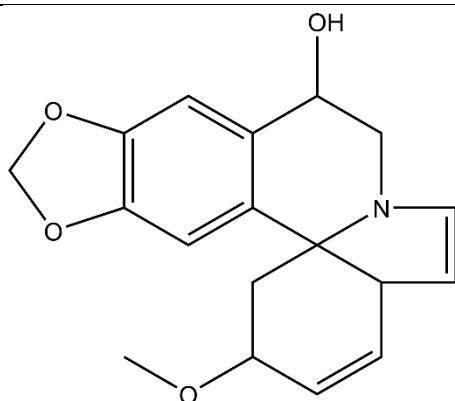
pool	Fragments	330.0843	312.0744	282.0565	165.0401
	p6(c)		•	•	•
p5(b)		•	•	•	•

24

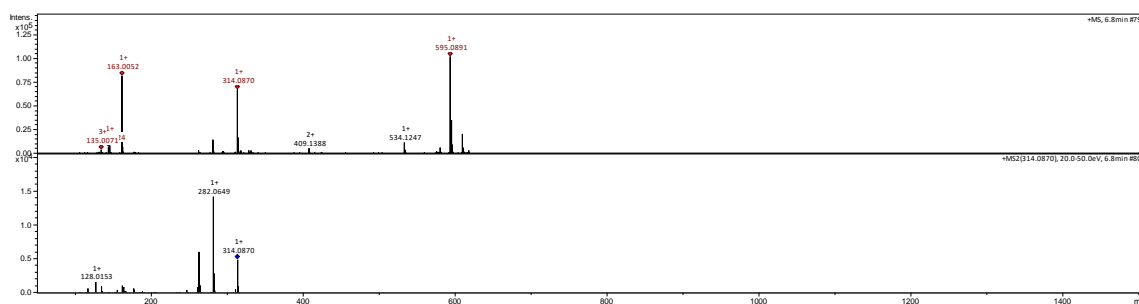
RT - 7.0 min  
 $m/z$  330.0829  $[M+H]^+$ 7,8,11-trimethoxy-1H,2H,4H,5H,12H-  
indolo[7a,1-a]isoquinolin-12-ol

pool \ Fragments	330.0843	262.0431	165.0401
p7(a)	●	●	●
p7(b)	●	●	●
p7(c)	●	●	●
p6(a)	●	-	-
p6(b)	●	-	●
p6(c)	●	-	●
p8(b)	●	-	●
p8(c)	●	-	●
p5(b)	●	●	●
p5(c)	●	-	●
p3(a)	●	-	●
p3(b)	●	-	●

25

RT - 6.8 min  
 $m/z$  314.0841  $[M+H]^+$ 19-methoxy-5,7-dioxa-13-  
azapentacyclo[11.7.0.0<sup>1</sup>,16.0<sup>2</sup>,10.0<sup>4</sup>,8]icosa-  
2,4(8),9,14,17-pentaen-11-ol

(Erythrine)



pool	Fragments	m/z				
		314.0894	282.0676	266,0754	264.0557	128.0103
	p7(a)	●	●	-	●	-
	p7(b)	●	●	-	●	-
	p7(c)	●	●	●	-	-
	p3(a)	●	●	●	●	●
	p3(b)	●	●	●	●	●
	p3(c)	●	●	●	●	●
	p2(a)	●	●	-	●	-

Os números (1-25) representam o número de alcalóides anotados e as letras (A, B, C e D) representam as possibilidades estruturais do mesmo alcalóide. A nomenclatura química está de acordo com a União Internacional de Química Pura e Aplicada (IUPAC). A massa/carga ( $m/z$ ), o tempo de retenção e o espectrograma (MS e MS/MS), junto com os principais fragmentos  $m/z$  de cada alcalóide em cada amostra também são mostrados na tabela; as *pools* 1-5 representa a amostra de sementes, enquanto que de 5-8 amostras de folhas.

**ANEXO II**

*Material Suplementar*

**Tabela S2.** Análise do perfil metabólito por *pool*.

<b>N° alkaloid</b>	<b>RT (min)</b>	<b><i>m/z</i></b>	<b>Estrutura</b>	<b><i>Pool</i></b>	<b>Abundância relativa</b>	<b>Altura do pico</b>
<b>1</b>	<b>3.8</b>	312.0715	Folha	7	<b>8</b>	2.81E+04
<b>2</b>	<b>4.4</b>	300.1101	Folha	6,7,8	<b>7</b>	2.70E+04
<b>3</b>	<b>5.1</b>	284.0787	Semente	1,2,3,4	<b>2</b>	6.23E+03
<b>4</b>	<b>5.5</b>	358.1098	Folha	6,7	<b>4</b>	1.60E+04
<b>5</b>	<b>3.8</b>	312.0738	Folha	6,7	<b>9</b>	3.33E+04
<b>6</b>	<b>6.3</b>	296.0805	Folha/semente	2,7	<b>1</b>	4.39E+03
<b>7</b>	<b>6.7</b>	314.0894	Folha	7	<b>92</b>	3.41E+05
<b>8</b>	<b>8.2</b>	298.0966	Folha/semente	3,5,7,8	<b>48</b>	1.75E+05
<b>9</b>	<b>9.9</b>	342.0816	Folha	7	<b>12</b>	4.26E+04
<b>10</b>	<b>13.7</b>	384.1203	Folha/semente	1,2,4,7	<b>2</b>	8.55E+03
<b>11</b>	<b>1.3</b>	466.1401	Semente	1,2,3,4	<b>2</b>	7.83E+03
<b>12</b>	<b>2.2</b>	480.1520	Semente	3	<b>6</b>	2.36E+04
<b>13</b>	<b>2.4</b>	318.1156	Folha/semente	3,7,8	<b>2</b>	6.73E+03
<b>14</b>	<b>3.4</b>	348.0348	Folha/semente	2,3,4,5,8	<b>9</b>	3.40E+04
<b>15</b>	<b>3.3</b>	478.1358	Semente	2,3	<b>8</b>	2.94E+04
<b>16</b>	<b>3.2</b>	332.0980	Folha/semente	1,6,7,8	<b>4</b>	1.65E+04
<b>17</b>	<b>4.1</b>	332.0982	Folha	7	<b>2</b>	8.84E+03
<b>18</b>	<b>5.7</b>	316.1044	Folha	7	<b>2</b>	7.84E+03
<b>19</b>	<b>5.1</b>	566.1426	Semente	2,3	<b>2</b>	5.63E+03
<b>20</b>	<b>5.6</b>	494.1265	Semente	2,3	<b>2</b>	5.54E+03
<b>21</b>	<b>5.4</b>	580.1255	Semente	2,3,4	<b>1</b>	4.94E+03
<b>22</b>	<b>6.2</b>	492.1120	Semente	1,2,3,4	<b>1</b>	3.69E+03
<b>23</b>	<b>6.2</b>	330.0781	Folha	5,6,	<b>8</b>	2.87E+04
<b>24</b>	<b>7.0</b>	330.0829	Folha/semente	3,5,6,7,8	<b>8</b>	2.87E+04
<b>25</b>	<b>6.8</b>	314.0841	Folha/semente	2,3,7	<b>7</b>	2.68E+04

**ANEXO III**

*Material Suplementar*

Tabela S3. Avaliação das bibliotecas do transcriptoma de *Erythrina velutina*.

Sample	Reads	Região vegetal	N° leituras brutas	N° rRNA	Porcentagem de rRNA reads	N° reads sem rRNA	Porcentagem de reads sem rRNA	N° final de reads limpas*	Porcentagem final de reads limpas	Nucleotídeo final após limpeza
<b>Pool 1</b>	<i>Forward</i>	Semente	38,517,758	2,409,427	6,26%	36,108,331	93,74%	27,465,395	71,31%	2,817,968,953
	<i>Reverse</i>	Semente	38,517,758	2,409,427	6,26%	36,108,331	93,74%	27,465,395	71,31%	2,674,741,368
<b>Pool2</b>	<i>Forward</i>	Semente	45,536,755	1,844,263	4,05%	43,692,492	95,95%	34,199,767	75,10%	3,516,006,145
	<i>Reverse</i>	Semente	45,536,755	1,844,263	4,05%	43,692,492	95,95%	34,199,767	75,10%	3,360,532,309
<b>Pool 3</b>	<i>Forward</i>	Semente	34,385,125	6,787,625	19,74%	27,597,500	80,26%	20,856,497	60,66%	2,143,877,575
	<i>Reverse</i>	Semente	34,385,125	6,787,625	19,74%	27,597,500	80,26%	20,856,497	60,66%	2,056,701,006
<b>Pool 4</b>	<i>Forward</i>	Semente	50,105,977	6,158,800	12,29%	43,947,177	87,71%	31,730,471	63,33%	3,211,592,730
	<i>Reverse</i>	Semente	50,105,977	6,158,800	12,29%	43,947,177	87,71%	31,730,471	63,33%	3,051,795,444
<b>Pool 5</b>	<i>Forward</i>	Folhas	42,873,185	2,620,503	6,11%	40,252,682	93,89%	30,967,679	72,23%	3,231,814,195
	<i>Reverse</i>	Folhas	42,873,185	2,620,503	6,11%	40,252,682	93,89%	30,967,679	72,23%	3,050,230,791
<b>Pool 6</b>	<i>Forward</i>	Folhas	43,766,448	5,643,469	12,89%	38,122,979	87,11%	28,741,073	65,67%	3,009,166,305
	<i>Reverse</i>	Folhas	43,766,448	5,643,469	12,89%	38,122,979	87,11%	28,741,073	65,67%	2,840,876,602
<b>Pool 7</b>	<i>Forward</i>	Folhas	38,583,198	3,685,525	9,55%	34,897,673	90,45%	26,473,356	68,61%	2,741,329,009
	<i>Reverse</i>	Folhas	38,583,198	3,685,525	9,55%	34,897,673	90,45%	26,473,356	68,61%	2,598,644,713
<b>Pool 8</b>	<i>Forward</i>	Folhas	57,391,786	2,941,368	5,13%	54,450,418	94,87%	42,558,990	74,16%	4,394,728,267
	<i>Reverse</i>	Folhas	57,391,786	2,941,368	5,13%	54,450,418	94,87%	42,558,990	74,16%	4,161,097,793

**ANEXO IV***Material Suplementar*

**Tabela S4.** Visão geral da análise e montagem do transcriptoma. Cada piscina corresponde a um tipo de estrutura (sementes ou folhas) localizada em uma área de colheita específica. Na área A (*pool* 1 - sementes / *pool* 5 - Folhas); B (*Pool* 2 - sementes / *Pool* 6- folhas); C (*Pool* 3 - sementes / *Pool* 7 - folhas); D (*pools* 4 - sementes / *pool* 8 - folhas). A e B correspondem aos municípios de Acari - RN e C e D aos municípios de Jardim do Seridó - RN. \* Leituras de alta qualidade: são leituras processadas após a remoção de RNAr de pares de bases com 20 ou mais índices de qualidade, e os valores médios de *Phred* acima de 33.

Localização	A Pool 1	B Pool 2	C Pool 3	D Pool 4	A Pool 5	B Pool 6	C Pool 7	D Pool 8
<b>Sequences</b>								
Número <i>reads</i>	38,517,758	45,536,735	34,385,125	50,105,977	42,873,185	43,766,448	38,583,198	57,391,786
<i>Reads</i> de alta qualidade*	27,465,395	34,199,767	20,856,947	31,730,471	30,967,679	28,741,073	26,473,356	42,558,990
Alcance de leitura (pb)	21-151	21-151	21-151	21-151	21-151	21-151	21-151	21-151
Comprimento médio (pb)	100	101	101	99	101	102	101	101
<b>Contigs</b>								
Número de <i>contigs</i>	71,747	91,072	68,050	68,804	88,190	93,207	87,819	95,089
Comprimento médio (pb)	1088.18	1102,95	984,88	871,50	1101.2	1096,8	1036,68	1085,99
N50	1606	1700	1535	1370	1683	1697	1596	1648

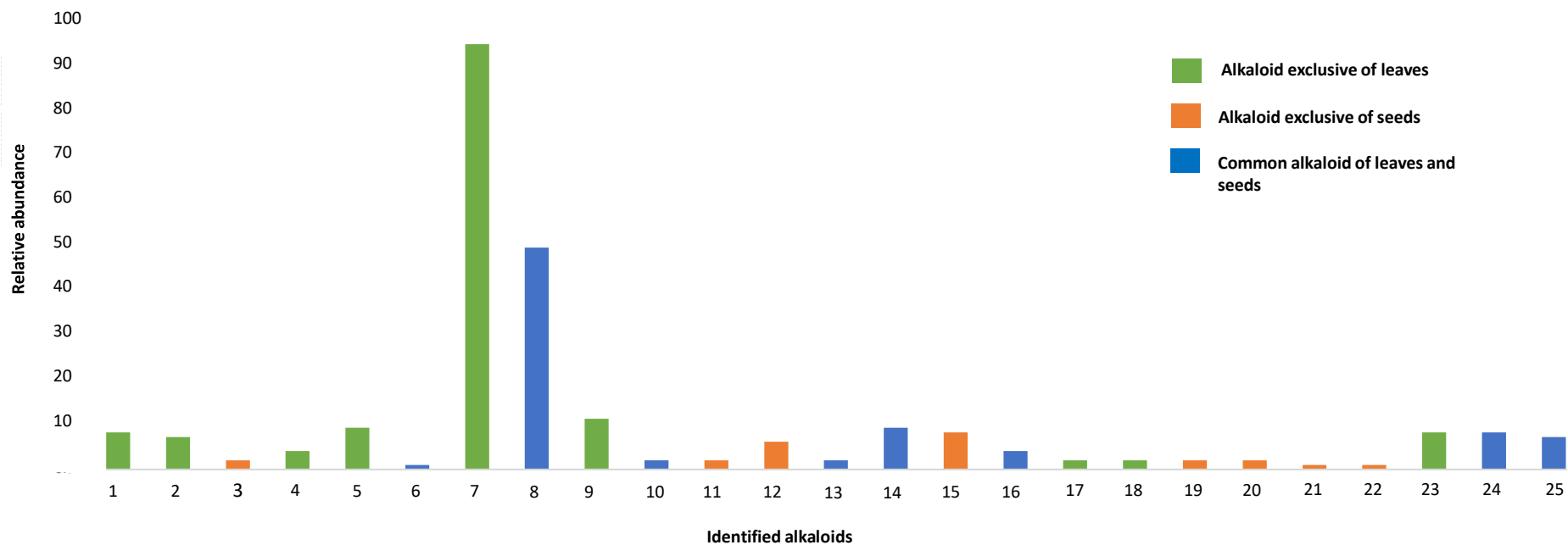
**ANEXO V***Material Suplementar*

**Tabela S5.** Resultados da análise *Bowtie2* de transcriptoma de sementes e folhas de *Erythrina velutina*.

<b>Amostras usadas para montagem</b>	<b>Região vegetal</b>	<b>Taxa de alinhamento geral</b>
Pool 1	Semente	97.90%
Pool 2	Semente	98.05%
Pool 3	Semente	97.86%
Pool 4	Semente	96.84%
Pool 5	Folhas	97.99%
Pool 6	Folhas	97.84%
Pool 7	Folhas	97.88%
Pool 8	Folhas	97.80%
Pool 1234 combinado	Sementes	97.75%
Pool 5678 combinado	Folhas	98.32%
Pool 1-8 combinado	Ambas as estruturas	98.77%

**ANEXO VI***Material Suplementar*

**Figura S6.** Análise semiquantitativa dos alcalóides da Eritrina identificados. Eixo Y: indica a proporção de abundância relativa entre entidades químicas; Eixo X: alcalóides identificados e anotados (1-25). As barras verdes e laranja representam alcalóides únicos nas folhas e sementes, respectivamente, enquanto as barras azuis representam alcalóides presentes nas folhas e sementes.



**ANEXO VII***Material Suplementar*

**Tabela S7.** Massa ( $m/z$ ), estrutura da planta, fórmula molecular, nível de anotação em nosso estudo e referência bibliográfica para os 25 alcalóides anotados

<b>Nº Alkaloids</b>	<b>Plant structure</b>	<b>Molecular formula</b>	<b><math>m/z</math> [M+H]</b>	<b>Annotation</b>	<b>Reference</b>
Alkaloid 1	Folhas	$C_{18}H_{17}NO_4$	312.0715	L2 or L3	u/n
Alkaloid 2	Folhas	(A) $C_{17}H_{17}NO_4$ (B) $C_{17}H_{17}NO_4$ (C) $C_{18}H_{21}NO_3$	300.1101	L2 or L3	(A) [9] (C) [12]
Alkaloid 3	Sementes	$C_{17}H_{17}NO_3$	284.0787	L2 or L3	u/n
Alkaloid 4	Folhas	(A) $C_{19}H_{19}NO_6$ (B,C) $C_{20}H_{23}NO_5$	358.1098	L2 or L3	u/n
Alkaloid 5	Folhas	(A,B,C) $C_{18}H_{17}NO_4$	312.0738	L2 or L3	(C) [2] [9],[11],[13] (B) [9]
Alkaloid 6	Folhas e sementes	$C_{18}H_{17}NO_3$	296.0805	L2 or L3	u/n
Alkaloid 7	Folhas	$C_{18}H_{19}NO_4$	314.0894	L2 or L3	(A,C) [9]
Alkaloid 8	Folhas e sementes	$C_{18}H_{19}NO_3$	298.0966	L2 or L3	(A) [9], [10]
Alkaloid 9	Folhas	$C_{19}H_{19}NO_5$	342.0816	L2 or L3	(A) [4],[7]

Alkaloid 10	Folhas e sementes	(A,B) $C_{21}H_{21}NO_6$ (C) $C_{20}H_{17}NO_7$	384.1203	L2 or L3	u/n
Alkaloid 11	Sementes	(A,B,C,D) $C_{23}H_{31}NO_9$	466.1401	L2 or L3	u/n
Alkaloid 12	Sementes	(A) $C_{24}H_{33}NO_9$	480.1520	L2 or L3	u/n
Alkaloid 13	Folhas e sementes	(A,B) $C_{18}H_{23}NO_4$	318.1153	L2 or L3	u/n
Alkaloid 14	Folhas e sementes	(A,B,C) $C_{19}H_{25}NO_5$	348.0348	L2 or L3	(A) [9]
Alkaloid 15	Sementes	(A,B) $C_{24}H_{31}NO_9$	478.1358	L2 or L3	u/n
Alkaloid 16	Folhas e sementes	(A,B,C) $C_{19}H_{25}NO_4$	332.0980	L2 or L3	(A) [2],[7],[9],[11],[8]
Alkaloid 17	Folhas	(A,B,C) $C_{18}H_{21}NO_5$	332.0982	L2 or L3	(A,B) [9]
Alkaloid 18	Folhas	(A,B,C,D) $C_{18}H_{21}NO_4$	316.1044	L2 or L3	(B) [1], [5], [6],[13] (C) [9]
Alkaloid 19	Sementes	-	566.1426	L2 or L3	u/n

Alkaloid 20	Sementes	(A,B,C) $C_{24}H_{31}NO_{10}$	494.1265	L2 or L3	u/n
Alkaloid 21	Sementes	$C_{28}H_{37}NO_{12}$	580.1255	L2 or L3	u/n
Alkaloid 22	Sementes	(A) $C_{25}H_{33}NO_9$ (B) $C_{24}H_{29}NO_{10}$	492.1120	L2 or L3	u/n
Alkaloid 23	Folhas	$C_{19}H_{23}NO_4$	330.0781	L2 or L3	u/n
Alkaloid 24	Folhas e sementes	$C_{19}H_{23}NO_4$	330.0829	L2 or L3	u/n
Alkaloid 25	Folhas e sementes	$C_{18}H_{19}NO_4$	314.0841	L2 or L3	[3],[11],[13]

As letras (A, B, C e D) representam as possibilidades estruturais do mesmo alcalóide de acordo com a tabela S1



**Reference:**

- [1] M.E. Amer, E. Shamma, A.J. Freyer, The Tetracyclic Erythrina Alkaloids. *Journal of Natural Products* (1991) 54, 329–363. doi:10.1021/np50074a001
- [2] Dictionary of Natural Products 29.2 (2020). Available from: <http://dnp.chemnetbase.com>.
- [3] D. S. Bhakuni, and s. Jain, Alkaloids of *cocculus laurifolius* D.C. *Tetrahedron* (1980) 36, 3107–3114. doi:10.1016/0040-4020(80)88040-9
- [4] A. S. Chawla, S. Chunchatprasert, and A.H. Jackson, Studies of Erythrina alkaloids: VII—<sup>13</sup>C NMR spectral studies of some Erythrina alkaloids. *Organic Magnetic Resonance*, (1983) 21, 39–41. doi:10.1002/omr.1270210110
- [5] H.-E. Zhao, Wu. J, Xu. F.-Q at al. Alkaloids from flowers of *Erythrina corallodendron*. *Natural Product Research* (2018), 1–6. doi:10.1080/14786419.2018.1472596
- [6] J. Wu, B-J. Zhang, Bao. M.-F, and X-H. Cai. A new erythrinan N-oxide alkaloid from *Erythrina stricta*. *Natural Product Research*, (2018)1 7. doi:10.1080/14786419.2018.1483924
- [7] C.C. Wanjala, B.F. Juma, G. Bojase, B.A. Gashe, R.R Majinda. Erythraline Alkaloids and Antimicrobial Flavonoids from *Erythrina latissima*. *Planta Medica*, 2002 68, 640–642. doi:10.1055/s-2002-32891
- [8] R. García-Mateos, M. Soto-Hernández, M. Martínez-Vázquez, A. Villegas-Monter. Isolation of alkaloids of *Erythrina* from tissue culture. *Phytochemical Analysis*, (1999) 10, 12–16. doi:10.1002/(sici)1099-1565(199901/02)10:1<12::aid-pca425>3.0.co;2-a
- [9] P. de A. Amaral, A.R. Antunes, J.W. Barlow. Isolation of erythrinan alkaloids from the leaves and flowers of *Erythrina speciosa*. *Revista Brasileira de Farmacognosia* (2019). doi:10.1016/j.bjrp.2019.01.007
- [10] K. Ito, H. Furukawa, and H. Tanaka. Structure of erythrinine, a new alkaloid from *Erythrina indica* Lam. *Journal of the Chemical Society D: Chemical Communications*, (1970). 17., 1076. doi:10.1039/c29700001076
- [11] J. Wu, B.-J. Zhang, M.-F. Bao, X.-H. Cai. A new erythrinan N-oxide alkaloid from

*Erythrina stricta*. Natural Product Research, (2018). 1–7.

**Tabela S8.** 25 alcaloides identificados e anotados e suas perdas características presentes no

Alcaloides anotados				
Cluster 1	Perdas características			
1	32u	50u	108u	-
2	32u	47u	-	-
3	18u	-	-	-
4	32u	91u	-	-
5	32u	50u	-	-
6	32u	-	-	-
7	32u	50u	149u	186u
8	32u	-	-	-
9	32u	78u	-	-
10	32u	106u	120u	-
Cluster 2	Perdas características			
11	162u	180u	-	-
12	162u	180u	-	-
13	18u	-	-	-
14	18u	-	-	-
15	162u	210u	-	-
16	18u	50u	-	-
17	18u	50u	-	-
18	18u	49u	138u	153u
19	44u	248u	266u	-
20	18u	180u	-	-
21	44u	266u	431u	-
22	212u	228u	-	-
23	18u	48u	165u	-
24	68u	165u	-	-
25	32u	48u	50u	186u

cluster 1 e 2