Erythrina velutina Willd. alkaloids: Piecing biosynthesis together from transcriptome analysis and metabolite profiling of seeds and leaves

Daisy Sotero Chacon, Taffarel Melo Torres, Ivanice Bezerra da Silva, Thiago Ferreira de Araújo, Alan de Araújo Roque, Francisco Ayrton Senna Domingos Pinheiro, Denise Selegato, Alan Pilon, Fernanda Priscila Santos Reginaldo, Cibele Tesser da Costa, Johnatan Vilasboa, Rafael Teixeira Freire, Eduardo Luiz Voigt, José Angelo Silveira Zuanazzi, Renata Libonati, Julia Abrantes Rodrigues, Filippe Lemos Maia Santos, Kátia Castanho Scortecci, Norberto Peporine Lopes, Leandro De Santis Ferreira, Leandro Vieira dos Santos, Alberto José Cavalheiro, Arthur Germano Fett-Neto, Raquel Brandt Giordani

Graphical Abstract

Harvest location

Plant structure

Syllaene alkaloids

Biocatalysis

Dew sample

Temperature

Relative humidity

Precipitation

Solar radiation

Active fire

Target molecule

RNA-seq

Protein

Metabolic profile

HPLC-MS

Peer review under responsibility of Cairo University.

* Corresponding author at: Gal. Gustavo Cordeiro de Farias Street, s/n, 59012-570 Natal, RN, Brazil.
E-mail address: raquebg@hotmail.com (R.B. Giordani).

https://doi.org/10.1016/j.jare.2021.01.017
2090-1232/© 2021 The Authors. Published by Elsevier B.V. on behalf of Cairo University.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Introduction

Plants are sessile organisms capable of producing over 200,000 specialized metabolites (also known as natural products of secondary metabolites) which represent chemical strategies for survival under environmental stresses and for interaction with other organisms [1]. Chemical diversity afforded by the specialized metabolism is a major source of bioactive molecules from plants, e.g., which have been used in folk medicine since ancient times, and it resonates with target molecules of biotechnological interest [2]. Specialized plant metabolites themselves or structures derived therefrom have been essential for medicinal chemistry and drug discovery [3].

The use of an integrative approach to understand a species in the context of its environment can accelerate and improve the identification of new chemical entities. During the course of evolution, plants have gained the ability to produce chemical compounds that are active against several threats, including biotic and abiotic stresses. This process has been driven by selective pressures inherent to the wide array of plant habitats, thereby contributing to the diversity of plant specialized metabolites. This is particularly true for regions characterized by extreme environments, such as the Caatinga, a Brazilian unique and understudied biome with semiarid edaphoclimatic features. Plants of this biome developed many metabolic adaptations to survive and resist in the face of several stress factors. The specialized metabolism fine-tunes plant fitness to habitat, requiring precise control of metabolic flux through specific pathways derived from central metabolism, and contributing to overall ecosystem functional dynamics. Responses to abiotic stresses encompass complex metabolite regulatory networks which can only be dissected using integrated omics analyses [4]. This scenario leads to the hypothesis that plants of the Caatinga constitute an important reservoir of novel bioactive molecules with an array of potential applications.

Bioanalytical pathways and metabolic regulation at different levels are complex issues that have been the subject of investigation mainly in model plants. Although fundamental, these studies fall short of providing a more representative picture of the variability and molecular wealth of all plants [5]. Targeted genome or transcriptome sequencing of medicinal plant species could pave the way for synthetic biology approaches aiming at producing specific bioactive compounds in microbial platforms, as has been done with artemisinin, a potent anti-malarial drug of Artemisia annua [6]. The expansion of knowledge will enhance our ability to modify plants and make rational use of plant biodiversity [5,7].

A major research effort involves molecular and biochemical integrated omic evaluations beyond plant model systems, as well as gauging the effective potential of these medicinal species to afford new bioactive natural products. As part of this effort, seeds and leaves of Erythrina velutina Willd., a species widely used in folk medicine were harvested for evaluation. The species was also chosen considering its family, Fabaceae, an alkaloid rich taxon. Alkaloids are organic nitrogenous bases, an important group of diversely distributed, chemically, biologically, and commercially significant natural products, whose nitrogen atoms originate from an amino acid. Their protective functions in planta range from defense against pathogens and herbivores to abiotic stress mitigation, e.g. drastic changes in light irradiance, water availability and temperature [8]. E. velutina is a native species of the Caatinga [9] having sedative, anticonvulsant, neuroprotective, and anxiolytic activities, mainly acting on central nervous system (CNS) targets [10,11].

This work reports on the combined omic approach of Next Generation Sequencing (NGS)-based transcriptome dataset production and targeted metabolic profiling of a non-model Fabaceae species grown in its natural Caatinga habitat. The results revealed the important chemodiversity of the species and provided insights on the biosynthetic pathway of its unique isoquinoline alkaloids.

Methodology

Plant material

Natural populations of Erythrina velutina Willd., from four locations were selected: two in the city of Acari (location 1: 6°27’44.5”S 36°38’30.6”W and location 2: 6°28’19.5”S 36°38’32.5”W at 270 m
above sea level) and two in the city of Jardim do Seridó (location 1: 6°33’58.2”S 36°43’00.6”W and location 2: 6°33’48.6”S 36°43’11.0”W at 218 m above sea level). Leaves and seeds were collected separately (Fig. S1) and samples composed of 5 to 10 trees for each population were used in a multi-omics approach (transcriptome and targeted metabolite profiling) organized in pools 1 to 8 (Table 1), in which pools 1 to 4 corresponded to seed samples (pool 2 and pool 3 - two harvest locations in the city of Acari; pool 1 and pool 4 - two harvest locations in the city of Jardim do Seridó) and 5 to 8 corresponded to leaf samples (pool 6 and pool 7 - two locations in the city of Acari; pool 5 and pool 8 - two locations in the city of Jardim do Seridó) in the state of Rio Grande do Norte, Brazil, in March 2018, during the dry season at daytime per-

tions in the city of Jardim do Seridó (in the state of Rio Grande do Norte, Brazil, in March 2018, during the dry season at daytime per-

The following criteria were adopted for selecting samples: dark green fully expanded apical leaves from top branches of the top of the tree most exposed to solar radiation, and red seeds attached to ripe pods still hanging from the trees. Samples were quickly washed with distilled water, immediately frozen in liquid nitrogen at the ends of the branches at the dark green fully expanded apical leaves from top branches with iod. The following criteria were adopted for selecting samples:

The acquisition parameter was capillary and cone voltage, 4500 V and 500 V, respectively, capillary temperature at 200 °C, N2 as a drying gas at 4.5 bar and a flow of 9.0 L/min. The spectra were acquired between 100 and 1500 Da. For high resolution mass analysis instruments and experiments were calibrated using an internal standard Na-TFA solution (4 mg/mL) at the end of each run.

The MS/MS experiments were carried out in Auto-MS/MS mode (positive mode) selecting up to 5 precursor ions within a cycle of 3.0 s. An energy gradient of 20 to 50 eV was used for fragmentation of precursor ions. Initially, the blank samples were analyzed to identify interferences and contaminants from the instrument and chromatographic column. These were then excluded from the Auto-MS/MS selection and fragmentation cycle by a function that allows their exclusions.

HR-LC-DAD-ESI-MS/AutoMS data acquisition
Separation and detection of medium-polar compounds were performed in a Shimadzu Class-LC 10, CBM 20A controller and a micrOTOF-QII (Bruker Daltonics) mass spectrometer using electro-
spray ionization (ESI) and Time of Flight (TOF) analyzer. An Xb C18 Kinetex column (100 × 2.1 mm, 2.6 μm, 100 Å), kept at 40 °C, flow rate of 0.400 mL/min, and UV-DAD detector operating between 190 and 400 nm were used. The chromatographic method was based on previous studies [12,13]. The mobile phase was composed of H2O (A) and acetonitrile (B), both with 0.1% (v/v) formic acid using the following gradient: 5–35% B (23 min); 35.0–100% B (17 min), holding at 100% B for 8 min followed by reducing from 100 – 5% B (2 min) and maintained in the initial condition (5% B) for 10 min, totaling 60 min of analysis. Mixed test solutions (mixture of standards - amino acids, flavonoids, etc.) and quality controls (QC - mixture of samples – 40 μL each) were injected to ensure reproducibility of the method, instrument, and chromatographic conditions.

The HRMS experiments were carried out in a micrOTOF-QII mass spectrometer (Bruker Daltonics) in positive ionization mode. The acquisition parameters were capillary and cone voltage, 4500 V and 500 V, respectively, capillary temperature at 200 °C, N2 as a drying gas at 4.5 bar and a flow of 9.0 L/min. The spectra were acquired between 100 and 1500 Da. For high resolution mass analysis instruments and experiments were calibrated using an internal standard Na-TFA solution (4 mg/mL) at the end of each run.

The MS/MS experiments were carried out in Auto-MS/MS mode (positive mode) selecting up to 5 precursor ions within a cycle of 3.0 s. An energy gradient of 20 to 50 eV was used for fragmentation of precursor ions. Initially, the blank samples were analyzed to identify interferences and contaminants from the instrument and chromatographic column. These were then excluded from the Auto-MS/MS selection and fragmentation cycle by a function that allows their exclusions.

**Table 1**

| Pool set | Pools | Harvest site | Geographical coordinates |
|----------|-------|--------------|--------------------------|
| A        | p1, p5 | Jardim do Seridó I | 6°33’58.2”S 36°43’00.6”W |
| B        | p2, p6 | Acari I       | 6°27’44.5”S 36°38’30.8”W |
| C        | p3, p7 | Acari II      | 6°28’19.5”S 36°38’32.5”W |
| D        | p4, p8 | Jardim do Seridó II | 6°33’48.6”S 36°43’11.0”W |
**MS/MS molecular networking analysis**

To perform the Molecular Networking analysis, all MS/MS data were initially converted to mzXML format using MSConvert software and processed using MzMine package (java script) for mass detection and calibration, peak deconvolution, mass grouping (isotopic pattern), and spectral alignment [14]. Once all the data were processed using MzMine software (http://mzmine.github.io/), the spectra were exported in a text format (.mgf) and uploaded to the Global Natural Products Social Molecular Networking (GNPS) platform in order to initiate the calculation of similarity among spectra (cosine similarity score). In this procedure, the spectra are converted into vectors which are paired and compared by a dot product calculation. The cosine similarity score assigned to each pair of vectors ranges from 0 to 1, in which 0 represents completely different and 1 represents identical spectra. The data were filtered by removing all MS2 peaks within ±17 u of the precursor m/z value. MS2 spectra were filtered by choosing only the top six peaks in the ±50 u window throughout the spectrum. The data were clustered with MCLCluster with a parent mass tolerance of 1.0 Da and a MS2 fragment ion tolerance of 0.3 Da to create consensus spectra (for consensus spectra, see https://ccms-ucsd.github.io/GNPSDocumentation/networking/). The network was created by filtering the edges to have a cosine score above 0.65 and more than 6 matched peaks. The spectra in the network were searched against the GNPS spectral libraries (GNPS: https://gnps.ucsd.edu/ProteoSAFe/libraries.jsp), using the same setup as that for the input data. Network visualization was performed in Cytoscape 2.8.3 and 3.4.3 (https://cytoscape.org). To avoid misinterpretation of HPLC contaminants, blank (mobile phase) injections were uploaded as a distinct sample group on GNPS workflow and excluded from the networks.

In addition, the nodes were colored based on the plant organ (leaf and/or seed) of Erythrina and the thickness of the edge attributed to the cosine similarity scores (thickness from 0.95 to 1). Thicker lines were used to indicate greater similarity. Subnets were generated in Cytoscape from isolated parts of the molecular family, in order to improve the analysis of metabolic dynamics, provide more elucidation of molecules, better visibility of edge thickness, and node connectivity. The subnets were based on classification among the Erythrina alkaloid classes.

**Transcriptome analysis**

**Total RNA extraction and cDNA library preparation**

Total RNA was isolated from seeds and leaves using the Trizol™ Plus RNA Purification kit (Life Technologies) according to the manufacturer’s protocol. To obtain high quality RNA, the samples were treated with the RNase-Free DNase set (Qiagen) and subsequently purified with the RNaseasy Mini kit (Qiagen). RNA quality was assessed using the Bioanalyzer 2100 RNA Nano Chip (Agilent) and quantification was performed using the NanoDrop ND-1000 (NanoDrop) spectrophotometer and the Qubit (Thermo Fischer Scientific) fluorometer. All samples that had an RNA Integrity Number (RIN) value above 6.0 were used in the next steps. Each pool was composed of three RNA samples and a total of 0.5 value above 6.0 were used in the next steps. Each pool was composed of three RNA samples and a total of 0.5

**RNA sequencing analysis and annotation**

Total RNA was sequenced according to the standard protocols of the NextSeq 500 (Illumina) platform, generating 150 paired-end readings. Composite samples obtained from at least three specimens yielded four seed RNA pools (1–4) and four leaves RNA pools (5–8), in accordance with the above mentioned harvest scheme. After RNA sequencing, FastQ files were generated and transferred from the GENIAL facility at CEAP-USP to the Bioinformatics, Bio-statistics and Computational Biology Core-B3C at Federal University of Semiarid (UFERSA-RN-Brazil). The massive analysis was performed in the Computational Cluster of High-Performance Computing Center-NPAD at UFRN. Initially, rRNA removal was performed using the SortMeRNA software (v2.1) (https://bioinfo.iilf.fr/RNA/sortmerNA) for all readings that combine the Forward (R1) and Reverse (R2) strands. This step was necessary to improve the quality of the readings. Sequence quality was evaluated by the FastQC software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The reading libraries were trimmed by Trimmmomatic (http://www.usadellab.org/cms/index.php?page=trimmomatic) to remove poor quality adapters and readings, in which the sliding window cut was 2:20.

The range of final readings was between 21 and 151 base pairs and indexed up to 20. The FastQ files described above were used in a Dassembly in the Trinity software (https://github.com/trinityrnaseq/trinityrnaseq/wiki) because the E. velutina full genome sequence is not available yet. A k-mer size range 21 was used to obtain the best results for mapping genes from plant orthologs in the BUSCO software (https://busco.ezlab.org/) aiming at validating the assemblies generated in Trinity. This tool provides quantitative measures for the assessment of genome assembly, gene set, and transcriptome completeness, based on evolutionarily-informed expectations of gene content from near-universal single-copy orthologs selected from OrthoDB v9 (https://www.orthodb.org/v9/index.html) and crossing with the Unigene database. Additionally, the software Bowtie2 (http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) was used to align sequencing reads with final assembled sequences aiming to validate the transcriptome.

The assembled transcripts were submitted to the InterProScan database strings (v5.30–69.0) (http://www.ebi.ac.uk/interpro/), which combines signatures from various databases (PROSITE, Pfam, PRINTS, ProDom, SMART, TIGRFAMs, PIRSF, SUPERFAMILY, GENE3D and PANTHER) for a primary classification. In this step, an identification parameter of at least 50% coverage and 70% identity of the functional domain was used before data input onto the platform of the aforementioned databases. After domain identification, transcripts containing sequences corresponding to candidate protein domains were selected to participate in the study of alkaloid biosynthesis according to data from public databases and from the literature. All identified sequences of E. velutina are “like sequences”, but the suffix was omitted for the sake of simplicity.

The sequences were subjected to a second run of BLASTx annotation against the 469,154 strings of UniProtKB/SWISS-PROT bank (https://www.uniprot.org/). A minimum of 30% identity, 50% coverage and 10E-value were used as parameters, in addition to the manual analysis of the alignment, in which sequence size and coverage were checked for the presence of substrate/active site. All transcripts were quantified as the number of transcripts per million (TPM) by the Salmon software [15].

**Climatic parameters**

Trimestral values of temperature, accumulated precipitation, and accumulated surface net solar radiation in the area where plants were harvested were obtained by using the European Centre for Medium-Range Weather Forecasts’ (ECMWF) ERA-Interim reanalysis [16]. For the same year, fire activity was analyzed based...
Erythrina alkaloids can be classified into dienoid, alkenoid, and lactone subclasses, based on their core moieties [10,11], a conserved and characteristic chemical nucleus. The molecular diversity among them is attributed, except for the lactone ring, to the substitution patterns, such as –OH, –CH₃, –CH₂O and glucosyl residues, usually found at C3, C8 and C11 of the alkaloidal main core skeleton. Herein, the targeted metabolic analysis was used to investigate the chemical diversity of the alkaloids in the crude extract, in a comparative way among leaves and seeds.

An initial analysis of the samples by 2D TOCSY-NMR (1H–1H) enabled us to compare the experimental data with an in-house database by including all the experimental 1H NMR spectral data reported to the alkaloids previously identified in *Erythrina* spp. The analysis criteria were related to spectral similarity values based on the algorithm of pattern recognition between the in-house database and the experimental data. These data pointed out the occurrence of alkenoid and dienoid scaffolds even though the unequivocal identification of the alkaloids was not possible due to the overlapping of chemical shifts present in different but similar molecules. Therefore, a more sensitive and orthogonal metabolomic approach was used in a complementary perspective. In order to obtain a more comprehensive analysis and to visualize structurally related molecules, an MS/MS Molecular Networking (MolNet) approach was employed as a dereplication strategy for metabolite annotation The MolNet protocol was based on cosine similarity score and fragmentation pattern with comparative targeted analysis, using the web-based platform Global Natural Products Social Molecular Networking (GNPS) [17], as well as an in-house database that compiles the previously reported mass spectral data on these alkaloids. The level of identification confidence in this study was 2 or 3 in a scale that varies from 1 to 4 [18,19] where it was possible to putatively annotated compounds or characterized compound classes. The 42 alkaloids detected (Fig. 1) were segregated in two clusters (both clusters have alkaloids of the dienoid and alkenoid subclass, but not lactonic) (Fig. 2B); these, 25 of them (60%) were annotated (Table S11) and 17 (40%) were unannotated (Table S11 and S12).

As the network produced by GNPS platform groups compounds with similar MS/MS spectra, the evaluation of MS/MS spectra could display information about Erythrina alkaloid scaffolds (dienoid, alkenoid, and lactone), once each subclass has different fragmentation pathways [20]. Retro Diels-Alder (RDA) is noted usually in alkenoid subclass while substituents elimination in Ring A, as for example CH₃OH (32 u), are commonly noted in dienoid subclass due to the resonance produced that aids to stabilize the charge [20].

Surely, other Erythrina alkaloid fragmentations in the mass spectra were observed as water, indole and sugar elimination although they are not always related to a specific subclass [20]. The analysis of these characteristic fragments of each subclass and the others fragments present in mass spectra aid to understand the compounds clustering including the unannotated compounds (Fig. 2C, Table S2, Table S11 and Table S12). The cluster 1 could be related mainly to dienoid alkaloids while other alkaloids including alkenoid subclass are included in cluster 2. The 17 unannotated alkaloids display characteristic mass spectrum of the Erythrina alkaloids, therefore were grouped in the clusters after MolNet analysis, however they did not match fully the previously reported alkaloid to the *Erythrina* genus. It underlines the chemical diversity of *E. velutina* when collected in its natural semiarid environment with its unique edaphoclimatic features. Such unexplored chemodiversity constitutes a reservoir of potentially useful metabolites still to be isolated, unequivocally assigned, and tested against pharmacological targets.

We hypothesized that the edaphoclimatic features in the Caatinga biome could select and elicit biosynthesis of novel and diverse metabolites in local species. Furthermore, native species from Caatinga have been scarcely studied so far at the metabolites level. These data provide relevant information, affording a representative picture of plant chemistry/interaction with the environment, which is particularly relevant under the multifactorial biotic and abiotic elicitors imposed by semiarid conditions. Details about the edaphoclimatic conditions prevailing at the harvesting site are provided in Supplementary Material (Fig. S3).

Given that trees used in the present experiments are under natural growth conditions, some alkaloids were heterogeneously distributed among the samples (Fig. S4 and Table S3). These differences were expected, in contrast to most experiments performed under controlled conditions, in which less variation is apparent. Modulation of plant secondary metabolites, including alkaloids, by macronutrients, light intensity, temperature, and soil salinity has been shown in some plant species [21]. Although it is not possible to establish a clear correlation between the production of different alkaloids and soil features in the present investigation, these factors must be considered as part of the multifactorial environmental equation. The highest chemical diversity was observed on the pools 2, 3 and 7, in which 11, 13 and 17 different structures of alkaloids were identified, respectively (Table S3). Interestingly, pools 3 and 7 (pool set C) were harvested at the same site in Acari. Soil analysis showed that this location has some peculiar characteristics, such as high soil pH, high calcium and potassium contents, low amounts of iron and the presence of chromium, a heavy metal not detected elsewhere (Table S7). Mineral nutrition, pH, and heavy metals have been shown to affect alkaloid accumulation. Levels of indole alkaloids in *Psychotria brachyceras*, for example, were significantly increased by exposure to NaCl or to micromolar concentrations of aluminum and silver, which has been partly attributed to moderate intensity oxidative stress [22].

Even though glycosylation constitutes an effective tool to increase bioavailability and bioactivity of benzylisoquinoline alkaloids [23], it is rare to detect the glycosyl form of Erythrina alkaloids and few of them have been reported so far [11]. Zhang and colleagues [24] identified five glycosylated Erythrina alkaloids by 1D and 2D NMR; in all of them glycosylation was at C11, whereas C15 and C16 had methoxy or methylenedioxy groups. This low availability of data on glycosyl form of alkaloids in the genus could be related to differences in post-harvest procedures, as well to the analytical strategy used in most studies, e.g. varying the pH for high and low levels, that are not suitable for preserve the original chemical structure for some substances [25]. It is likely that the rigorous experimental procedures involved in the present analysis with a targeted metabolome perspective contributed to preserve the original sample; in addition, the environmental conditions of the Caatinga could be involved in promoting the production of this kind of metabolite. Some general features of the annotated alkaloids can be highlighted regarding structure and organ distribution. Seed-exclusive alkaloids showed hydroxyl substituents at C3, C11, and, to a lesser extent, at C2, whereas C3 could also present an O-methylation (numbering alkaloid main scaffold in Fig. 2A). Interestingly, seven glycosylated alkaloids were detected exclusively in seeds. To the best of our knowledge, their proposed structures
are unprecedented (Table S1 - alkaloid number 11, 12, 15, 19, 20, 21, 22). Glycosyl residues frequently were attached at C15 and C16; otherwise these positions could be involved in a methylene-dioxy bridge (R-O-CH₂-O-R) (Fig. 4B). Although the biological functions of alkaloids in seeds are not completely defined [26], they may provide defense against herbivores or pathogens and perhaps take part in nitrogen metabolism during germination [8]. Erythraline-11β-O-glucopyranoside, identified in seeds of *E. christa-galli*, showed inhibitory effect against tobacco mosaic virus [27]. The effects of glycosylation of specialized metabolites are manifold, including modulation of transport, stability, storage, and stress hormones [28]. The presence of glycosylated alkaloids in the examined *E. velutina* organs may be relevant as non-enzymatic antioxidants, as reported for other classes of glycosylated alkaloids [29]. The seasonally anticipated or directly acting extreme conditions of the Caatinga dry season (heat, excess irradiance and low water availability, Fig. S3) may also have promoted alkaloid accumulation aiming at redox regulation in plants. An additional reason for the relatively higher abundance of glycosylated alkaloids in seeds of *E. velutina* is the fact that lipids represent the second most abundant reserve in this species. Lipid reserve mobilization may generate hydrogen peroxide during germination and early seedling establishment, particularly at higher temperatures [30,31].

Alkaloids solely found and annotated in leaves share the feature of O-methylation at C3, and the oxidized form (C=O) is predominant at C11 and, to a lesser extent, at C2. In the alkaloids shared by both organs main characteristics included predominance of OH at C2, an O-methylation pattern at C3, whereas C15 and C16 were mainly involved in a methylenedioxy ring. The role of BIA in planta includes protection against herbivores, fungi and bacteria, which contributes to increased plant reproductive success [8].

Divergent metabolite accumulation profiles could be observed in leaves and seeds and establishing intrinsic patterns is not trivial. The relative signal intensity in m/z 492.113 (alkaloid 22) with peak height of 3.69E+03 was used to normalize the abundance of all alkaloids (Table S3 and Fig. S4). The m/z 298.0966 alkaloid, tentatively annotated as coccolinine or isococcolinine (alkaloid 8, Table S3) at peak height of 1.75E+05, was the most prominent alkaloid present in both leaves and seeds. In addition to coccudinone (alkaloid 2 – Table S1), this molecule was recorded for the first time in the *Erythrina* genus. Previously, these structures were identified in *Cocculus laurifolius* (Ranunculales order, Menispermaceae). Despite the taxonomic distance, they share several structural fea-
Alkaloids represent one of the most diverse classes of plant metabolites both in terms of structure and bioactivity. About 20% of plant species accumulate alkaloids [34], which can also be found in several classes of other living organisms, forming a group of chemically diverse natural products, abundant both in number and form [35]. These molecules also constitute an interesting source of chemical structures with medicinal properties [36]. Some of the best-known alkaloid-based medicines are the benzylisoquinoline alkaloids (BIAs) morphine and codeine. Indeed, opium poppy (Papaver somniferum) has emerged as a powerful “non-model model” system for studying biosynthesis and production of BIAs relevant in the pharmaceutical field [37,38]. BIA diversity is remarkable, spanning over 2,000 known structures, also present in related taxa, such as Coptis japonica, Thalictrum flavum and Eschscholzia californica [39]. However, beyond the classical Ranunculales distribution of BIAs, lie unique molecules, such as those found in Fabaceae (Erythrina alkaloids). The comparative analysis based on gene similarity data from BIA-producing plants represents a useful tool to suggest candidate genes in E. velutina and associate them with steps in the proposed alkaloid biosynthetic pathway.

Candidate genes for a biosynthetic pathway of E. velutina alkaloids

Data from RNA-seq analysis has been used to identify several genes encoding biosynthetic pathways from across plant specialized metabolism, including alkaloids [40,41]. However, this scenario is more complex in non-model plants that do not yet have a genome sequence available. Methods based on alignment with reference species and de novo assembly can generate specific and testable hypotheses at the molecular level about metabolic pathways in non-model plants [42]. Therefore, the transcriptome of seeds and leaves of E. velutina was analyzed trying to assess main candidate genes involved in the alkaloid biosynthesis.

The primary analysis generated a total of 6,251,021 (Table S5) transcripts among the eight pools. The numbers of transcripts identified in each pool are listed in Table S4. Databases and previously reported information on other BIA-producing species were also searched in order to select suitable domains for candidate genes involved in E. velutina alkaloid biosynthetic pathways. The annotated and manually selected transcripts were then subjected to a local analysis via BLASTx. Several transcripts showed similarity with species-specific sequences of other BIA-producing plants (Table S8). Based on the transcripts identified as candidate genes, on the annotated alkaloid features, and on previous reports, a puta-

Fig. 2. Molecular networking and annotated alkaloids. (A) Representation of the tetracyclic spiroamine of the alkaloids of the genus Erythrina and their respective classes (B) Molecular networking with MS/MS data from hydroethanolic crude extracts analysis of Erythrina velutina. Green nodes: alkaloids exclusively in leaves; orange nodes: alkaloids exclusively in seeds; yellow nodes: alkaloids in both leaves and seeds, and gray nodes (leaves: pentagon, seeds: square): alkaloids identified but not annotated (they did not match fully the mass spectrum of alkaloids previously reported in Erythrina spp.). The high resolution mass value for each alkaloid is shown in the node. Thicker and closer lines indicate greater similarity between the chemical entities. (*) highlight the mass values that were repeated in the cluster (may represent isomers). Red-bordered nodes indicate glycosylated alkaloids. Clusters 1 and 2 are not related to the division of categories of the dienoid and alkenoid classes. (C) Structures and characteristic losses due to substituent elimination observed in dienoid subclass fragmentation that were also present in MS/MS spectra of the alkaloids annotated in cluster 1.

Daisy Sotero Chacon, Taffarel Melo Torres, Ivanice Bezerra da Silva et al. Journal of Advanced Research 34 (2021) 123–136

129
Fig. 3. Putative biosynthetic pathway of Erythrina alkaloids. Erythrina alkaloids are isoquinoline-type, forming a component of the phenylalanine precursor. The pathway begins with branches initiated with 4-HPP and tyrosine, which by enzymatic action form 4-HPAA and dopamine, respectively, producing intermediate (S)-norcoclaurine and coclaurine. The CYP80B3 part (S)-methylnorlaudanosoline will be formed and is one of the crucial points in the chemical cascade of Erythrina alkaloids, as part of it may generate a bifurcation to (S)-norreticuline (pathways in green) and norprotosinomenine (pathways in pink). Pathways in yellow color represent the alkaloids annotated in this study (in accordance with Table S1); Dotted arrows indicate reactions not yet fully established; Blue transcripts were identified in our transcriptome data set according to the Blastx result. Structures of the seven glycosylated alkaloids [11,12,15,19,20,21,22] are shown in Fig. 4. Abbreviations: 4-HPP, 4-Hydroxyphenylpyruvate; 4-HPAA, 4-hydroxyphenylacetaldehyde; 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-METHYL-COCCLAURINE 3-OHYDROXYLASE; 4OMT, 3-OHYDROXY-N-METHYL-(S)-COCLAURINE 4-O-METHYLTRANSFERASE; N4OMT, NORBELLADINE 4-O-METHYLTRANSFERASE; BBE, BERBERINE BRIDGE ENZYME; CYP719B1, SALUTARIDINE SYNTHASE; CYP712A2, (S)-STYLOPINE SYNTHASE; SALR, SALUTARIDINE REDUCTASE; SALAT, SALUTARIDINOL 7-O-ACETYLTRANSFERASE; CYP96T1, NOROXOMARITIDINE SYNTHASE; T60DM, THEBAINE 6-O-DEMETHYLASE; COR, NADPH-DEPENDENT CODEINONE REDUCTASE; SARED, SANGUINARINE REDUCTASE.
Fig. 4. Gene expression heatmap. The transcriptional expression profile of the candidate genes involved in the Erythrina alkaloid biosynthesis pathway is illustrated by a color gradient. All transcript expression values were scaled by TPM between low and high expression. (A) Orange different color gradients indicate low or high number of transcripts in each identified pool after annotation by domain; Palette with different color gradients in blue indicate the low or high expressed transcripts. Transcripts considered as possible isoforms were identified and named as “like”. Heatmap in green and orange box indicate transcripts identified exclusively in leaves and seeds, while in blue box, found in both plant structures. (B) Structure of the glycosylated alkaloids identified in the metabolic profile. (C) Heatmap of transcripts selected by domain with catalytic glycosylation function and subsequently identified using Blastx in seeds and leaves (UDP-glycosyltransferase). (D) Heatmap of the number of transcripts identified with the UDP-glycosyltransferase domain. P1-p4: transcripts expressed in seeds and p5-p8: transcripts expressed in leaves.
The biosynthetic pathway for benzylisoquinoline alkaloids derives from aromatic amino acids, starting with the activity of enzymes encoded by *EvTAT* and *EvTYDC*. Next, the condensation of 4-hydroxyphenylacetaldehyde (4-HPAA) and dopamine, two L-Tyr derivatives, is catalyzed by norcoclaurine synthase (*EvNCS*), which produces \((S)-\)norcoclaurine [43]. The transcript corresponding to *EvNCS* was identified only in leaves (Fig. 4). Subsequent methylation of \((S)-\)norcoclaurine produces \((S)-\)coclaurine [44,45], followed by the synthesis of the intermediate \((S)-\)methylnorlaudanosoline by \((S)-N\)-methylcoclaurine 3'-hydroxylase (encoded by *EvCYP80B3*, which was expressed in both leaves and seeds). In opium poppy, the overexpression of *PsCYP80B3* increased total alkaloid yield, as well as morphine biosynthesis [46].

The transcriptome analysis also yielded the gene candidates *Ev6OMT* and *Ev7OMT* as involved in the biosynthesis of \((S)-\)coclaurine. *Ev7OMT* was compatible with that of *P. somniferum* [47,48]; however, *Ev6OMT* was more common (Fig. 4). OMTs contribute to the formation of several alkaloid chemical skeletons, such as the isoquinoline monoterpene alkaloid emetine [49] and the last step of the biosynthesis of vindoline, a *Catharanthus roseus* indole alkaloid [50]. Expression of OMT genes seems to be tightly regulated, and the involvement of miRNAs in their regulation has been shown in *P. somniferum* [51].

\((S)-\)methylnorlaudanosoline could be one of the pivotal branching points in the biosynthesis of *Erythrina* alkaloids. This com-
pound is a common intermediate in the production of both (S)-norprotosinomenine – part of the “abnormal” Erythrina alkaloids (cocculine alkaloids pathway) - and (S)-norreticuline (Fig. 3).

(S)-norreticuline, produced by 4’-O-methyltransferase (encoded by Ev4OMT), present in leaves and seeds, is involved in the pathway of the so-called “common Erythrina alkaloids”. These are metabolites usually found in the genus *Erythrina* which contain a tetra cyclic spiroamine in their chemical skeleton [52]. 4OMT has been previously described in BIA-producing species, such as *Coptis japonica* [53], *Papaver somniferum* [54], *Eschscholzia californica* [55] and *Narcissus* spp. [56]. A transcript present in one seed pool, encoding norbelladine 4’-O-methyltransferase (N4OMT), was previously reported in Amaryllidaceae as being involved in galantamine biosynthesis, a BIA alkaloid clinically used to treat Alzheimer’s disease [57].

In general, methyltransferases previously described in BIA biosynthesis accept a wide variety of multi-structure alkaloid substrates. An example is cocaine N-methyltransferase (encoded by *EvCNMT*), previously characterized as part of the biosynthesis of papaverine, an important isoquinoline opium poppy alkaloid [47-53]. The enzyme acts on the central pathway of BIA and converts a non-methylated substance from secondary amine to a tertiary one [53,54].

Subsequently, after formation of the tertiary amine, C-C bonds can be catalyzed by the proteins encoded by berberine bridge enzyme (*EvBBE*)/EvCYP719B1, through the stereospecific conversion of the methyl portion of intermediate (I) to (II) and methane dioxi bridges formation catalyzed by EvCYP986T1/EvCYP719A, that converts the intermediate (II) to noramarine. BBE of cultured opium poppy cells was characterized and its gene successively expressed in *Saccharomyces cerevisiae*, yielding catalytically active enzyme [60-62]. Transcripts encoding cytochrome P450 enzymes are important in contributing to the structural diversity of molecules, including BIAS [63]. Several enzyme members of the subfamily [CYP719A9, CYP719A11, CYP719A2, CYP719A3, CYP719A1 and CYP80B3] are involved in isouquinoline moiety biosynthesis [64,65], whereas CYP986T1 was capable of catalyzing the formation of C-C coupling in Amaryllidaceae alkaloid biosynthesis [66]. Most transcripts encoding CYP719A and CYP80B2 enzymes were highly expressed in both leaves and seeds. CYP719A2 and CYP719A3 cloned from *Eschscholzia californica* cells showed catalytic activity in methylenedioxy bridge formation and methyl jasmonate treatment resulted in the coordinated induction of these genes [66].

The core skeleton of Erythrina alkaloids, a tetra cyclic spiroamine, would be produced by subsequent steps of rearrangement to form the cation neospirinic, dibenz azonine to yield a diallylic cation from noramarine [56,67]. We propose that the result of this pathway would be salutaridine (*EvSALR*), from the intermediate (IV) (reaction A), an enzyme dehydrogenase member and short chain reductase family that in *P. somniferum* use the NADPH as coenzyme, to afford salutaridinol, mainly involved in the morphine and codeine biosynthesis [68].

The next step could be catalyzed by salutaridinol 7-O-acetyltransferase (*EvSALAT*) in our proposal through reaction B, which specifically acetylates salutaridinol epimer-7 into salutaridinol-7-O-acetate. This step was previously described in the formation of BIA in opium poppy. The acetyl group could be spontaneously eliminated (catalyzing the conversion of intermediate (V) to (VI), in a pH-dependent way as reported in Opium poppy, leading to the formation of an oxide bridge between C4 and C5 to produce thebaaine, the first pentacyclic alkaloid prior to morphine formation [37]. Evidence from a gene silencing study suggests that *SALR* and *SALAT* interact and may act as an enzyme complex in opium poppy [69]. *EvSALAT* was found only in seeds. After spontaneous deacetylation, thebaaine 6-O-dimethylase (*EvT6ODM*), a member of the Fe(II)/2-oxoglutarate-dependent dioxygenase family what form neopinone [70], would removes the methyl group and changes the unsaturated bonds in the neighboring carbon (reaction C). Regarding Erythrina alkaloid biosynthesis, these kinds of reactions are also necessary to produce the oxygenated and unsaturated A ring, present in several alkaloids isolated from this genus. Therefore, we propose *SALAT*, and *T6ODM*, both of which seed-specific, as candidate genes in the pathway (Fig. 4).

Finally, the unsaturation of a C-C bond is necessary in the heterocycle of several Erythrina alkaloid moieties. Therefore, we propose the occurrence of an equilibrium between two *Erythrina* intermediates, like the previously observed between neopinone-/codeinone intermediates in *P. somniferum* [71]. The subsequent steps involve oxidation reactions; transcripts compatible with codeinone reductase (*EvCOR*) [72] and sanguinarine reductase isolated from cell cultures of *Eschscholzia californica* (*EvSARED*) [73] were found in the leaves and seeds transcriptome dataset. These candidate genes encode enzymes that would enable oxidation and reduction reactions to achieve the final structure of the Erythrina alkaloids.

Identifying cytochrome P450 enzymes in plants is a challenge, and the information about them is still rather limited; among all crystallized CYPs only five of them are from plants [74]. Therefore, besides the CYPs discussed and pointed as gene candidates in the Erythrina alkaloids biosynthesis since they matched our methodologic parameters, other CYPs reported in plant metabolism and identified in the *E. velutina* transcriptome data set were highlighted (Fig. 4). Among them, CYP82N4 and CYP71D12 are also involved in the alkaloid biosynthetic pathways [75,76].

As pointed out above, predominant annotation of seed-exclusive glycosylated alkaloids in *Erythrina* was observed, however, almost twice as many transcripts with a functional domain of UDP-glycosyltransferase were detected in leaves (Fig. 4 and Table S9). Although UDP-glycosyltransferases are a large family of proteins involved in diverse metabolic processes, these results might indicate the possibility of translocation of alkaloids between organs and/or variations of aglycone modifying enzymes in different plant cell types.

Some alkaloids detected in the metabolic profiling dataset belong to a class named cocculine alkaloids or “abnormal Erythrina alkaloids” (Fig. 3). The chemical skeleton of such alkaloids contains a tetra cyclic spiroamine moiety, but they were originally characterized from *Cocculus laurifolius*. Data regarding the biosynthesis of the cocculine alkaloids are scarce, but norprotosinomenine has been indicated as an intermediate with stronger isotope incorporation than norreticuline [32,33]. The occurrence of both uncommon and common Erythrina alkaloids in *E. velutina* and the presence of candidate genes possibly involved in both pathways (Fig. 3) suggest that the most appropriate term to refer to these alkaloids is “tetra cyclic spiroamine BIA”. This would acknowledge the fact that these alkaloids could be synthesized from both precursors: (S)- norprotosinomenine as well as (S)-norcocularine.

The species *Cocculus laurifolius* (order Ranunculales) and *Erythrina velutina* (order Fabales) are included in the great clade of eudicots. The Ranunculales order is a valuable source of benzylisoquinoline alkaloids with important alkaloid-producing families: Menispermaceae, that includes species of the genus *Cocculus*. Berberidaceae, such as the genus *Berberis*, which comprises *B. vulgaris* (the main natural source of berberine) [77], and Papaveraceae (*Papaver somniferum*), which accumulate BIAs, such as morphine, codeine, and noscapine [39]. Fabaceae is taxonomically more distant but is included in the same monophyletic group [78].

Phylogenetic, biochemical and molecular approaches were considered to conduct an empirical investigation on the evolution of BIA synthesis in plants indicating that the ability of angiosperms to produce BIAs appears to have originated shortly before the appearance of eudicots [79]. These studies support a monophyletic
evolution of proteins involved in the biosynthesis of BIAs in angiosperms and suggest that the ecophysiological functions of specific BIAs was generated by selective environmental pressures and/or the emergence of specialized physiological mechanisms that contributed to the maintenance or reactivation of the pathway. The large clade of eudicots seems to be an important evolutionary branch for the synthesis of several bioactive tyrosine derivatives and formation of alkaloids.

In recent years, efforts have been made on biosynthesis research and pharmaceutical use of isoquinoline alkaloids; however, the results are mostly focused on opium poppy and some Amaryllidaceae species. The present study provides a first omic analysis of the importance of Fabaceae and E. velutina in this scenario, bringing to the spotlight a non-model plant growing in its naturally harsh environment. Overall, there was good agreement between transcriptome and metabolome data, affording a reasonable picture of BIA biosynthesis and part of its regulation in this species.

Conclusion

This detailed macro and micro molecular approach applied to seeds and leaves of E. velutina harvested in Caatinga from four different locations revealed 42 alkaloids by HR-LC-MS/MS and MolNet in a targeted metabolomic approach. Furthermore, based on the combined evidence, 24 gene candidates were put together in a putative biosynthetic pathway. These results provide a framework to advance on further findings to understand the real role of each step in the route as well as the flux controller enzymes leading to the singular alkaloid diversity of this species. In addition, these results contribute to show potential biotechnological targets and to stimulate future studies on Brazilian biodiversity detailed at the level of biosynthesis of bioactive natural products, notably alkaloids. Finally, in an important and necessary way, the datasets presented herein add new data to public databases on metabolome and transcriptome from non-model medicinal plants.

Accession numbers

The transcriptome sequence readings were deposited in the Sequence Read Archive (SRA) under the BioProject accession number: PRJNA668524.

Financing

This work was supported by the Serrapilheira Institute (grant number Serra-1709-19691), by the Ministry of Science, Technology, Innovation and Communications – MCTIC, CNPQ/National Council of Science and Technology – INCT BioNat, [grant number 465637/2014-0], São Paulo Research Foundation (FAPESP) [grant INCTBioNat 2014/50926-0] and by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) [CAPES – Finance Code 001].

Author contributions

D.S.C., I.B.S., F.P.S.R., A.A.R., D.S. and A.P. carried out the experimental procedures. D.S.C., F.P.S.R., D.S., A.P., F.A.S.D.P., R.F., A.S.P., J.A.S.Z., A.J.C., L.S.F. and N.P.L. analyzed and discussed chemical data. D.S.C., I.B.S., C.T.C., J.V., T.M.T., T.F.A., E.L.V., K.C.S., L.V.S. analyzed and discussed molecular data. R.L., J.A.R. and F.L.M.S. analyzed and discussed the edaphoclimatic data. R.B.G. and A.G.F.N., planned and discussed molecular data. R.L., J.A.R. and F.L.M.S. analyzed and discussed the edaphoclimatic data. All authors discussed the major conclusions and contributed to write the manuscript.

Compliance with Ethics Requirements

This article does not contain any studies with human or animal subjects.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the research supported by LNB – Brazilian Biorenewables National Laboratory (CNPEM/MCTIC) during the use of the High Throughput Sequencing (NGS) open access facility. Also, we thank the Core Facility for Scientific Research–University of São Paulo (CEFAP-USP/GENIAL (Genome Investigation and Analysis Laboratory)) for the NextSeq 500 (illumina) analysis. Ministry of Science, Technology, Innovation and Communications – MCTIC, FAPESP (FAPESP: grant numbers 2009/54083-9R, 2013/07600–3, 2014/05935–0 [DMS], 2017/06446–2 [DMS], and 2016/13292–8 [ACP]) and INCT BioNat. D.S.C and F.P.S.R. are recipients of fellowship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). JASZ acknowledges CNPQ for the research fellowship.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jare.2021.01.017.

References

[1] Yonekura-Sakakibara K, Saito K. Functional genomics for plant natural product biosynthesis. Nat Prod Rep 2009:1486–87. doi: https://doi.org/10.1039/b817077e
[2] Matsuura HN, Malik S, de Costa F, Yousefzadi M, Mirjalili MH, Arroo R, et al. Specialized Plant Metabolism Characteristics and Impact on Target Molecule Biotechnological Production. Mol Biotechnol 2018:169–83. doi: https://doi.org/10.1007/s12095-017-9697-0
[3] Rodrigues T, Becker D, Schneider P, Schneider G. Counting on natural products for drug design. Nat Chem 2016:531–40. doi: https://doi.org/10.1038/nchem.2479
[4] Debnath M, Pandey M, Bisen PS. An Omics Approach to Understand the Plant Abiotic Stress, OMICS: A Journal of. Integr Biol 2011:739–62. doi: https://doi.org/10.1089/omi.2010.0468
[5] Chang C, Bowman JL, Meyerowitz EM. Field Guide to Plant Model Systems. Cell 2016:325–39. doi: https://doi.org/10.1016/j.cell.2016.08.031
[6] Pham J, Vilm A, Feliz M, Maffetone N, Walker JR, et al. A Review of the Microbial Production of Bioactive Natural Products and Biologics. Front Microbiol 2019. doi: https://doi.org/10.3389/fmicb.2019.01404
[7] Zhu H, Li C, Gao C. Applications of CRISPR-Cas in agriculture and plant biotechnology. Nat Rev Mol Cell Biol 2020:1–17. doi: https://doi.org/10.1038/s41580-020-00288-9
[8] Matsuura HN, Fett-Neto AG. Plant alkaloids: main features, toxicity, and mechanisms of action. Plant toxins 2015:1–15. doi: https://doi.org/10.1007/s41786-014-0001-7
[9] Beissa MC, Lacerda CF, Amarino AV, Bezerra AM, Lima AD. Mechanisms of salt tolerance in seedlings of six woody native species of the Brazilian semi-arid 2017:157–65.
[10] Rambo DF, Biegelmeier R, Tonson NSB, Dresch RR, Moreno PRH, Henriques AT. The genus Erythrina L.: A review on its alkaloids, preclinical, and clinical studies. Phytother Res 2019:258–275. doi: https://doi.org/10.1002/ptr.6321
[11] Fahmy NM, Al-sayed E, El-shazly M, Singab AN. Alkaloids of genus Erythrina: An updated review. Nat Prod Res 2019:1–22. doi: https://doi.org/10.1080/17435390.2018.154500
[12] De Vos RC, Moc M, Lommen A, Keurentjes JJ, Bino RJ, Hall RD. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat Protoc 2007:778–91. doi: https://doi.org/10.1038/nprot.2007.9
[13] Lise JC, Schauer N, Kopka J, Willmitzer L, Fernie AR. Gas chromatography mass spectrometry-based metabolite profiling in plants. Nat Protoc 2006:387–96. doi: https://doi.org/10.1038/nprot.2006.96
benzophenanthridine alkaloids in the response of plants to pathogenic attack, Proc Natl Acad Sci. 1991;9969–73. http://doi.org/10.1073/pnas.88.22.9969.  

[62] Winkler A, Puhl M, Weber H, Kutschau TM, Gruber K, Macheroux P. Berberine bridge enzyme catalyzes the six-electron oxidation of (S)-reticuline to dehydroscoulerine. Phytochemistry 2009;1092–7. doi: https://doi.org/10.1016/j.phytochem.2009.05.005.  

[63] Zeng J, Liu Y, Liu W, Liu X, Liu F, Huang P, et al. Integration of transcriptome, proteome and metabolism data reveals the alkaloids biosynthesis in Macleoya cordata and Macleoya microcarpa. PLoS ONE 2013;1–18. doi: https://doi.org/10.1371/journal.pone.0053409.  

[64] Ikezawa N, Iwasa K, Sato F. CYP719A subfamily of cytochrome P450 oxygenases and isoquinoline alkaloid biosynthesis in Eschscholzia californica. Plant Cell Rep 2008:123–33. doi: https://doi.org/10.1007/s00299-008-0624-8.  

[65] Kilgore MB, Augustin MM, May GD, Crow JA, Kutschau TM. CYP96T1 of Narcissus sp. aff. pseudonarcissus Catalyzes Formation of the Para-Para' C-C Phenol Couple in the Amaryllidaceae Alkaloids, Frontiers. Plant Sci 2016;7. doi: https://doi.org/10.3389/fpls.2016.00225.  

[66] Ikezawa N, Iwasa K, Sato F. Molecular cloning and characterization of methylenedioxy bridge-forming enzymes involved in stylopine biosynthesis in Eschscholzia californica. Febs J 2007:1019–35. doi: https://doi.org/10.1111/j.1742-4658.2007.05652.x.  

[67] Barton DHR, James R, Kirby GW, Turner DW, Widdowson DA. Phenol Oxidation and Biosynthesis. Part XVlll. The Structure and Biosynthesis of Erythrina Alkaloids. J Chem Soc 1968:1529–37. doi: https://doi.org/10.1039/j39680001529.  

[68] Higashi Y, Kutschau TM, Smith TJ. Atomic Structure of Salutaridine Reductase from the Opium Poppy (Papaver somniferum). J Biol Chem 2010:6532–41. doi: https://doi.org/10.1074/jbc.m109.088989.  

[69] Kempe K, Higashi Y, Frick S, Sabarna K, Kutschau TM. RNAi suppression of the morphine biosynthetic gene SALAT and evidence of association of pathway enzymes. Phytochemistry 2009;579–89. doi: https://doi.org/10.1016/j.phytochem.2009.03.002.  

[70] Hagel JM, Facchini PJ. Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. Nat Chem Biol 2010:273–5. doi: https://doi.org/10.1038/nchembio.317.  

[71] Lippen B, Miller JAC, Allen RS, Chitty JA, Gerlach WL, Frick S, et al. Increasing morphinan alkaloid production by over-expressing codeine reductase in transgenic Papaver somniferum. Plant Biotechnol J 2007;26–37. doi: https://doi.org/10.1111/j.1467-7652.2006.00212.x.  

[72] Unterlinner B, Lenz R, Kutschau TM. Molecular cloning and functional expression of codeine reductase: the penultimate enzyme in morphine biosynthesis in the opium poppy Papaver somniferum. Plant J 1999:465–75. doi: https://doi.org/10.1046/j.1365-313X.1999.00470.x.  

[73] Vogel M, Lawson M, Spilw, Conrad U, Roos W. Structure and mechanism of sanguinarine reductase, an enzyme of alkaloid detoxification. J Biol Chem 2010:18397–406. doi: https://doi.org/10.1074/jbc.m110.168633.  

[74] Beaudoin GAW, Facchini PJ. Isolation and characterization of a cDNA encoding (S)-cis-N-methylstylopine 14-hydroxylase from opium poppy, a key enzyme in sanguinarine biosynthesis. Biochem Biophys Res Commun 2009:597–603. doi: https://doi.org/10.1016/j.bbrc.2012.12.120.  

[75] Besseau S, Fellner F, Lanoue A, Thamm AMK, Salim V, Schneider B, et al. A Pair of Tabersonine 16-Hydroxylases Initiates the Synthesis of Vindoline in an Organ-Dependent Manner in Catharanthus roseus. Plant Physiol 2013:1792–803. doi: https://doi.org/10.1104/pp.113.222828.  

[76] Neag MA, Mocan A, Echeverría J, Nandi OI, Facchini PJ. Evidence for the monophyletic evolution of benzylisoquinoline alkaloid biosynthesis in angiosperms. Phytochemistry 2005:1374–93. doi: https://doi.org/10.1016/j.phytochem.2005.04.029.