Ecometabolomic Analysis of Wild Populations of *Pilocarpus pennatifolius* (Rutaceae) Using Unimodal Analyses

Daniella M. Allevato[1]*, Eduardo Kiyota[2], Paulo Mazzafera[2,3] and Kevin C. Nixon[1]

[1] Section of Plant Biology, School of Integrative Plant Science, Cornell University, Ithaca, NY, United States, [2] Departamento de Biologia Vegetal, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, Brazil, [3] Departamento de Produção Vegetal, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, Brazil

Studies examining the diversity of plant specialized metabolites suggest that biotic and abiotic pressures greatly influence the qualitative and quantitative diversity found in a species. Large geographic distributions expose a species to a great variety of environmental pressures, thus providing an enormous opportunity for expression of environmental plasticity. *Pilocarpus*, a neotropical genus of Rutaceae, is rich in alkaloids, terpenoids, and coumarins, and is the only commercial source of the alkaloid pilocarpine for the treatment of glaucoma. Overharvesting of species in this genus for pilocarpine, has threatened natural populations of the species. The aim of this research was to understand how adaptation to environmental variation shapes the metabolome in multiple populations of the widespread species *Pilocarpus pennatifolius*. LCMS data from alkaloid and phenolic extracts of leaf tissue were analyzed with environmental predictors using unimodal unconstrained and constrained ordination methods for an untargeted metabolomics analysis. PLS-DA was used to further confirm the chemoecotypes of each site. The most important variables contributing to the alkaloid variation between the sites: mean temperature of wettest quarter, as well as the soil content of phosphorus, magnesium, and base saturation (V%). The most important contributing to the phenolic variation between the sites: mean temperature of the wettest quarter, temperature seasonality, calcium and soil electrical conductivity. This research will have broad implications in a variety of areas including biocontrol for pests, environmental and ecological plant physiology, and strategies for species conservation maximizing phytochemical diversity.

Keywords: ecological metabolomics, unimodal ordination, *Pilocarpus pennatifolius*, chemoecotype, ecophysiology, canonical correlation analyses

INTRODUCTION

The vast array of compounds synthesized in plants has led to much debate over the function or adaptive significance for such a diversity of compounds. When plant specialized metabolites (PSM) were first discussed, they were often seen as waste products from primary metabolism, accidents or aberrant metabolism (Haslam, 1986). Today, PSMs are believed to have multiple functions and

have been shown to be as important as primary metabolites in many cases (Wink, 1999; Cipollini, 2000; Izhaki, 2002).

The qualitative and quantitative PSM diversity in plants appears to change over time (phenology) and space, encompassing many ecological roles (Moore et al., 2014). Studies examining the diversity of PSM suggest that in the environment, biotic and abiotic pressures greatly influence the qualitative and quantitative diversity of the metabolome found in a species. The Climatic Variability hypothesis suggests that locations with constant warm temperatures and little seasonal variation would lead to a lower capacity for environmental plasticity (Molina-Montenegro and Naya, 2012). Large geographic distributions expose a species to a great variety of environmental pressures, thus providing an opportunity for an enormous range of chemoecotypes (Sultan, 1987; Tack et al., 2012; Gratani, 2014).

Pilocarpus, a genus of Rutaceae in the sub-tribe Pilocarpine, is known for its high bioactivity due to the presence of many alkaloids, terpenoids and coumarins (Santos and Moreno, 2004). Specifically, the genus has been sought after for one specific imidazole alkaloid, pilocarpine, as it is the only source of this compound that is used to treat glaucoma, as well as xerostomia. The genus *Pilocarpus* is the only known genus that contains this compound, and many of the species in this genus have become threatened with extinction due to overharvesting of wild populations (Pinheiro, 1997, 2002). Besides two chemical variants found in the genus *Casimoroa* (Rutaceae), Imidazole alkaloids have not been found in any other genus of plants. The genus *Pilocarpus* comprises 16 species and has a widespread, but strictly Neotropical distribution from southern Mexico and the Antilles, to northern Argentina (Skorupa, 1996).

Previous studies on other species of *Pilocarpus* have found that abiotic factors such as environmental variation can affect compound expression (Abreu et al., 2007; Sawaya et al., 2011). One study examining imidazole alkaloid concentrations in leaf tissue in *Pilocarpus microphyllus* plantations in Southern Brazil has found that there is variation of compound concentrations and compound presence depending on the season. The exposure to seasonality in the South is very different from the uniform seasons where *P. microphyllus* is naturally found, in the tropical rainforest and terra firme regions of the Amazonian state of Para. The seasonal variation of alkaloid production in *P. microphyllus* has demonstrated a strong association between specific groups of compounds and seasons, and has led to the hypothesis that distinct biosynthetic pathways are active in different seasons (Abreu et al., 2007; Sawaya et al., 2011). In addition, studies looking at salt stress, wounding, and hypoxia in *Pilocarpus* reported significant reductions in certain compounds following exposure to these conditions, further supporting the plasticity of certain genotypes to environmental stresses (Avancini et al., 2003). Further studies using *P. microphyllus* calluses and seedlings showed variation in pilocarpine production with pH variation, absence/excess of nutrients N, K, P, and NaCl, and amino acid precursors histidine and threonine (Avancini et al., 2003; Abreu et al., 2005; Andreazza et al., 2009).

One species in the genus, *Pilocarpus pennatifolius*, has a large latitudinal distribution as well as some of the greatest diversity of compounds present in the genus (Santos and Moreno, 2004; Sawaya et al., 2011). As *P. pennatifolius* has one of the largest distributions, one would expect that there would be a greater variation in chemistry since it is exposed to a larger variety of environmental factors. Using *P. pennatifolius* as an example of a species with a widespread latitudinal distribution, we were interested in determining which environmental factors are associated with the greatest variation among metabolomes of *P. pennatifolius* populations (**Supplementary Figure 1**). We hypothesize that there are patterns of chemistry, or chemoecotypes, correlated with varying environmental conditions among wild populations of *P. pennatifolius*.

Previous plant comparative metabolomics studies have utilized linear methods to investigate correlations between PSM and the environment (Kaplan et al., 2004; Arbona et al., 2013; Lankadurai et al., 2013; Sampaio et al., 2016). However, with the larger environmental gradient present in the wild populations of this study, the response of chemical variation in these situations follows a non-linear and unimodal distribution (Whittaker, 1956). In situations of smaller environmental gradients, a linear response could approximate the conditions well; however, when there is a unimodal response we expect that the linear model will not suffice (Šmilauer and Lepš, 2014). PSM production in plants can display a linear or unimodal response when faced with different abiotic factors. A unimodal response can display maximum compound abundance at an optima with low compound production at the extremes (e.g., no nutrient or excessive nutrient), or there can be maximum compound abundance at the extremes when a plant faces major environmental stressors (**Supplementary Figure 2**; Yan et al., 2004; Pennycooke et al., 2005; Albert et al., 2009; Ncube et al., 2012; Šmilauer and Lepš, 2014). Although unimodal methods can approximate linear responses well, the reverse is not true with linear methods. Therefore, in the case of our study with a large environmental gradient and a unimodal response of traits, only the unimodal method can accurately assess the response. Multivariate methods can address the issues related to unimodal distributions and have been used extensively in microbial ecology and community ecology, but they have not previously been applied to plant environmental metabolomics (Braak, 1986; Harborne and Jeffrey, 1993; Sardans et al., 2011; Paliy and Shankar, 2016).

The objective of our study was to test the hypothesis that there are correlations between chemotypes and environmental conditions for wild populations of *P. pennatifolius* using multivariate methods to analyze leaf metabolomics data. This included an unconstrained ordination analysis to assess overall variation in the metabolome among sites, a constrained ordination to assess metabolomic variation attributed to environmental predictors, and discriminant analysis to assess the group classifications for chemoecotypes of the various wild populations. To avoid inaccurate estimates from linear analysis, in this study we have used the following unimodal analyses, Correspondence Analysis (CA) (Hill, 1973, 1974) and Canonical Correspondence Analysis (CCA), to accurately assess variation in the metabolomes of wild populations of *P. pennatifolius*.

MATERIALS AND METHODS

Field Collections of Plant Material and Plant Sampling Design

To examine variation in alkaloid and phenolic profiles due to environmental variation, this study focused on *P. pennatifolius*. This species was chosen because it has the largest latitudinal range (greater variation of climatic variables) that is both current and historic, and it is not threatened with extinction in natural populations. Using bioclimatic data from the WorldClim website[1], distribution data from 2189 Brazilian herbarium specimens from the SpeciesLink website[2], and the Ecocrop function in DIVA-GIS, localities with the greatest disparity of climatic variables were chosen for collection sites while planning our fieldwork itinerary (Dell et al., 2014).

Field collected plants were sampled from six locations for *P. pennatifolius*, obtaining as many individuals as possible from each site (**Table 1**). Specimens were collected from Parque Estadual Lago Azul, Campo Mourão Site 1, Campo Mourão Site 2, Foz do Iguaçu, Estação Ecológica St. Tereza, and Cruz do Pedro. Herbarium specimens were made for each site and can be accessed at the Bailey Hortorium Herbarium (BH) and USP Ribeirão Preto Herbarium (SPFR) (**Supplementary Table 1** and **Supplementary Figure 1**). Multiple leaves were collected from each individual. Specifically, the third leaflet down from the terminal leaflet of the compound leaves was chosen; each sample collection was obtained from a different compound leaf each time.

Soil Sample Field Collections and Analysis

At each collection locality, two soil samples were collected at 20 cm depth (van Rai, 2001; Camargo et al., 2009) and soil analysis was carried out at the Instituto Agronômico in Campinas, SP, Brazil following established methods (van Rai, 2001). Fe, Mn, Zn, P. K, Ca, Mg, S, B, Na, Al, N, organic matter, pH, soil electrical conductivity (EC), Base saturation (V%), cation exchange capacity (CEC) were determined in the soil samples (**Supplementary Table 2**).

Bioclimatic Data Retrieval

Bioclimatic variables (BIO1-BIO19) at 30-arc seconds were extracted from the WorldClim website (see text footnote 1) using the longitude and latitude coordinates of each collection locality. In addition, altitude data was extracted using coordinates for each collection on the program DIVA-GIS (**Supplementary Table 3**).

Biochemical Extraction of Leaf Tissues for LCMS

To validate biochemical extractions of alkaloids and coumarins from field collected leaf tissue, a degradation study on two live specimens of *P. pennatifolius* from the New York Botanic Garden was conducted. Silica-dried leaf tissue on ice showed the

[1] www.worldclim.org

[2] http://splink.cria.org.br

TABLE 1 | Summary of *P. pennatifolius* population sites and final set of environmental variables.

Sites in Brazil	*	Geographic coordinates	#	TS	MTWQ	MTCM	P (mg/dm3)	Mg (mmolc/dm3)	Ca (mmolc/dm3)	V% (%)	EC (dS/m)	H+Al (mmolc/dm3)
A. Foz do Iguaçu, PR		−25°26'48.98" −54°35'6.45"	3	359.78	23.73	23.73	37	27	201	94	1.3	16
B. Campo Mourao A, PR		−24°2'01.4" −52°20'50.7"	5	291.01	23.5	16.5	100.5	6.5	83.5	39.5	1.1	139
C. Campo Mourao B, PR		−24°2'00.7" −52°20'50.7"	4	291.01	23.5	16.5	66	38.5	362	92.5	0.9	30
D. Parque Lago Azul, PR		−24°06'05.8" −52°18'42.8"	9	290.26	23.35	16.38	57	35	328	93	1.5	29
E. Cruz do Pedro, SP		−21°17'42.27 −47°53'56.46"	5	204.70	23.13	18.33	32	18	73	76	0.7	29
F. Estação Ecológica St. Tereza, SP		−21°13'12" −47°50'54"	5	203.01	23.23	18.46	186	36	246	96	2.2	11

*P, Mg, Ca, V% (base saturation), EC (soil electrical conductivity), and H+Al were determined by soil analysis at each site: #Refers to number of individuals. Other bioclimatic and soil variables measured are in **Supplementary Tables 2, 3**. TS, temp seasonality; MTWQ, Mean Temp Wettest Qtr; MTCM, Mean Temp Coldest Month. *Colors refer to color-coding for **Figures 1, 3** only.*

least degradation and was not significantly different from the immediately frozen tissue. This method of preservation of leaf tissue for biochemical extractions has been used in many studies that require plant collections in remote areas, as the removal of water reduces enzymatic activity in the leaf tissue and the cold temperature helps to prevent enzymatic reactions due to handling (Fine et al., 2013).

Biochemical extractions to elucidate alkaloids and phenolics present in *Pilocarpus* leaf tissue were performed on silica-dried leaf tissue, at UNICAMP in Brazil. Alkaloid extraction protocols followed modifications made by Dr. Sawaya (Dr. Mazzafera Lab) on the method developed by Avancini et al. (2003). Alkaloid extracts were done on approximately 10 mg leaf tissue (normalized by exact weight in LCMS data prep). Leaf tissue was extracted with 1 ml of 0.1% Formic acid and placed in an ultrasonic water bath for 20 min. Supernatant was removed and 1 ml of 0.1% Formic acid step was redone twice, filtered, and analyzed. The phenolic extraction protocol followed the methods developed by Vialart et al. (2012) and Durand-Hulak et al. (2015). These phenolic extracts were done on approximately 200 mg tissue (normalized by exact weight in LCMS data prep). Tissue was homogenized in 2 ml 80% ethanol for 1 min, centrifuged for 10 min, and the supernatant removed to a fresh tube. Each individual plant had a total of three biological replicates for each extraction.

UPLC-ESI-MS Methods
Alkaloids
Five microliters of each sample were injected into an Acquity UPLC coupled with a TQD triple-quadrupole mass spectrometer (Micromass-Waters, Manchester, United Kingdom). Mass spectrometer conditions were: capillary 3.0 kV, cone 30 V, extractor 1V, ion source temperature 150°C, desolvation temperature 300°C and column temperature 30°C, in electrospray ionization in positive mode. The chromatographic column used was a Polaris 3 C18-A 100 mm × 2.0 mm column (Varian) and the elution was carried out using an ammonium acetate buffer, 10 mmol/L, pH 3.0 (solvent A) and acetonitrile (solvent B) in a gradient ranging from 5 to 25% of solvent B in 8 min. The retention time and m/z were used to identify alkaloid compounds in the samples.

Phenolics
Sample extracts were diluted in 80% ethanol (ethanol and water) 1:4. Four microliters of each sample were injected into an Acquity UPLC coupled with a TQD triple-quadrupole mass spectrometer (Micromass-Waters, Manchester, United Kingdom). Mass spectrometer conditions were: capillary 3.0 kV, cone 35 V, extractor 1V, ion source temperature 120°C, desolvation temperature 350°C and column temperature 30°C, in electrospray ionization in positive mode. The chromatographic column used was a Waters ACQUITY C18-BEH (2.1 mm × 100 mm, 1.7 μm) column and the elution was carried out using Milli-Q water with 0.1% formic acid (solvent A) and acetonitrile (solvent B) in a gradient ranging from 10 to 100% of solvent B in 8 min. The retention time and m/z were used to identify coumarin compounds in the samples

using a SIM (single ion monitoring) mode. For the metabolomics analysis samples were run in TIC (total ion chromatogram) mode for all phenolics.

LCMS Data Processing
Data processing was conducted using LCMS raw data in MZmine2 (Pluskal et al., 2010) and XCMS in R (Smith et al., 2006). Waters raw LCMS data were converted into mzXML format using the MSConvert package in Proteowizard v.3.0.10051 (Chambers et al., 2012). The raw data in mzXML format was then processed in MZmine2 following protocols described by Earll (2012): mass detection (m/z tolerance of 0.002) and filtering, chromatogram builder, peak deconvolution using Wavelets XCMS algorithm, peak alignment of all the LCMS runs, adduct search, and gap filling (Earll, 2012). The raw data was also analyzed using XCMS in R, this included peak picking, integration, and non-linear retention time correction of peaks between samples. LCMS data of all the metabolite features were exported, including EIC (extracted ion chromatograms) for manual analysis and a csv file of the processed samples containing: peak m/z, peak RT, and peak intensity areas for multivariate analyses. Some compounds were identified via standard references as well as comparison with literature (Sawaya et al., 2011; Hiserodt and Chen, 2012; Tine et al., 2017; **Supplementary Tables 4A,B**). All metabolite features used in the analysis are in **Supplementary Table 5A** (alkaloid extract) and **Supplementary Table 5B** (phenolic extract). LCMS data was normalized by dry weight of leaf tissue, log transformed, and pareto scaled to analyze the metabolome of individuals from these six sites.

LCMS Multivariate Analysis
Unsupervised multivariate methods, CA and HCA were employed in CANOCO 5. CA was used to reduce the dimensionality of the LCMS data. Hierarchical cluster analysis (HCA) was utilized to create a dendrogram, utilizing Ward's method with Euclidean distances to confirm affinities from CA.

Constrained ordination is not possible with 39 environmental predictors, as the result will be unconstrained; therefore, the 39 environmental predictors were reduced to eleven environmental predictors, to reduce multicollinearity and the arch effect due to many environmental variables (Draper and Smith, 1998; Šmilauer and Lepš, 2014). Factor analysis of the environmental variables revealed five components for the 39 environmental predictors. These complex variables were made up of many environmental predictors and the variables did not load clearly; consequently, the interpretability would be lost if these complex variables were used in the constrained ordination (Draper and Smith, 1998; Šmilauer and Lepš, 2014). Evaluation of the sites on Canoco software revealed environmental variables with low contributions/arrows, which were removed first from the data set. Forward stepwise analysis was performed on Canoco software using a semi-automated procedure, determining the eleven most informative variables that explained the residual variation in chemical traits. This stepwise analysis also analyzed correlated variables, assessing the difference in eigenvalues of species-environmental correlations to determine if wrong variables or too many variables were removed from correlated sets. The final set

of included variables is presented in **Table 1**; however, for both analyses only four variables had significant contributions to the chemical variation.

Canonical Correspondence Analysis of environmental predictors was used to assess variation of metabolic profiles, associated with environmental variables. The three environmental variables with greatest significant variation affecting profiles were depicted on CCA.

Supervised multivariate analyses such as Partial Least Squares Discriminant Analysis (PLS-DA) using the R pls package through MetaboAnalyst, and Random Forest analysis using the randomForest package in R through MetaboAnalyst, were used to confirm classification of chemotype groups (Xia and Wishart, 2002). PLS-DA is excellent at managing both noisy data and multicollinearity, and it was used to identify significant variables in chemotypes through the determination of the regression coefficient, loading plots and variable importance on projection (VIP). Random Forest Analysis was also used for class prediction and provides error rates for each class as well as outlier measures, OOB error, and variable importance measures.

Validation of Multivariate Analyses

The significance of the class discriminations described by PLS-DA models were statistically validated using performance accuracy measures $R2$ (sum of squares captured by the model) and $Q2$ (cross-validated $R2$), cross-validation with different numbers of components and a permutation test using the optimal number of components from the cross-validation.

RESULTS

Pilocarpus pennatifolius Sample Sites and Environmental Variables

The choice of environmental variables to use in the model can affect the analysis; therefore, forward selection of environmental predictors in Canoco was used to reduce the environmental variables for the analysis (**Table 1**). Leaf tissue underwent two separate extraction protocols, an alkaloid and a phenolic extraction, and was normalized by dry weight of leaf tissue and log transformed, to analyze the metabolome of individuals from these six sites. All identified and unidentified peaks were used for the metabolomic analyses (**Supplementary Tables 5A,B**).

Correspondence Analysis to Assess Overall Variation Among Sites

The relationship between the compound abundance across samples, versus the measured values of the environmental variables was evaluated. As many of our environmental variables encompass a broad range, we analyzed the distribution in SPSS and found that very few variables followed a linear response; therefore, it was preferable to use unimodal analysis methods for our analysis (Šmilauer and Lepš, 2014). For our indirect gradient analysis we used the unimodal method of correspondence analysis.

Correspondence analysis, also known as reciprocal averaging, is based on the weighted averaging of compound scores to estimate the latent variables that best predict values for all compounds in a model (Lepš and Šmilauer, 2003). CA of metabolomics data from the populations of the six sites depicts about four major groups of variation in the alkaloid extract (**Figure 1A**) and about three major groups of variation in the phenolic extract (**Figure 1C**). The explained variation of each axis of the CA ordination is shown in **Table 2**. The first few axes of CA are usually attributed to environmental variables; however, as this is a secondary comparison to environmental gradients, it is an indirect gradient analysis and is viewed as the overall chemical variation between samples. When each compound is graphed on the CA diagram we can see how each grouping of samples is differentiated by a greater abundance of select compounds in the chemical profile (**Supplementary Figures 3A,B**).

Canonical Correspondence Analysis to Assess Variation Due to Environmental Factors

To directly test the environmental gradient, CCA, a multivariate extension of CA was applied. Here the axes are restricted or constrained to be linear combinations of environmental variables, and therefore the axes represent those variables that present the maximum separation or variation in the samples. The total environmental variables collected for each site consisted of 19 bioclimatic variables from the WorldClim website (see text footnote 1) and 20 soil variables analyzed from soil samples at each site. Using all 39 environmental variables for the CCA would create an ordination diagram that appears more comparable to an unconstrained CA than to a constrained CCA. Therefore, after removing collinear variables we performed a forward selection analysis of explanatory variables in CANOCO, to test which variables would improve the fit (**Tables 1, 3**). The CCA ordination diagram of the alkaloid extraction depicts the four most important variables contributing to the chemical variation between the sites as: mean temperature of wettest quarter as well as the content of phosphorus, magnesium, and V% – base saturation in the soil (**Figure 1B**). The CCA ordination diagram of the phenolic extraction depicts the four most important variables contributing to this chemical variation between the sites as: mean temperature of the wettest quarter, temperature seasonality, calcium, and soil EC (**Figure 1D**). The explained variation of each axis of the CCA ordination diagrams is shown in **Table 2**. Next, we tested the constrained axes to verify the significance of the axes: the first and second axes were significant for the alkaloid extract, whereas only the first axis was significant for the phenolic extracts (**Table 4**).

Canonical Correspondence Analysis ordination diagrams can also be plotted as a biplot of the response variables (compounds) and the explanatory variables (environmental variables). The compounds-environmental variables CCA biplot of the entire metabolome (**Supplementary Figure 4A**: Conceptual Diagram, **Supplementary Figure 4B**: Alkaloid extraction, and **Supplementary Figure 4C**: Coumarin extraction) depicts how different groups of the metabolome are affected by
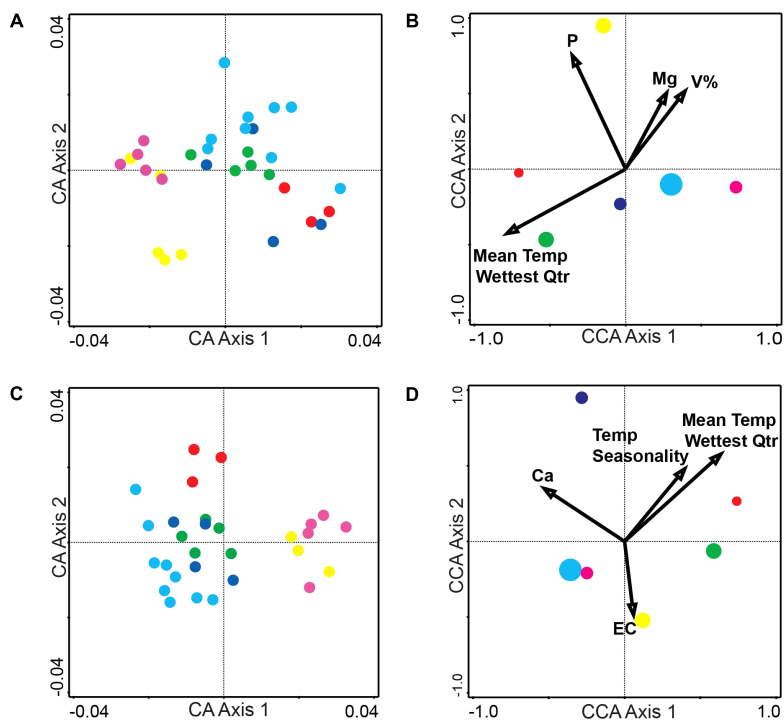
FIGURE 1 | Correspondence Analysis (CA) and Canonical Correspondence Analysis (CCA) for the alkaloid and phenolic extractions. **(A)** CA ordination plot of alkaloid extraction at the six sites with each dot representing an individual plant (average of three biological replicates) at a site **(B)** CCA ordination biplot of alkaloid extractions displaying environmental variables (arrows) and each site (circle with size representing the number of individuals collected) **(C)** CA ordination plot of phenolic extraction at the six sites with each dot representing an individual plant (average of three biological replicates) at a site **(D)** CCA ordination biplot of phenolic extraction displaying environmental variables (arrows) and each site (circle with size representing the number of individuals collected). Number of individuals collected and site location (A–F) is also noted in **Table 1**. Numerical results of CA and CCA are in **Table 2**. Site A is Red, Site B is Green, Site C is Dark Blue, Site D is Light Blue, Site E is Magenta, and Site F is Yellow.

TABLE 2 | Explained variation from axes in the CA and CCA ordination diagrams.

	Axis 1	Axis 2	Axis 3	Axis 4
Alkaloid extraction				
Unconstrained (CA) explained variation (cumulative)	30.95	47.65	58.04	65.61
Constrained (CCA) explained variation (cumulative)	11.23	18.14	21.41	23.33
Pseudo-canonical correlation	0.629	0.715	0.658	0.586
Phenolic extraction				
Unconstrained (CA) explained variation (cumulative)	29.09	41.28	50.48	58.84
Constrained (CCA) explained variation (cumulative)	13.98	20.47	24.70	27.83
Pseudo-canonical correlation	0.751	0.845	0.746	0.649

TABLE 3 | Interactive forward selection of environmental variables.

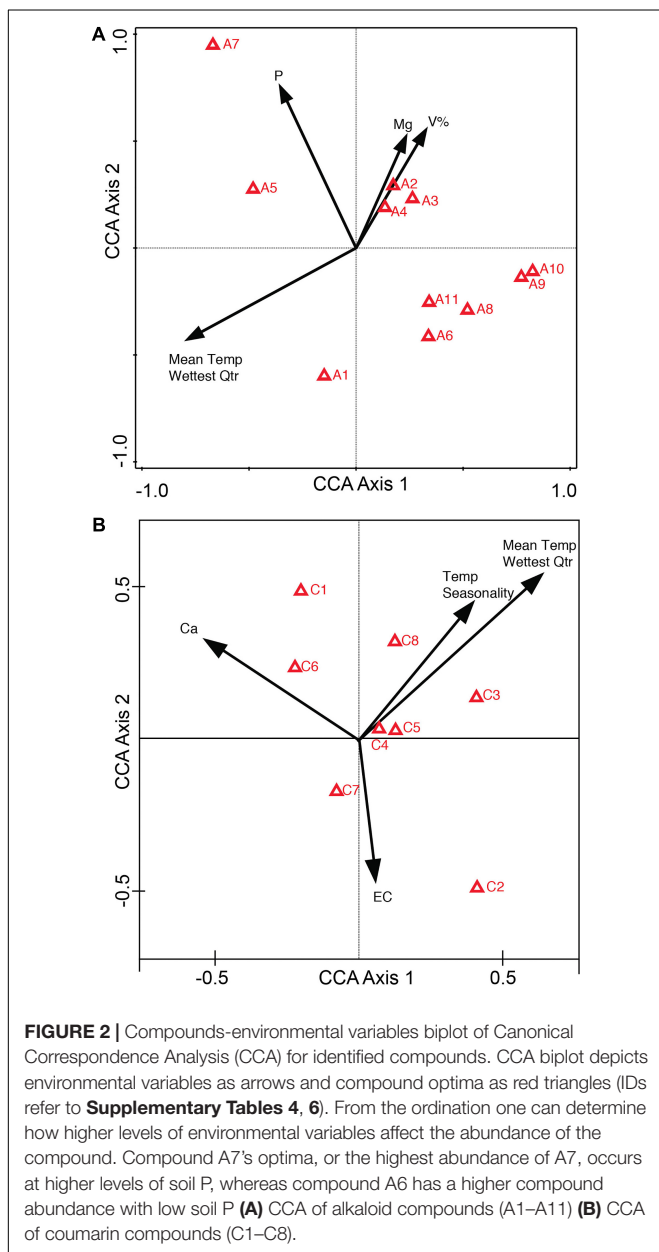|	Explains %	Contribution %	*P*-value
Alkaloid extraction			
Mean Temp Wettest Qtr.	11.1	35.4	0.008
P	8.9	28.4	0.018
V% (Base Saturation)	4.5	14.4	0.377
Mg	3.5	11.1	0.664
Phenolic extraction			
Mean Temp Wettest Qtr.	8.5	28.6	0.032
Ca	7.4	24.8	0.073
EC (Soil Electrical Conductivity)	6.0	20.1	0.167
Temp seasonality	4.7	15.8	0.397

higher and lower levels of the environmental variables present across the collection sites. Next, the compound optima (red triangles) in the compounds-environmental CCA biplot (**Supplementary Figure 4**) were reduced to depict only the identified compounds. This was done to analyze how higher/lower levels of environmental variables across the sites affect the abundance of these specific compounds (**Figure 2A**: Alkaloids, **Figure 2B**: Coumarins). **Supplementary Table 6** is

provided to summarize the levels of environmental variables that have been associated with abundance optima of the following identified compounds.

In the Alkaloid CCA biplot (**Figure 2A**), compounds A7 [3-Hydroxymethyl-4-(3-methyl-3H-imidazol-4-yl)-1-phenylbutan-1-one] and A6 [3-(Hydroxy-phenyl-methyl)-4-(3H-imidazol-4-ylmethyl)-dihydro-furan-2-one] illustrate high compound abundance at opposite levels of environmental variables. A7 has a higher abundance in the sites with higher soil P, Mg,

TABLE 4 | Significance of constrained axes of CCA diagrams.

	Axis 1	Axis 2	Axis 3	Axis 4
Alkaloid extraction				
Explained by constrained axis	14.71%	8.60%	3.06%	1.53%
P-value	0.007	0.03	0.68	0.894
Phenolic extraction				
Explained by constrained axis	13.95%	6.48%	4.21%	1.86%
P-value	0.013	0.163	0.4	0.867

FIGURE 2 | Compounds-environmental variables biplot of Canonical Correspondence Analysis (CCA) for identified compounds. CCA biplot depicts environmental variables as arrows and compound optima as red triangles (IDs refer to **Supplementary Tables 4**, **6**). From the ordination one can determine how higher levels of environmental variables affect the abundance of the compound. Compound A7's optima, or the highest abundance of A7, occurs at higher levels of soil P, whereas compound A6 has a higher compound abundance with low soil P **(A)** CCA of alkaloid compounds (A1–A11) **(B)** CCA of coumarin compounds (C1–C8).

optimum of pilocarpine (A4) appears to be associated with the same environmental variables as the optima of two other identified compounds: A3 (pilocarpidine) and A2 [4-(3H-imidazol-4-ylmethyl)-3-methyl-5H-furan-2-one]. These three compounds have greater compound abundances when present at sites with high levels of soil Mg and V% (Base Saturation), mid-to-high levels of soil P, and low Mean Temperature in the Wettest Quarter. The CCA biplot of the coumarin compounds (**Figure 2B**) shows less indication of groupings. C1 (scopoletin) and C2 (psoralen) appear to have the most distinct response to environmental variables. C1 has its highest abundance at sites with high soil Ca, high temperature seasonality, high Mean Temperature in Wettest Quarter, and low soil EC. However, the opposite is true for C2, where the highest abundance of C2 is associated with the sites that have low soil Ca, low Temperature Seasonality, low Mean Temperature in the Wettest Quarter and high soil EC.

Classification of Chemoecotypes

To further confirm the classification of chemoecotypes by site we used Partial Least Squares-Discriminant Analysis and Random Forests Analysis. PLS-DA was used on the log-transformed LCMS data. Possible hypotheses for the chemical variation at sites include: no variation between sites leading to one chemotype for all sites (**Supplementary Figure 5A**), visible groupings or chemoecotypes identified per site (**Supplementary Figure 5B**), or random chemical variation within and among sites (**Supplementary Figure 5C**). For both alkaloid and phenolic extractions we were able to confirm group classifications/chemoecotypes for each site. For the alkaloid extraction, we note a clear grouping of individuals from each site in the PLS-DA 3-D scores plot, though we note that Site C has 1 outlier (**Figure 3A**). In the phenolic extraction, we also see a grouping of individuals from each site on the PLS-DA 3-D scores plot (**Figure 3**). PLS-DA 3-D scores plot of all biological replicates at all sites was also completed, and a dense grouping at each site is noted (**Supplementary Figures 6A,B**). Statistical validations of the class separations of PLS-DA were done through the cross-validation of different numbers of components and permutation tests of randomly assigned class labels using the optimal number of components (**Supplementary Figure 7** and **Supplementary Table 7**).

Random Forest Analysis was also utilized for class prediction. In **Figures 4A,B** we see that the classification of sites based on the alkaloid extraction (**Figure 4A**) is more defined than the phenolic extraction (**Figure 4B**). This is also confirmed by the OOB error and error rates for each class seen in **Table 5**. The variation in certain groups or sites could be due to sample size; however, it is not possible to acquire many individuals of this species at every location.

DISCUSSION

Just as it is possible to determine species growing in an area based on environmental conditions present at a location (Braak, 1986; Harbert and Nixon, 2015), it is also possible
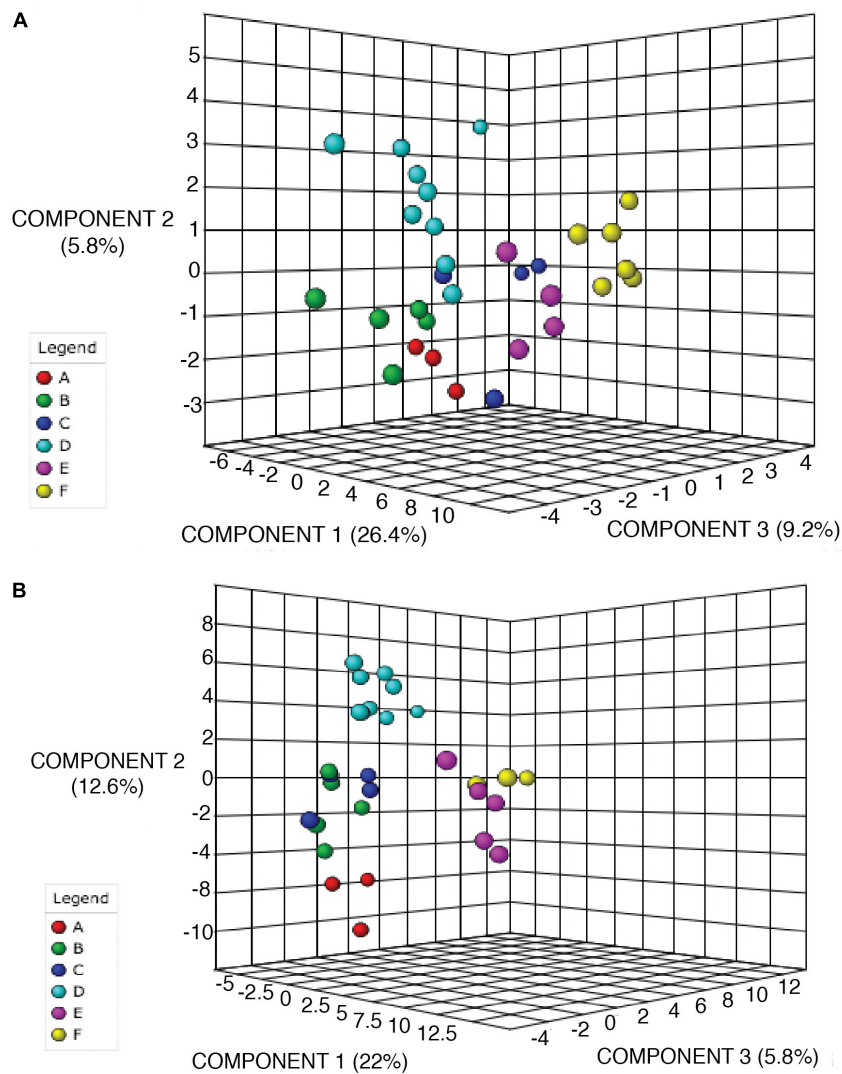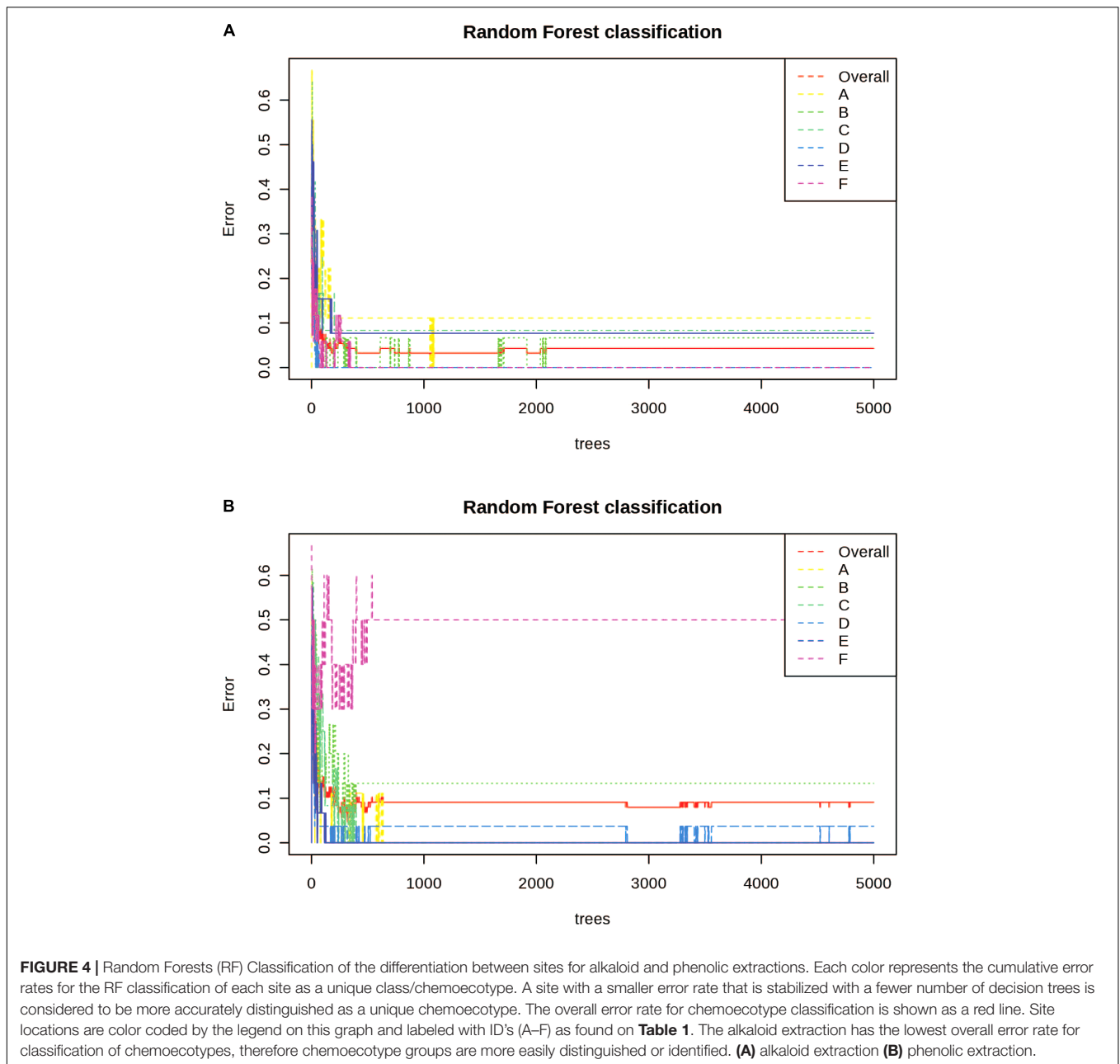
V%, as well as mid-to-high values of Mean Temperature in the Wettest Quarter. A6 on the other hand has a higher abundance in the sites with lower soil P, Mg, V% as well as lower values of Mean Temperature in the Wettest Quarter. The abundance

FIGURE 3 | PLS-DA of alkaloid and phenolic extractions depicts differentiation between sites. 3D scores plot between selected PCs. The explained variances are shown in brackets. Site locations (A–F) are color-coded and ID's refer to **Table 1**. **(A)** PLS-DA for alkaloid extraction **(B)** PLS-DA for phenolic extraction.

to predict which compounds or compound profiles will be more abundant at environmental optima through the use of direct gradient analyses. Metabolomic expression across the distribution of a plant species can follow a non-linear, non-monotonic relationship with environmental variables, and this can be verified in situations of large environmental gradients (Whittaker, 1956). To accurately analyze non-linear responses, unimodal methods of ordination diagrams are preferred.

In our study, we analyzed the metabolite profiles of *P. pennatifolius* populations using two extraction protocols, alkaloid and phenolic. For the alkaloid extract, we found that the most significant environmental variables contributing to the variation in alkaloid profiles were mean temperature of the wettest quarter, as well as the soil content of phosphorus, magnesium, and V% (base saturation). Temperature and

precipitation are known to affect alkaloid content in plants, and thus variations in these bioclimatic factors could provide variations in chemical profiles (Waller, 1978; Harborne and Jeffrey, 1993). It is usually expected that nitrogen availability has a strong effect on alkaloid production as alkaloids are N containing compounds, but this is not always true. Previous stress studies irrigating nutrient solutions on *P. microphyllus* seedlings found that N omission in nutrient solutions did not affect pilocarpine production as much as K omission, as is further shown in our study (Avancini et al., 2003). This low response could be due to a reduced availability of nitrogen in an accessible form or other limiting nutrients that are needed in conjunction with nitrogen for a balanced nutrition. Base saturation (V%) indicates the amount of basic cations in the soil, and is essential for determining soil fertility as it specifies the percentage of

FIGURE 4 | Random Forests (RF) Classification of the differentiation between sites for alkaloid and phenolic extractions. Each color represents the cumulative error rates for the RF classification of each site as a unique class/chemoecotype. A site with a smaller error rate that is stabilized with a fewer number of decision trees is considered to be more accurately distinguished as a unique chemoecotype. The overall error rate for chemoecotype classification is shown as a red line. Site locations are color coded by the legend on this graph and labeled with ID's (A–F) as found on **Table 1**. The alkaloid extraction has the lowest overall error rate for classification of chemoecotypes, therefore chemoecotype groups are more easily distinguished or identified. **(A)** alkaloid extraction **(B)** phenolic extraction.

sodium, calcium, potassium, and magnesium that is a part of the CEC. Sites A and B have high CEC values (250.50 and 232.05); however, their respective V% values (94 and 39.5%) demonstrate the great difference in fertility of these soils. The poor soil fertility of Site B is also verified by the lower pH and higher H+Al values. Minerals such as phosphorus and magnesium are very important in various enzymatic processes throughout the plant, and they have large roles in promoting plant development and growth. Studies examining the effect of increases in phosphorous and magnesium have had contradictory results in regards to the effect on alkaloids, ranging from no effect to both positive and negative effects (Waller, 1978; Bramble and Graves, 1991; Mazzafera, 1999). A study that examined the effect of Phosphorus

omission on *P. microphyllus* found no effect on the pilocarpine content; however, the authors did not look at the effect on the overall alkaloid profile (Avancini et al., 2003; Abreu et al., 2005).

The most significant environmental variables contributing to the variation in the phenolic profiles in the phenolic extracts were mean temperature in the wettest quarter, temperature seasonality, as well as the soil content of calcium and the soil EC. Temperature seasonality has been known to have a strong effect on compound expression, since areas of greater seasonality are faced with more variation and could lead to greater variation in chemistry (Molina-Montenegro and Naya, 2012). Specifically, work on the synthesis of phenolics has found that they are regulated not only by developmental signals, but also

TABLE 5 | Random Forests classification performance matrix.

	A	B	C	D	E	F	Class error
Alkaloid extraction: OOB Error = 0.0323							
A	9	0	0	0	0	0	0
B	0	14	0	1	0	0	0.0667
C	0	0	11	1	0	0	0.0833
D	0	0	0	27	0	0	0
E	0	0	0	0	12	1	0.0769
F	0	0	0	0	0	17	0
Phenolic extraction: OOB Error = 0.0909							
A	9	0	0	0	0	0	0
B	0	13	0	2	0	0	0.133
C	0	0	12	0	0	0	0
D	0	0	1	26	0	0	0.037
E	0	0	0	0	15	0	0
F	0	1	0	0	4	5	0.5

signals from the environment (Figueiredo et al., 2008; Carbone et al., 2009; Jaakola, 2013). The potential role of calcium is vast as it is essential for membrane permeability, in addition to being a messenger for pathways in development, responses to environmental stimuli and in response to plant defense (Hepler, 2005; Lecourieux et al., 2006; Zhang et al., 2014). In fact, various studies have shown that calcium treatment has led to an increase in phenolics, due to the upregulation of important genes in phenolic compound metabolism (Xu et al., 2014; Zhang et al., 2014; Ahmad et al., 2016). On the other hand, it appears that Phenylalanine ammonia-lyase (PAL), the first enzyme of the phenylpropanoid biosynthetic pathway, is negatively affected by high concentrations of calcium; therefore, it is likely that a variation in calcium could affect the phenolic profile in *P. pennatifolius* (Chishaki and Horiguchi, 1997; Tomás-Barberán et al., 1997; Teixeira et al., 2006; Xu et al., 2014). PAL is indispensible for the synthesis of a range of metabolites, and it has been found that PAL activity also varies when faced with environmental stimuli such as thermal stress (Christie et al., 1994; Hunter et al., 1996; Rivero et al., 2001; Payyavula et al., 2012). Soil EC is also important for plant defense and stress management. EC is the ability of the soil to conduct electricity, and it is determined by a variety of factors including soil moisture, clay content, soil temperature and salinity (including salts such as Na, Ca, K, Mg, Cl, SO4, CO3) (Hanlon, 1993). Various studies have found that higher soil conductivity can indicate greater fertility when there is a higher CEC or a higher sum of bases (Ca, Mg, and K) (Fraisse et al., 2001; Officer et al., 2004; Moral et al., 2010). This greater availability of nutrients can increase the fertility of the soil, thereby promoting growth and subsequently metabolites produced, or it could also present a decrease in metabolites as per the growth-defense trade-off (Coley et al., 1985; Herms and Mattson, 1992; Hunter et al., 1996; Fine et al., 2006; Endara and Coley, 2011). In this study the sum of bases is highly positively correlated with CEC (Pearson's $r = 0.94$); however, EC has a moderately positive correlation with CEC (Pearson's $r = 0.28$) and EC has a moderately positive correlation with the sum of bases (Pearson's $r = 0.34$). As the correlation of EC and CEC is weak, it is possible that other factors are affecting the value

of EC in these soil samples. These factors could be related to soil moisture/temperature or the clay soil type present, which in combination can affect the soil microbiome, thus affecting the phenolic profile present (Sudduth et al., 2003; Badri et al., 2013). Further work is required to assess soil texture and water drainage contributions to the EC values at these sites.

There are many factors that can affect a plant's metabolome and the response of each plant depends on its genetic profile, its environment, as well as the interaction between genotype and the environment. Plasticity is the term used to describe how one genotype can have multiple phenotypes, depending on the environment (Thoday, 1953). Local adaptation depends on both genetic and environmental factors to determine phenotypes with the greatest fitness; however, there is a gradient of importance for each quality (Spichtig and Kawecki, 2004). Several studies have found that areas of stable environmental factors, such as constant temperatures (i.e., the tropics), are correlated with reduced plasticity of PSM (Sultan, 1987; Molina-Montenegro and Naya, 2012; Tack et al., 2012; Gratani, 2014). As plants are immobile, variation in abiotic and biotic interactions leads to an adaptation to the environment, through the modification of a plant's chemical defense. Consequently, stability in the environment would reduce the overall stresses facing the plant, decreasing the selective forces favoring plasticity of the plant to change its metabolite phenotype for survival, which reduces its metabolite variation. There are many instances of a single plant species having different phytochemical profiles in different locations (Hunter et al., 1996; Koricheva, 1999; Andrade-Neto et al., 2002; Hu et al., 2007; Endara and Coley, 2011; O'Reilly-Wapstra et al., 2013; Moore et al., 2014; Moustafa et al., 2016). This could be due to different chemotypes (distinct genetically determined phytochemical profiles that are not evolutionary plastic), plasticity of a single genotype, or gradients of plasticity with genetic variance (Santesson, 1968; Keefover-Ring et al., 2009; Gratani, 2014). Determining the plasticity of a species is difficult without a common garden or transplant experiment.

As our study used only ordination methods to associate environmental variables with variation in chemical profiles of *P. pennatifolius*, and correlation does not imply causation, there are still many laboratory and greenhouse experiments needed, such as seedling stress tests that can be done to further confirm the effect of environmental variables on the chemical profile. In addition, further identification of the compounds found to be significantly different between chemoecotypes through CA, CCA, PLS-DA, and RF could lead to a more targeted analysis and understanding of biological pathways. Common garden experiments, planting different genotypes in the same location, could also be very useful to detect whether differences in chemical variation are indeed due to environmental factors or genetic differences. Either result would be beneficial for breeding as one could breed individuals for specific genes, or one could add/omit nutrients and stressors to modify compound yields. In the case of *P. pennatifolius*, a rare and slow growing tree, we are unable to run a common garden experiment. Therefore this exploratory analysis in wild populations is advantageous and essential as it has reduced the environmental variables, and will allow for a more guided experimental analysis.

CONCLUSION

Analysis of wild populations of *P. pennatifolius* has provided us with potential environmental variables that should be followed up with greenhouse and in-field experiments to determine their importance in alkaloid and phenolic biosynthesis. Of course, the expression of compound profiles in plants is related to a variety of conditions, plant-environment interactions, plant genotype and genotype-environment interactions. Future work in our laboratory will examine population genetic differences among the various sites to assess variation that can be attributed to specific genotypes.

As the field of metabolomics is expanding into larger ecological studies examining wild populations, it is important to determine the best way to assess the response curves of compound profiles present in samples of a large population gradient. As there are non-monotonic and non-linear relationships between compound profiles and environmental variables, the use of unimodal methods (CA, CCA) instead of linear methods (PCA, RDA) will be much more accurate to describe the variation in ecometabolomic studies. As such our study has used exploratory data analysis methods common in community and microbial ecology studies, to perform an ecological metabolomics study of wild populations of *P. pennatifolius*. These unimodal methods can also be used in studies involving analysis of morphology, behavior, pollination, or physiological responses in wild populations across large gradients.

DATA AVAILABILITY

All datasets generated for this study are included in the manuscript and/or the **Supplementary Files**.

AUTHOR CONTRIBUTIONS

DA conceived and designed the project, performed most of the experiments, conducted and interpreted the analysis with contributions from all of the authors, and wrote the article with contributions from all the authors. PM and KN supervised the experiments. EK conducted and performed the LCMS runs. KN supervised and complemented the writing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.00258/full#supplementary-material

REFERENCES

Abreu, I. N., Mazzafera, P., Eberlin, M. N., Zullo, M. A. T., and Sawaya, A. C. H. F. (2007). Characterization of the variation in the imidazole alkaloid profile of *Pilocarpus microphyllus* in different seasons and parts of the plant by electrospray ionization mass spectrometry fingerprinting and identification of novel alkaloids by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* 21, 1205–1213. doi: 10.1002/rcm.2942

Abreu, I. N., Sawaya, A. C. H. F., Eberlin, M. N., and Mazzafera, P. (2005). Production of pilocarpine in callus of jaborandi (*Pilocarpus microphyllus* stapf). *In Vitro Cell. Dev. Biol. Plant* 41, 806–811. doi: 10.1079/IVP2005711

Ahmad P, Abdel Latef, A. A., Abd Allah, E. F., Hashem, A., Sarwat, M., Anjum, N. A., et al. (2016). Calcium and potassium supplementation enhanced growth, osmolyte secondary metabolite production, and enzymatic antioxidant machinery in cadmium-exposed chickpea (Cicer arietinum L.). *Front. Plant Sci.* 7:513. doi: 10.3389/fpls.2016.00513

Albert, A., Sareedenchai, V., Heller, W., Seidlitz, H. K., and Zidorn, C. (2009). Temperature is the key to altitudinal variation of phenolics in arnica montana L. cv. ARBO. *Oecologia* 160, 1–8. doi: 10.1007/s00442-009-1277-1

Andrade-Neto, M., Cunha, U., Mafezoli, J., and Silveira, E. R. (2002). Volatile constituents of different populations of pilocarpus spicatus saint hill. (Rutaceae) from the northeast of brazil. *J. Essent. Oil Res.* 14, 319–324. doi: 10.1080/10412905.2002.9699870

Andreazza, N. L., Abreu, I. N., Sawaya, A. C. H. F., Eberlin, M. N.,and Mazzafera, P. (2009). Production of imidazole alkaloids in cell cultures of jaborandi as affected

by the medium pH. *Biotechnol. Lett.* 31, 607–614. doi: 10.1007/s10529-008-9895-y

Arbona, V., Manzi, M., Ollas, C., and Gómez-Cadenas, A. (2013). Metabolomics as a tool to investigate abiotic stress tolerance in plants. *Int. J. Mol. Sci.* 14, 4885–4911. doi: 10.3390/ijms14034885

Avancini, G., Abreu, I. N., Saldaña, M. D. A., Mohamed, R. S., and Mazzafera, P. (2003). Induction of pilocarpine formation in jaborandi leaves by salicylic acid and methyljasmonate. *Phytochemistry* 63, 171–175. doi: 10.1016/S0031-9422(03)00102-X

Badri, D. V., Zolla, G., Bakker, M. G., Manter, D. K., and Vivanco, J. M. (2013). Potential impact of soil microbiomes on the leaf metabolome and on herbivore feeding behavior. *New Phytol.* 198, 264–273. doi: 10.1111/nph.12124

Braak, C. J. F. T. (1986). Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67, 1167–1179. doi: 10.2307/1938672

Bramble, J., and Graves, D. (1991). Calcium and phosphate effects on growth and alkaloid production in coffea arabica: experimental results and mathematical model. *Biotechnol. Bioeng.* 37, 859–868. doi: 10.1002/bit.260370910

Camargo, O. A., Moniz, A. C., Jorge, J. A., and Valadares, J. M. A. S. (2009). *Métodos de Análise Química, Mineralógica e Física de Solos do Instituto Agronômico de Campinas.* Campinas, SP: Instituto Agronomico.

Carbone, F., Preuss, A., De Vos, R. C., D'Amico, E., Perrotta, G., Bovy, A. G., et al. (2009). Developmental, genetic and environmental factors affect the expression of flavonoid genes, enzymes and metabolites in strawberry fruits. *Plant Cell Environ.* 32, 1117–1131. doi: 10.1111/j.1365-3040.2009.01994.x

Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., et al. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* 30, 918–920. doi: 10.1038/nbt.2377

Chishaki, N., and Horiguchi, T. (1997). Responses of secondary metabolism in plants to nutrient deficiency. *Plant Nutri. Sust. Food Prod. Environ.* 78, 341–345. doi: 10.1007/978-94-009-0047-9_101

Christie, P. J., Alfenito, M. R., and Walbot, V. (1994). Impact of low-temperature stress on general phenylpropanoid and anthocyanin pathways: enhancement of transcript abundance and anthocyanin pigmentation in maize seedlings. *Planta* 194, 541–549. doi: 10.1007/BF00714468

Cipollini, M. L. (2000). Secondary metabolites of vertebrate-dispersed fruits: evidence for adaptive functions. *Rev. Chil. Hist. Nat.* 73, 421–440. doi: 10.4067/S0716-078X2000000300006

Coley, P. D., Bryant, J. P., and Chapin, F. S. (1985). Resource availability and plant antiherbivore defense. *Science* 230, 895–899. doi: 10.1126/science.230.4728.895

Dell, M., Zuccolo, A., Tuna, M., Gianfranceschi, L., and Pè, M. E. (2014). Targeting environmental adaptation in the monocot model Brachypodium distachyon: a multi-faceted approach. *BMC Genomics* 15:801. doi: 10.1186/1471-2164-15-801

Draper, N. R., and Smith, H. (1998). *Applied Regression Analysis.* New York, NY: Wiley-Interscience. doi: 10.1002/9781118625590

Durand-Hulak, M., Dugrand, A., Duval, T., Bidel, L. P., Jay-Allemand, C., Froelicher, Y., et al. (2015). Mapping the genetic and tissular diversity of 64 phenolic compounds in citrus species using a UPLC-MS approach. *Ann. Bot.* 115, 861–877. doi: 10.1093/aob/mcv012

Earll, M. (2012). *MZmine Tutorial.* Basel: Syngenta

Endara, M.-J, and Coley, P. D. (2011). The resource availability hypothesis revisited: a meta-analysis: revisiting the resource availability hypothesis. *Funct. Ecol.* 25, 389–398. doi: 10.1111/j.1365-2435.2010.01803.x

Figueiredo, A. C., Barroso, J. G., Pedro, L. G., and Scheffer, J. J. C. (2008). Factors affecting secondary metabolite production in plants: volatile components and essential oils. *Flavour Fragr. J.* 23, 213–226. doi: 10.1002/ffj.1875

Fine, P. V., Metz, M. R., Lokvam, J., Mesones, I., Zuñiga, J. M., Lamarre, G. P., et al. (2013). Insect herbivores, chemical innovation, and the evolution of habitat specialization in amazonian trees. *Ecol.* 94, 1764–1775. doi: 10.1890/12-1920.1

Fine, P. V., Miller, Z. J., Mesones, I., Irazuzta, S., Appel, H. M., Stevens, M. H., et al. (2006). The growth-defense trade-off and habitat specialization by plants in amazonian forests. *Ecology* 87, S150–S162. doi: 10.1890/0012-9658(2006)87[150:TGTAHS]2.0.CO;2

Fraisse, C. W., Sudduth, K. A., and Kitchen, N. R. (2001). Delineation of site-specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity. *Trans. ASAE* 44, 155–166. doi: 10.13031/2013.2296

Gratani, L. (2014). Plant phenotypic plasticity in response to environmental factors. *Adv. Bot.* 2014, 1–17. doi: 10.1155/2014/208747

Hanlon, E. A. (1993). *Soil pH and Electrical Conductivity: A County Extension Soil Laboratory Manual.* UF/IFAS Extension, University of Florida: Gainesville, FL.

Harbert, R. S., and Nixon, K. C. (2015). Climate reconstruction analysis using coexistence likelihood estimation (CRACLE): a method for the estimation of climate using vegetation. *Am. J. Bot.* 102, 1277–1289. doi: 10.3732/ajb.1400500

Harborne, J. B., and Jeffrey, B. (1993). *Introduction to Ecological Biochemistry.*? San Diego, CA: Academic Press.

Haslam, E. (1986). Secondary metabolism – fact and fiction. *Nat. Prod. Rep.* 3, 217–249. doi: 10.1039/np9860300217

Hepler, P. K. (2005). Calcium: a central regulator of plant growth and development. *Plant Cell* 17, 2142–2155. doi: 10.1105/tpc.105.032508

Herms, D. A., and Mattson, W. J. (1992). The dilemma of plants: to grow or defend. *Q. Rev. Biol.* 67, 283–335. doi: 10.1086/417659

Hill, M. O. (1973). Reciprocal averaging: an eigenvector method of ordination. *J. Ecol.* 61, 237–249. doi: 10.2307/2258931

Hill, M. O. (1974). Correspondence analysis: a neglected multivariate method. *J. R. Stat. Soc. Ser. C* 23, 340–354. doi: 10.2307/2347127

Hiserodt, R., and Chen, L. (2012). "An LC/MS/MS Method for the Analysis of Furocoumarins in Citrus Oil", in *ACS Symposium Series. Recent Advances in the Analysis of Food and Flavors.* American (Chemical Society: Washington). 71–88. doi: 10.1021/bk-2012-1098.ch006

Hu, Y., Zhang, Q., Xin, H., Qin, L. P., Lu, B. R., Rahman, K., et al. (2007). Association between chemical and genetic variation of *Vitex rotundifolia* populations from different locations in China: its implication for quality control of medicinal plants. *Biomed. Chromatogr.* 21, 967–975. doi: 10.1002/bmc.841

Hunter, M. D., Malcolm, S. B., and Hartley, S. E. (1996). Population-level variation in plant secondary chemistry, and the population biology of herbivores. *CHEMOECOLOGY* 7, 45–56. doi: 10.1007/s10886-011-0018-1

Izhaki, I. (2002). Emodin – a secondary metabolite with multiple ecological functions in higher plants. *New Phytol.* 155, 205–217. doi: 10.1046/j.1469-8137.2002.00459.x

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 18, 477–483. doi: 10.1016/j.tplants.2013.06.003

Kaplan, F., Kopka, J., Haskell, D. W., Zhao, W., Schiller, K. C., Gatzke, N., et al. (2004). Exploring the temperature-stress metabolome of Arabidopsis. *Plant Physiol.* 136, 4159–4168. doi: 10.1104/pp.104.052142

Keefover-Ring, K., Thompson, J. D., and Linhart, Y. B. (2009). Beyond six scents: defining a seventh *Thymus vulgaris* chemotype new to southern france by ethanol extraction. *Flavour Fragr. J.* 24, 117–122. doi: 10.1002/ffj.1921

Koricheva, J. (1999). Interpreting phenotypic variation in plant allelochemistry: problems with the use of concentrations. *Oecologia* 119, 467–473. doi: 10.1007/s004420050809

Lankadurai, B. P., Nagato, E. G., and Simpson, M. J. (2013). Environmental metabolomics: an emerging approach to study organism responses to environmental stressors. *Environ. Rev.* 21, 180–205. doi: 10.1139/er-2013-0011

Lecourieux, D., Ranjeva, R., and Pugin, A. (2006). Calcium in plant defence-signalling pathways. *New Phytol.* 171, 249–269. doi: 10.1111/j.1469-8137.2006.01777.x

Lepš, J., and Šmilauer, P. (2003). *Multivariate Analysis of Ecological Data using CANOCO.* New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511615146

Mazzafera, P. (1999). Mineral nutrition and caffeine content in coffee leaves. *Bragantia* 58, 387–391. doi: 10.1590/S0006-87051999000200018

Molina-Montenegro, M. A., and Naya, D. E. (2012). Latitudinal patterns in phenotypic plasticity and fitness-related traits: assessing the climatic variability hypothesis (CVH) with an invasive plant species. *PLoS One* 7:e47620. doi: 10.1371/journal.pone.0047620

Moore, B. D., Andrew, R. L., Külheim, C., and Foley, W. J. (2014). Explaining intraspecific diversity in plant secondary metabolites in an ecological context. *New Phytol.* 201, 733–750. doi: 10.1111/nph.12526

Moral, F. J., Terrón, J. M., and Marques da Silva, J. R. (2010). Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. *Soil Tillage Res.* 106, 335–343. doi: 10.1016/j.still.2009.12.002

Moustafa, M. F., Hesham, A. E. L., Quraishi, M. S., and Alrumman, S. A. (2016). Variations in genetic and chemical constituents of Ziziphus spina-christi L.

Populations grown at various altitudinal zonation up to 2227m height. *J. Genet. Eng. Biotechnol.* 14, 349–362. doi: 10.1016/j.jgeb.2016.09.001

Ncube, B., Finnie, J. F., and Van Staden, J. (2012). Quality from the field: the impact of environmental factors as quality determinants in medicinal plants. *S. Afr. J. Bot.* 82, 11–20. doi: 10.1016/j.sajb.2012.05.009

Officer, S. J., Kravchenko, A., Bollero, G. A., Sudduth, K. A., Kitchen, W. J., Wiebold, H. L., et al. (2004). Relationships between soil bulk electrical conductivity and the principal component analysis of topography and soil fertility values. *Plant Soil* 258, 269–280. doi: 10.1023/B:PLSO.0000016557.94937.ed

O'Reilly-Wapstra, J. M., Miller, A. M., Hamilton, M. G., Williams, D., Glancy-Dean, N., Potts, B. M., et al. (2013). Chemical variation in a dominant tree species: population divergence, selection and genetic stability across environments. *PLoS One* 8:e58416. doi: 10.1371/journal.pone.0058416

Paliy, O., and Shankar, V. (2016). Application of multivariate statistical techniques in microbial ecology. *Mol. Ecol.* 25, 1032–1057. doi: 10.1111/mec.13536

Payyavula, R. S., Navarre, D. A., Kuhl, J. C., Pantoja, A., and Pillai, S. S. (2012). Differential effects of environment on potato phenylpropanoid and carotenoid expression. *BMC Plant Biol.* 12:39. doi: 10.1186/1471-2229-12-39

Pennycooke, J. C., Cox, S., and Stushnoff, C. (2005). Relationship of cold acclimation, total phenolic content and antioxidant capacity with chilling tolerance in petunia (Petunia x hybrida). *Environ. Exp. Bot.* 53, 225–232. doi: 10.1016/j.envexpbot.2004.04.002

Pinheiro, C. U. (1997). Jaborandi (Pilocarpus sp., rutaceae): a wild species. *Econ. Bot.* 51, 49–58. doi: 10.1007/BF02910403

Pinheiro, C. U. (2002). Extrativismo, cultivo e privatização do jaborandi (Pilocarpus microphyllus Stapf ex Holm.; Rutaceae) no Maranhão, Brasil. *Acta Bot. Brasilica* 16, 141–150. doi: 10.1590/S0102-33062002000200002

Pluskal, T., Castillo, S., Villar-Briones, A., and Orešiè, M. (2010). MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11:395. doi: 10.1186/1471-2105-11-395

Rivero, R. M., Ruiz, J. M., Garcìa, P. C., López-Lefebre, L. R., Sánchez, E., Romero, L., et al. (2001). Resistance to cold and heat stress: accumulation of phenolic compounds in tomato and watermelon plants. *Plant Sci.* 160, 315–321. doi: 10.1016/S0168-9452(00)00395-2

Sampaio, B. L., Edrada-Ebel, R., and Costa, F. B. D. (2016). Effect of the environment on the secondary metabolic profile of *Tithonia diversifolia*: a model for environmental metabolomics of plants. *Sci. Rep.* 6:29265. doi: 10.1038/srep29265

Santesson, R. (1968). *Svensk Naturvetenskap*. Stockholm: Allf.

Santos, A. P., and Moreno, P. R. H. (2004). Pilocarpus spp: a survey of its chemical constituents and biological activities. *Rev. Brasileira Ciências Farm.* 40, 116–137. doi: 10.1590/S1516-93322004000200002

Sardans, J., Penuelas, J., and Rivas-Ubach, A. (2011). Ecological metabolomics: overview of current developments and future challenges. *Chemoecology* 21, 191–225. doi: 10.1007/s00049-011-0083-5

Sawaya, A. C. H. F., Vaz, B. G., Eberlin, M. N., and Mazzafera, P. (2011). Screening species of pilocarpus (Rutaceae) as sources of pilocarpine and other imidazole alkaloids. *Genet. Resour. Crop Evol.* 58, 471–480. doi: 10.1007/s10722-011-9660-2

Skorupa, L. A. (1996). *Revisao Taxonomica de Pilocarpus Vahl (Rutaceae)*. Ph.D. thesis, university of São Paulo, Brazil.

Šmilauer, P., and Lepš, J. (2014). *Multivariate Analysis of Ecological Data using CANOCO 5*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9781139627061

Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* 78, 779–787. doi: 10.1021/ac051437y

Spichtig, M., and Kawecki, T. J. (2004). The maintenance (or not) of polygenic variation by soft selection in heterogeneous environments. *Am. Nat.* 164, 70–84. doi: 10.1086/421335

Sudduth, K. A., Kitchen, N. R., Bollero, G. A., Bullock, D. G, and Wiebold, W. J. (2003). Comparison of electromagnetic induction and direct sensing of soil electrical conductivity. *Agron. J.* 95, 472–482. doi: 10.2134/agronj2003.0472

Sultan, S. (1987). "Evolutionary Implications of Phenotypic Plasticity in Plants", in *Evolutionary Biology*, eds Hecht, M. K., Wallace, B., and Prance, G. T. (Springer, Boston, MA). doi: 10.1007/978-1-4615-6986-2_7

Tack, A. J. M., Johnson, M. T. J., and Roslin, T. (2012). Sizing up community genetics: it's a matter of scale. *Oikos* 121, 481–488. doi: 10.1111/j.1600-0706.2011.19926.x

Teixeira, A. F., Andrade, A. B., Ferrarese-Filho, O., and Ferrarese, M. L. L. (2006). Role of calcium on phenolic compounds and enzymes related to lignification in soybean (Glycine max L.) root growth. *Plant Growth Regul.* 49, 69–76. doi: 10.1007/s10725-006-0013-7

Thoday, J. M. (1953). Components of fitness. Symposium Soc. Exp. Biol. 7, 96–113

Tine, Y., Renucci, F., Costa, J., Wélé, A., and Paolini, J. (2017). A method for LC-MS/ms profiling of coumarins in zanthoxylum zanthoxyloides (Lam.) B. zepernich and timler extracts and essential oils. *Molecules* 22:E174. doi: 10.3390/molecules22010174

Tomás-Barberán, F. A., Gil, M. I., Castañer, M., Artés, F., Saltveit, M. E. (1997). Effect of selected browning inhibitors on phenolic metabolism in stem tissue of harvested lettuce. *J. Agric. Food Chem.* 45, 583–589. doi: 10.1021/jf960478f

van Rai, J. B. (2001). *Análise química para avaliação da fertilidade de solos tropicais*. Campinas: Instituto Agronômico.

Vialart, G., Hehn, A., Olry, A., Ito, K., Krieger, C., Larbat, R., et al. (2012). A 2-oxoglutarate-dependent dioxygenase from Ruta graveolens L. exhibits p-coumaroyl CoA 2′-hydroxylase activity (C2′H): a missing step in the synthesis of umbelliferone in plants. *Plant J.* 70, 460–470. doi: 10.1111/j.1365-313X.2011.04879.x

Waller, G. R. (1978). *Alkaloid Biology and Metabolism in Plants*. New York, NY: Plenum Press. doi: 10.1007/978-1-4684-0772-3

Whittaker, R. H. (1956). Vegetation of the great smoky mountains. *Ecol. Monogr.* 26, 1–80. doi: 10.2307/1943577

Wink, M. (Ed.) (1999). *Biochemistry of Plant Secondary Metabolism*. Sheffield: WileyBlackwell.

Xia, J., and Wishart, D. S. (2002). Using metaboanalyst 3.0 for comprehensive metabolomics data analysis. *Curr. Protoc. Bioinform.* 55, 14.10.1–14.10.91.

Xu, W., Peng, H., Yang, T., Whitaker, B., Huang, L., Sun, J., et al. (2014). Effect of calcium on strawberry fruit flavonoid pathway gene expression and anthocyanin accumulation. *Plant Physiol. Biochem.* 82, 289–298. doi: 10.1016/j.plaphy.2014.06.015

Yan, X., Wu, S., Wang, Y., Shang, X., and Dai, S. (2004). Soil nutrient factors related to salidroside production of rhodiola sachalinensis distributed in chang bai mountain. *Environ. Exp. Bot.* 52, 267–276. doi: 10.1016/j.envexpbot.2004.02.005

Zhang, L., Du, L., and Poovaiah, B. W. (2014). Calcium signaling and biotic defense responses in plants. *Plant Signal. Behav.* 9:e973818 doi: 10.4161/15592324.2014.973818