Complimentary and personal copy for

Paula Carolina Pires Bueno, Luis Francisco Salomé Abarca, Naira Buzzo Anhesine, Maíra Silva Giffoni, Fabiola Manhas Verbi Pereira, Roseli Buzaneli Torres, Rayran Walter Ramos de Sousa, Paulo Michel Pinheiro Ferreira, Claudia Pessoa, Alberto José Cavalheiro

www.thieme.com

Infraspecific Chemical Variability and Biological Activity of *Casearia sylvestris* from Different Brazilian Biomes

DOI 10.1055/a-1301-0183
Planta Med

This electronic reprint is provided for non-commercial and personal use only: this reprint may be forwarded to individual colleagues or may be used on the author's homepage. This reprint is not provided for distribution in repositories, including social and scientific networks and platforms.

Publishing House and Copyright:
© 2020 by
Georg Thieme Verlag KG
Rüdigerstraße 14
70469 Stuttgart
ISSN 0032-0943

Any further use
only by permission
of the Publishing House

Thieme

Infraspecific Chemical Variability and Biological Activity of *Casearia sylvestris* from Different Brazilian Biomes

Authors
Paula Carolina Pires Bueno[1], Luis Francisco Salomé Abarca[2], Naira Buzzo Anhesine[1], Maíra Silva Giffoni[1], Fabiola Manhas Verbi Pereira[1,3], Roseli Buzaneli Torres[4], Rayran Walter Ramos de Sousa[5], Paulo Michel Pinheiro Ferreira[5], Claudia Pessoa[6], Alberto José Cavalheiro[1]

Affiliations
1 Institute of Chemistry, Department of Biochemistry and Organic Chemistry, São Paulo State University (UNESP), Araraquara/SP, Brazil
2 Natural Products Laboratory, Institute of Biology, Leiden University, Leiden, The Netherlands
3 National Institute of Alternative Technologies for Detection Toxicological Assessment and Removal of Micropollutants and Radioactive Substances (INCT-DATREM) Araraquara, São Paulo State, Brazil
4 Agronomic Institute of Campinas, Herbarium IAC, Campinas/SP, Brazil
5 Department of Biophysics and Physiology, Laboratory of Experimental Cancerology, Federal University of Piauí, Teresina/PI, Brazil
6 Department of Physiology and Pharmacology, Faculty of Medicine, Federal University of Ceará, Fortaleza/CE, Brazil

Key words
Casearia sylvestris, Salicaceae, clerodane diterpenoids, glycosylated flavonoids, biological activities, plant ecophysiology, metabolomics, multivariate data analysis

received June 28, 2020
accepted after revision October 26, 2020
published online

Bibliography
Planta Med 2020
DOI 10.1055/a-1301-0183
ISSN 0032-0943
© 2020. Thieme. All rights reserved.
Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Correspondence
Prof. Dr. Alberto José Cavalheiro
Institute of Chemistry, Department of Biochemistry and Organic Chemistry, São Paulo State University (UNESP)
CEP 14800-060, Rua Prof. Francisco Degni 55,
14800-900 Araraquara, SP, Brazil
Phone: + 55 16 33 01 96 67, Fax: + 55 16 33 22 23 08
albjcava@iq.unesp.br

Supplementary material is available under https://doi.org/10.1055/a-1301-0183

ABSTRACT

Casearia sylvestris is an outstanding representative of the *Casearia* genus. This representability comes from its distinctive chemical profile and pharmacological properties. This species is widespread from North to South America, occurring in all Brazilian biomes. Based on their morphology, 2 varieties are recognized: *C. sylvestris* var. *sylvestris* and *C. sylvestris* var. *lingua*. Despite the existence of data about their chemical composition, a deeper understanding of the specialized metabolism correlation and variation in respect to environmental factors and its repercussion over their biological activities was still pending. In this study, an UHPLC-DAD-based metabolomics approach was employed for the investigation of the chemical variation of 12 *C. sylvestris* populations sampled across 4 Brazilian biomes and ecotones. The correlation between infraspecific chemical variability and the cytotoxic and antioxidant activities was achieved by multivariate data analysis. The analyses showed that *C. sylvestris* var. *lingua* prevailed at Cerrado areas, and it was correlated with lower cytotoxic activity and high level of glycosylated flavonoids. Among them, narcissin and isorhamnetin-3-*O*-*α*-L-rhamnopyranosyl-(1 → 2)-*α*-L-arabinopyranoside showed good correlation with the antioxidant activity. Conversely, *C. sylvestris* var. *sylvestris* prevailed at the Atlantic Forest areas, and it was associated with high cytotoxic activity and high content of clerodane diterpenoids. Different casearins showed good correlation (R^2 = 0.3–0.70) with the cytotoxic activity. These findings highlighted the great complexity among different *C. sylvestris* populations, their chemical profile, and the related biological activities. Consequently, it can certainly influence the medicinal properties, as well as the quality and efficacy, of *C. sylvestris* phytomedicines.

Electronic reprint for personal use

Introduction

Plants produce a broad range of metabolites, many of which are species specific, presenting many biological roles [1,2]. Even within species, plant populations can exhibit a large metabolic variation that can be triggered and regulated by multiple factors, such as induction of cryptic genes and/or environmental factors [3,4]. The occurrence of such metabolic variation is especially relevant for the selection and use of medicinal plants. Depending on the geographical origin or time of collection of the plant specimens, the contents of bioactive metabolites can differ considerably among them [5,6]. In general, qualitative and quantitative variations can occur not only for a single compound but also for many of them [6].

In this regard, *Casearia sylvestris* Sw. (Salicaceae), one of the 180 species of the *Casearia* genus, is a well-known pantropical plant species showing morphological and chemical variability across different ecosystems [7–9]. Moreover, its high adaptive capability enables it to proliferate not only in Brazil but also from Mexico to Uruguay [7]. In Brazil, this species is popularly known as "guaçatonga" and occurs in several biomes including the Atlantic Forest, Amazon Forest, Cerrado, Caatinga, and Pantanal [9].

Based on their morphological traits, 2 distinguished varieties are recognized as *C. sylvestris* var. *sylvestris* and *C. sylvestris* var. *lingua* [7,8]. *C. sylvestris* var. *sylvestris*, commonly found at the Brazilian Atlantic Forest, is a tree taller than 2 m, while *C. sylvestris* var. *lingua* is a shrub commonly found in the Cerrado biome [7,10]. In Brazil, as well as in other Latin American countries, *C. sylvestris* is used as leaves' crude extracts and infusions in the folk medicine. It is mainly employed in the treatment of snakebites, skin diseases, diarrhea, inflammations, fever, and gastric ulcers, among others [11,12]. Due to its ethnopharmaceutical relevance, *C. sylvestris* is one of the 71 plant species listed in the Brazilian National List of Medicinal Plants of Interest of the Brazilian Unified Health System [13].

The chemistry of *C. sylvestris* is characterized by the presence of oxygenated clerodane diterpenoids, which are also considered chemotaxonomic markers for the *Casearia* genus. These compounds, which are also called casearins, can be found in several plant organs, including leaves, stems, roots, and seeds [14–17]. In addition, some tannin derivatives were identified in the *Casearia* species [18,19]. Moreover, 14 3-*O*-glycosylated flavonoids and catechin derivatives were recently isolated from the leaves of *C. sylvestris* var. *lingua*, providing the first clue about its metabolic differentiation from *C. sylvestris* var. *sylvestris* [20]. In this context, the correlation between the production of glycosylated flavonoids and clerodane diterpenoids in the varieties *lingua* and *sylvestris* respectively has been demonstrated [20,21].

Biological activities of *C. sylvestris* extracts, fractions, and isolated compounds have been extensively studied. Anti-ulcer, anti-inflammatory, and cytotoxic bioassays attribute the biological activity to the clerodane terpenoidal contents [11,22,23]. In this regard, cytotoxicity studies have shown special relevant antitumor potential of several clerodane diterpenoids isolated from *C. sylvestris* against a series of tumor cell lines. As an example, the ethanolic extract of *C. sylvestris* leaves and its isolated casearins A–F showed cytotoxic activity against Chinese hamster lung fibroblasts and murine sarcoma 180 cells [24]. Caseargrewiin F and casearin X showed high cytotoxic activity against cancer cell lines MOLT-4 (leukemia), MDA-MB-435 (melanoma), HCT-8 (colon), SF-295 (glyoblastoma), and L-929 (normal fibroblasts) [25]. On the other hand, the ethanolic leaf extract of *C. sylvestris* var. *lingua* did not show cytotoxic activity when challenged against J774 macrophage cells, which could be correlated to the absence of clerodane diterpenoids in this variety [26].

Despite the extensive research about the bioactivity potential of *C. sylvestris*, the effect of the geographical origin on its chemical composition and its biological effects has not been systematically demonstrated. This is a quintessential requirement especially for the use of *C. sylvestris* materials in folk medicine as well as a resource of bioactive metabolites. In this regard, metabolomics has been proven to be a high throughput and robust approach to unveil such metabolic changes in plants. The correlation between plants' metabolic variations and their potential pharmacological efficacy can be scrutinized by multivariate data analysis [5,27].

On the other hand, the environmental interactions occurring at the guaçatonga niches can constitute a complex network that make the metabolic analysis difficult to interpret. Studies on the genetic diversity and in field observations in south-east Brazil supported the occurrence of 2 varieties as independent biological units, as well hybridization between them. However, the genetic relationships between the 2 varieties may differ in other localities, such as in the central northern Brazil, where the 2 varieties may also occur sympatrically [7].

Therefore, the aims of the present research were to perform a comprehensive study to assign the geographical distribution of *C. sylvestris* according to its morpho-anatomic traits; subsequently to perform a correlation between the metabolic contents of the different sampled *C. sylvestris* populations and their geographical origin; and finally to study the relationship between their metabolic variation and their cytotoxic and antioxidant activities. To do so, LC-DAD-based metabolomics complemented by offline LC-MS analyses and multivariate data analysis were employed.

Results and Discussion

In the present study, leaves from *C. sylvestris* var. *lingua, C. sylvestris* var. *sylvestris*, and an intermediate morphotype were sampled at 12 different geographical locations of Brazil including 4 different biomes: Atlantic Forest, Cerrado, Pantanal, and Pampa, as well as their ecotones. The individual samples were morphologically characterized and chemically profiled (n = 61) (▶ **Fig. 1** and **Table 1**) [28]. All samples were processed and analyzed under the same conditions, according to a previously validated UHPLC-DAD method [21]. The chromatograms were plotted at 254 nm from 0 to 25 min, which comprised the glycosylated flavonoids eluting region, defined by their characteristic UV spectra and previously isolated standards. From 25 to 45 min, the chromatograms were plotted at 235 nm, which comprised the diterpenes eluting region. The peak areas were normalized to the internal standard. The resulting data matrix (61 samples × 6751 variables) was mean centered prior to principal component analysis (PCA).

The resulting model consisted of 4 PCs, which explained 56% of the total variance. The score plot of PC1 (27.7%) against PC2

Electronic reprint for personal use

▶ **Fig. 1** Geographical limits of biomes and ecotones used as sampling regions for *C. sylvestris* in Brazil. Source: Thomas Heinemann.

(10.7%) showed a clear separation between varieties *lingua* and *sylvestris* along PC1 (▶ **Fig. 2a**). Samples showing intermediary morphology formed a third group between the 2 varieties' clusters (▶ **Fig. 2a**). From this perspective, the observed clusters highlighted the main driving factors for the chemical differentiation of the samples. Such factors were varieties and geographical origin effects; nonetheless, a stronger effect was observed in varieties. To clarify and validate these results, a supervised analysis was employed. In order to give the same weight to all variables and suppress effects noncorrelated to varieties' effects, an orthogonal projection to latent structures discriminant analysis (OPLS-DA) scaled by unit variance (UV) method was constructed. The factor variety was set as class, with a total of 3 classes. The OPLS-DA model was highly validated by the permutation ($Q^2 = 0.74$) and CV-ANOVA ($p < 0.01$) tests, and the plot showed a clear separation among the 3 varieties (**Fig. 1S**, Supporting Information).

The PCA loading plot determined the retention time of 11.3 min as the main variable contributing to the clustering in the positive direction of PC1 (▶ **Fig. 2b**). The characteristic UV spectra (UV max = 255 and 356 nm), the HRMS (623.1620 $[M - H]^-$), and the MS/MS fragmentation pattern (30 eV: 623 → 314 [315]) allowed the annotation of this variable as isorhamnetin-3-*O*-rutinoside (narcissin). This information was further confirmed by comparison with the retention time of a previously isolated compound from the same plant species. The cluster including samples classified as *C. sylvestris* var. *lingua* and samples with intermediary morphology were growing surrounded by Cerrado's vegetation from different Brazilian States. These results matched with previous research showing the prevalence of samples rich in phenolic compounds at Cerrado and ecotones areas from São Paulo State [21]. On the other hand, the occurrence of clerodane diterpenoids was correlated with *C. sylvestris* var. *sylvestris* collected at Atlantic Forest and remnant areas in the negative direction of PC1 (▶ **Fig. 2a**). Other compounds responsible for the sample clustering were identified as casearin D (27.5 min), caseargrewiin F (29.1 min), casearin S (29.5 min), and casearin X (30.7 min) (▶ **Table 2** and **Figs. 2S–5S**, Supporting Information). All those diterpenes showed characteristic UV spectra (UV max 236, 230, 234, and 228, respectively), and HRMS compatible with previously isolated compounds (571.2849 $[M + Na]^+$, 527.2623 $[M + Na]^+$, 497.2292 $[M + Na]^+$, and 555.2935 $[M + Na]^+$, respectively). The MS/MS fragmentation pattern of all those compounds matched with the characteristic cleavages of the acetate and butanoate groups (losses of 60 and 88 Da, respectively), attached to the decalin system nucleus of such diterpenoids [29].

Samples showing intermediary morphology from Pampa and Pantanal were also clustered together with *C. sylvestris* var. *sylvestris*. Interestingly, those samples were collected in ecotones comprising Cerrado and flooded areas, with high water availability in the soil. This fact provided a glimpse into the effect of environmental factors over the chemical shaping of those ecotypes. In this context, it is worth addressing the fact that ecotypes can result at the same geographic region with different ecosystems or from separated regions with related ecological conditions [30]. This might be the case of samples showing intermediary morphology overlapping the different varieties. Moreover, this also highlights the use of their chemical profiles to complement the variety assignation made by morpho-anatomical features.

To score the effect of the geographical origin, the sample set was divided according to the sample variety. The PCA ($R^2 = 0.59$) analysis of *C. sylvestris* var. *lingua* showed a good separation between the samples along both negative and positive directions of PC1 (▶ **Fig. 3a**). When an OPLS-DA model was constructed setting geographical origin as classes (▶ **Fig. 3b**), the plot also showed a good separation and high Q^2-value (0.50) but was still not validated by the CV-ANOVA test. This could indicate the need for a higher number of samples per class. However, the data from PCA indicated a metabolic differentiation dictated by the samples' geographical origin. A similar trend was observed for *C. sylvestris* var. *sylvestris* (▶ **Fig. 3c**). Nonetheless, the sample separation in the PCA ($R^2 = 0.63$) was less evident than in *C. sylvestris* var. *lingua*. The OPLS-DA analysis showed a good separation, but it was not validated (▶ **Fig. 3d**). The samples with intermediary morphology did not show good separation in the PCA plot (▶ **Fig. 3e**). This might indicate 2 possible explanations. First, *C. sylvestris* var. *lingua* is more prone to present metabolic changes caused by environmental factors across geographical regions than *C. sylvestris* var. *sylvestris* and the intermediate group. Second, the ecological niches at Cerrado areas where *C. sylvestris* var. *lingua* grows contrast more among each other, thus influencing bigger metabolic changes among populations.

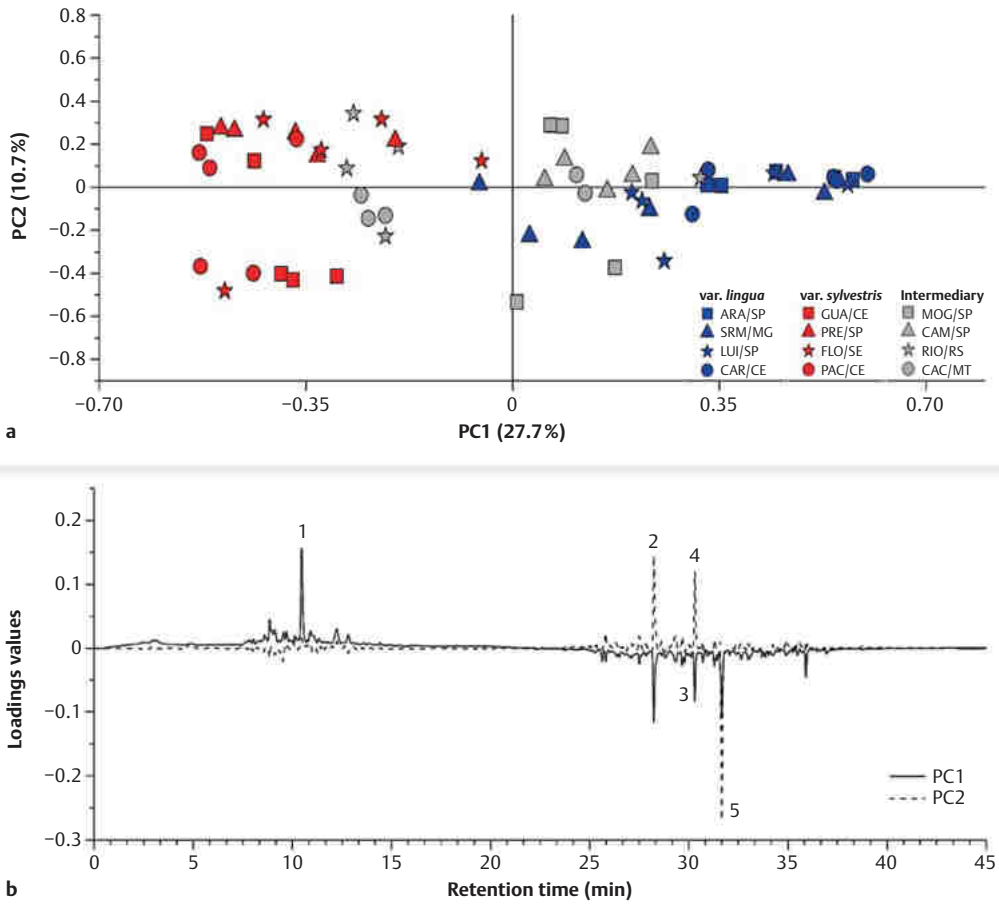
It is worth noting that *C. sylvestris* var. *lingua,* mainly found in Cerrado or open and xeric areas, are constantly subjected to high sunlight exposition and hot weather. Thus, these specimens suffer higher abiotic stress when compared to those located at areas where the canopy filters the excess of light, as in the Atlantic Forest. Therefore, this could be one of the reasons explaining *C. syl-*

Electronic reprint for personal use

▶ **Table 1** *C. sylvestris* representatives collected in different Brazilian biomes.

Location[1]	Abbreviation	N[2]	Variety[3]	Biome	Geographical coordinates	Elevation (m)	Climate
Araraquara/SP	ARA/SP	5	*lingua*	Cerrado	21°48′ 55′′ S 48°11′ 48′′ W	629	hot/dry
São Roque de Minas/MG	SRM/MG	6	*lingua*	Cerrado	20°08′ 47′′ S 46°39′ 47′′ W	1202	hot/dry
Cariri/CE	CAR/CE	5	*lingua*	Ecotone[4]	07°09′ 19′′ S 39°43′ 14′′ W	626	hot/dry
Luis Antônio/SP	LUI/SP	5	*lingua*	Ecotone[5]	21°33′ 22′′ S 47°53′ 31′′ W	604	hot/dry
Mogi-Guaçu/SP	MOG/SP	5	*intermediary*	Ecotone[5]	22°13′ 05′′ S 47°09′ 34′′ W	649	hot/dry
Rio Grande/RS	RIO/RS	5	*intermediary*	Pampa[6]	32°04′ 54′′ S 52°09′ 48′′ W	2	warm/dry
Cáceres/MT	CAC/MT	5	*intermediary*	Pantanal[6]	16°04′ 29′′ S 57°40′ 07′′ W	120	warm/dry
Campinas/SP	CAM/SP	5	*intermediary*	Cerrado[7]	22°48′ 58′′ S 47°03′ 45′′ W	623	warm/dry
Guaramiranga/CE	GUA/CE	5	*sylvestris*	Atlantic Forest	04°12′ 42′′ S 38°56′ 50′′ W	806	hot/humid
Pacoti/CE	PAC/CE	5	*sylvestris*	Atlantic Forest	04°10′ 23′′ S 38°52′ 15′′ W	334	hot/humid
Florianópolis/SC	FLO/SC	5	*sylvestris*	Atlantic Forest	27°40′ 33′′ S 48°33′ 54′′ W	5	hot/humid
Presidente Venceslau/SP	PRE/SP	5	*sylvestris*	Atlantic Forest[8]	21°53′ 10′′ S 51°52′ 28′′ W	449	hot/humid

[1] Location or population names refer to the municipalities where samples were collected; [2] number of collected individuals; [3] *C. sylvestris* varieties classification based on botanical characterization (morphology); [4] ecotone Cerrado/Caatinga; [5] ecotone Cerrado/Atlantic Forest; [6] ecotone Cerrado/flooded areas; [7] ciliary forest; [8] remnant area



▶ **Fig. 2** **a** Score plot (PC1 vs. PC2) of *C. sylvestris* samples (n = 61) colored according to variety morphologies (*C. sylvestris* var. *lingua*, *C. sylvestris* var. *sylvestris*, and individuals showing intermediary morphology) and shaped by population. **b** Loadings plot (PC1 vs. PC2) for variables discrimination: (1) narcissin; (2) casearin D; (3) caseargrewiin F; (4) casearin S; (5) casearin X.

Electronic reprint for personal use

▶ **Table 2** Compounds annotated in the 12 *C. sylvestris* lyophilized extracts by LC-DAD-ESI-MS/MS (HRMS and MS/MS in negative or positive mode).

Code	*RR*t (min)	Compound name	UV max (nm)	Negative or positive ionization (*m/z*)		Exact mass		Error (ppm)
				HRMS	MS/MS	experimental	calculated	
F-4	8.7	Isorhamnetin-3-*O*-trihexoside	254; 350	769.2154 $[M-H]^-$	30 eV: 769 → 314	770.2227	770.2270	5.609
F-7	9.5	Quercetin-3-*O*-rutinoside (rutin)	250; 350	609.1462 $[M-H]^-$	30 eV: 609 → 300 (301)	610.1535	610.1534	− 0.131
F-8	10.4	Isorhamnetin-3-*O*-neohesperidoside	253; 353	623.1620 $[M-H]^-$	35 eV: 623 → 314 (315)	624.1693	624.1690	− 0.449
F-10	11.3	Isorhamnetin-3-*O*-rutinoside (narcissin)	255; 356	623.1587 $[M-H]^-$	30 eV: 623 → 314 (315)	624.1660	624.1690	4.838
F-11	11.7	Quercetin-3-*O*-hexoside	250; 353	433.0803 $[M-H]^-$	35 eV: 433 → 300 (301)	434.0876	434.0849	− 6.174
F-12	12.2	Isorhamnetin-3-*O*-hexoside	254; 350	477.1019 $[M-H]^-$	30 eV: 477 → 314 (315)	478.1092	478.1111	4.016
		Kaempferol-3-*O*-dihexoside		563.1394 $[M-H]^-$	30 eV: 563 → 284 (285)	564.1467	564.1479	2.163
F-13	13.0	Isorhamnetin-3-*O*-*α*-L-rhamnopyranosyl-(1 → 2)-*α*-L-arabino-pyranoside	254; 356	593.1516 $[M-H]^-$	30 eV: 593 → 314 (315)	594.1589	594.1585	− 0.640
D-2	24.0	Casearin L/A	233	557.2719 $[M+Na]^+$	n. a.	534.2838	534.2829	− 1.685
D-13	27.5	Casearin M/P/D	236	571.2849 $[M+Na]^+$	30 eV: 571 → 481, 423	548.2996	548.2985	− 2.006
D-15	28.8	Casearin H	232	527.2615 $[M+Na]^+$	n. a.	504.2723	504.2723	0.000
D-16	29.1	Caseargewiin F	230	527.2623 $[M+Na]^+$	n. a.	504.2731	504.2723	− 1.586
D-18	29.5	Casearin S	234	497.2292 $[M+Na]^+$	35 eV: 497 → 437, 349	474.2617	474.2618	0.211
D-22	30.7	Casearin X/I	228	555.2924 $[M+Na]^+$	n. a.	532.3032	532.3036	0.751
D-23	30.8	Casearin J	236	585.3027 $[M+Na]^+$	0 eV: 585 → 525, 497, 437	562.3138	562.3142	0.711
D-24	32.0	Casearin X/I	225	555.2934 $[M+Na]^+$	n. a.	532.3042	532.3036	1.086
D-25	34.0	Casearin C	232	669.3603 $[M+Na]^+$	n. a.	646.3711	646.3717	0.928

n. a. = not available

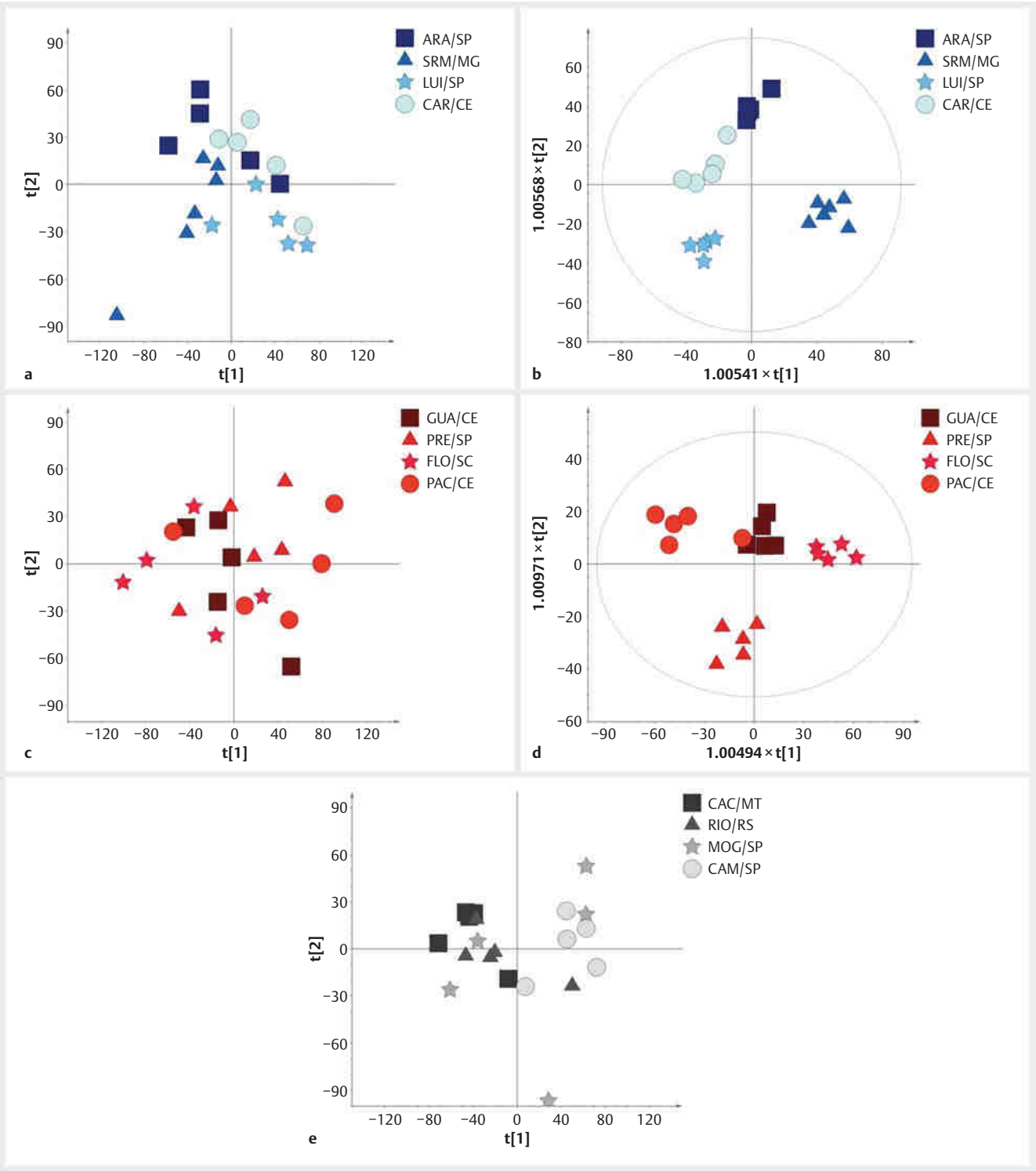
vestris var. *lingua's* distinctive chemical signature composed of phenolics compounds. It is well known that temperature and light radiance, among others environmental stressors, are important factors that causes oxidative stress in plants [31]. Depending on the light intensity to which plants are subjected, different molecular adaptation strategies will occur in synchrony to modulate the rate of light absorbed [32, 33]. The results suggest that the expression of genes encoding enzymes related to the phenylpropanoid and flavonoid biosynthesis pathways in *C. sylvestris* var. *lingua* are increased by exposure to UV radiation and higher oxidative stress [30, 34, 35].

To scrutinize the correlation between different *C. sylvestris* populations and their biological activity degree, the dried and powdered leaf material of each individual was combined according to its corresponding geographical origin. Therefore, the 12 populations were represented by 5 individuals per location. The resulting powders were extracted, lyophilized, and subjected to quantitative analysis by UHPLC-DAD with 3 replicates per sample (n = 36). ▶ **Table 2** and **Fig. 4** show the annotated compounds as well as the chromatographic profiles of each population. The detailed quantitative data is found in **Table 1S** (Supporting Information).

Subsequently, the 12 *C. sylvestris* lyophilized extracts were tested for their cytotoxic and antioxidant activities (**Table 2S**, Supporting Information). These bioactivity tests were chosen based on the tradiconal uses and biological activities reported for *C. sylvestris*. Correlation analyses were performed using the targeted quantitative data of clerodane diterpenoids and glycosylated flavonoids. Taking into account the characteristic UV spectra absorption of each compound class, narcissin and caseargrewiin F were selected to build the analytical curves for quantifications. In the data matrix, each quantified compound constituted a variable, conforming 38 variables (13 flavonoids and 25 diterpenoids). The cytotoxic and antioxidant activity constituted the Y variables in the OPLS analysis scaled by UV method.

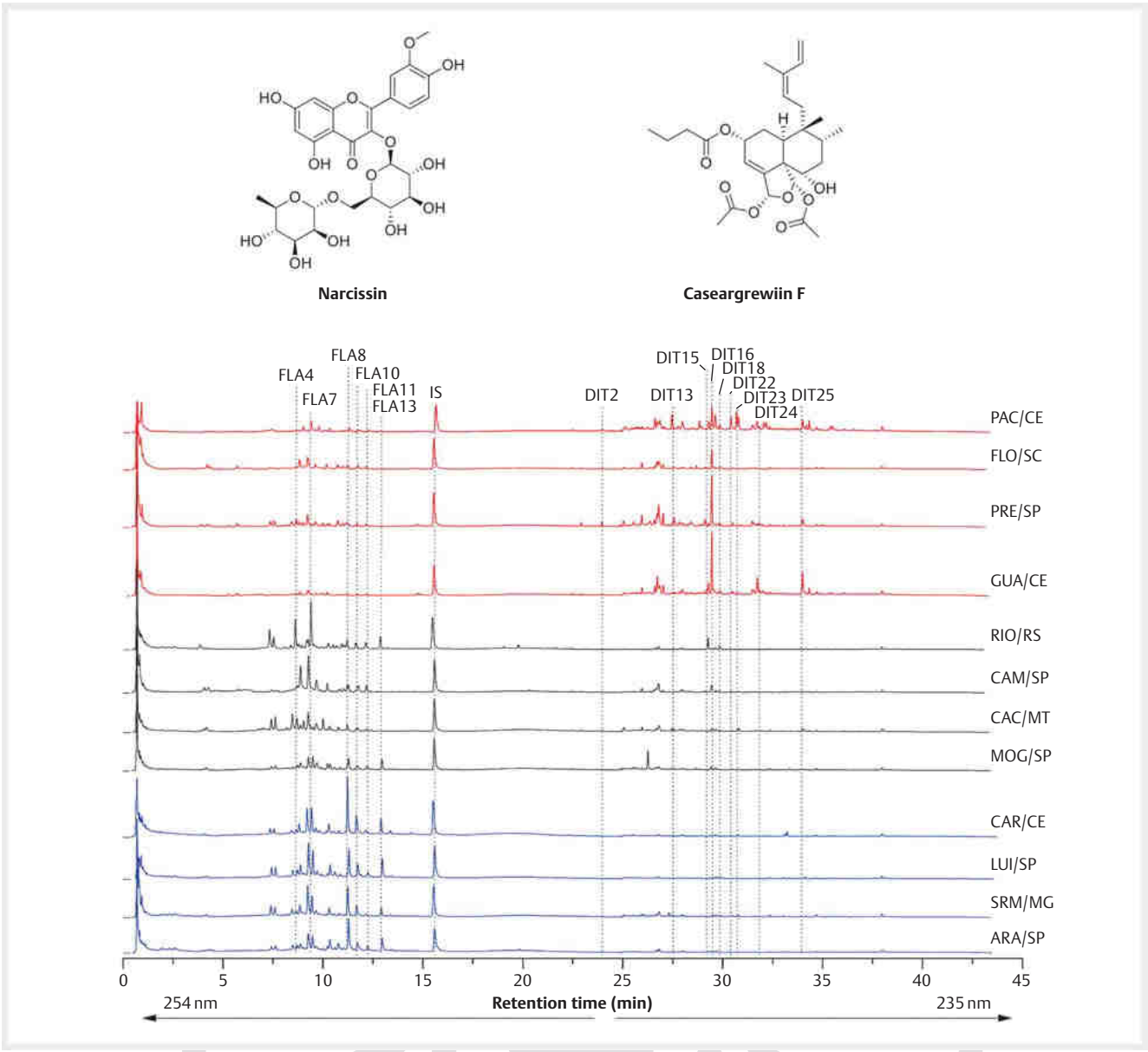
The individual OPLS analysis of the cytotoxic activity against HCT-116, PC-3, and SF-295 cell lines were validated with Q^2 values of 0.96, 0.90, and 0.97 respectively, and all of them with $p < 0.01$ in a CV-ANOVA test (▶ **Fig. 5a–c**). According to the loading plot, only variation in diterpenoids compounds correlated to the activity variation. According to the variable importance for the projection (VIP) plot, a correlation between the variation of clerodane diterpenoids and the cytotoxic activity degree of the *C. sylvestris*

Electronic reprint for personal use



▶ **Fig. 3** Principal component analysis (PCA) and ortogonal projection to latent structures discriminant analysis (OPLS-DA) of *C. sylvestris* var. *sylvestris*, *C. sylvestris* var. *lingua*, and individuals showing intermediary morphology. **a** PCA of 4 populations of *C. sylvestris* var. *lingua* colored according to their geographical origin; **b** OPLS-DA of 4 populations of *C. sylvestris* var. *lingua* using geographical origin as classes; **c** PCA of 4 populations of *C. sylvestris* var. *sylvestris* colored according to their geographical origin; **d** OPLS-DA of 4 populations of *C. sylvestris* var. *sylvestris* using geographical origin as classes; **e** PCA of 4 populations of individuals showing intermediary morphology colored according to their geographical origin.

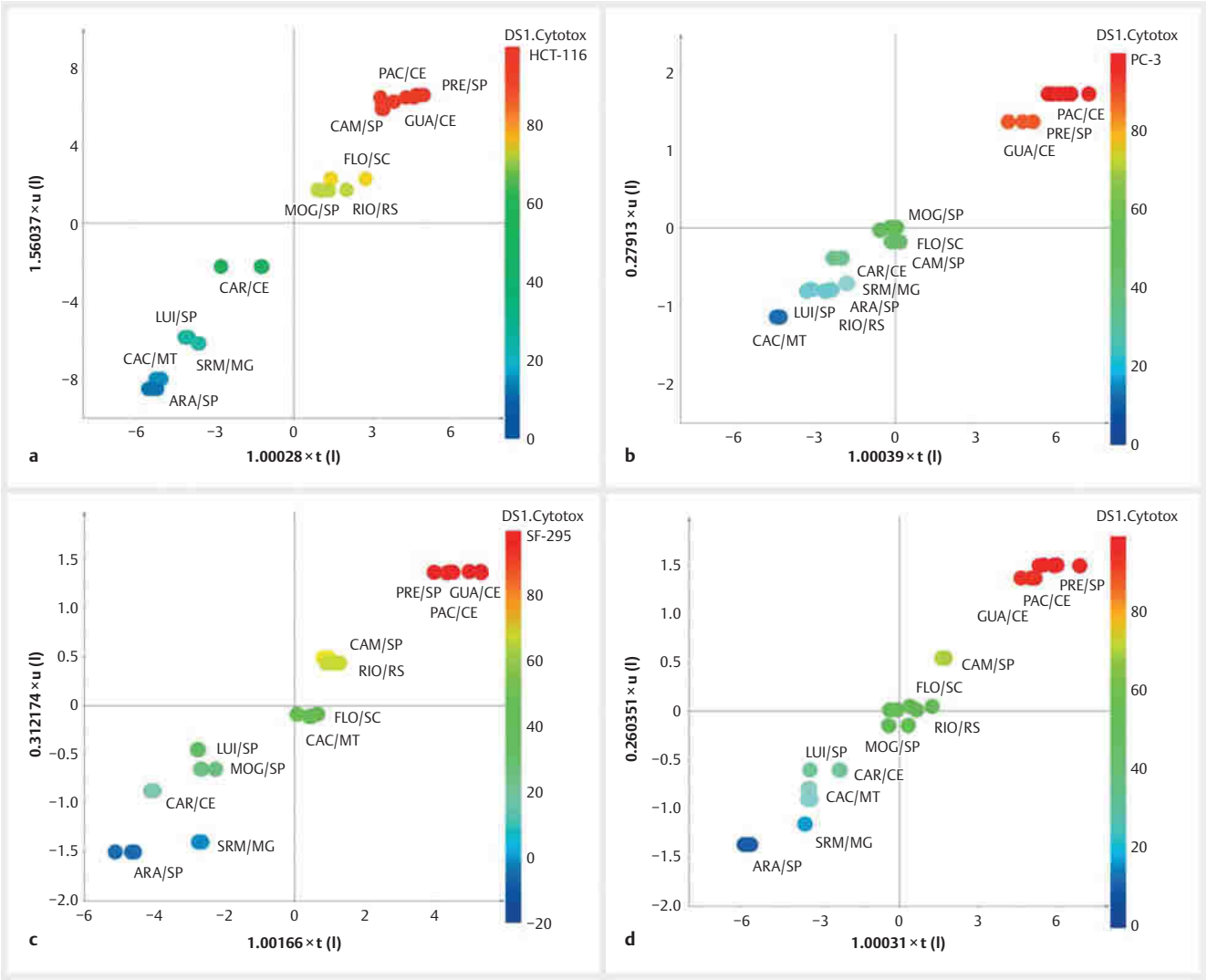
Electronic reprint for personal use



▶ **Fig. 4** UHPLC-UV fingerprints of 12 *C. sylvestris* lyophilized extracts. Chromatograms were plotted at 254 nm (from 0 to 25 min) and 235 nm (from 25 to 45 min). The IS signal corresponds to the internal standard butyl gallate. Peaks eluting from 6–14 min correspond to glycosylated flavonoids, as represented by narcissin; peaks eluting from 24–34 min correspond to clerodane diterpenoids, as represented by caseargrewiin F. Chromatographic conditions: Column: Kinetex 150 mm × 2.1 mm; 2.6 µm, 100 Å. Mobile phase: water (**a**) and acetonitrile (**b**). Gradient elution: 10–25 % of B from 0 to 15 min, 25–90 % of B until 35 min, holding 90 % of B until 40 min and returning to the initial condition in 2 min. Reequilibration time: 3 min. Flow rate: 400 µL/min. Oven temperature: 35 °C. Injection volume: 2 µL.

groups was stablished and confirmed (**Fig. 6S A–C**, Supporting information). Casearin S was one of the most correlated compounds with the cytotoxic activity against the cell lines. Interestingly, when performing individual linear regressions between individual metabolites and the activity range variation, casearin S showed a quite low correlation degree (R^2 < 0.20) with activity variation against all cell lines (▶ **Table 3**). This might suggest a possible interaction between casearin S and other diterpenoids such as casearin D, casearin X, and caseargrewiin F, which were also correlated to the cytotoxic activity variation in *C. sylvestris* groups. The OPLS analysis of the averaged cytotoxic activity against all cell lines was highly validated (Q^2 = 0.97, p < 0.01) (▶ **Fig. 5 d**). The VIP plot pointed out casearin S as one of the main metabolites correlated to the general cytotoxic activity variation (**Fig. 6 SD**, Supporting information). The other metabolites associated to the cytotoxic activity degree and variation also showed good correlation degree (R^2 = 0.30–0.70) when individually analyzed (▶ **Table 3**). These findings support the dependency of the presence of clerodane diterpenoids for the reported cytotoxic activities of *C. sylvestris*.

Electronic reprint for personal use



▶ **Fig. 5** Orthogonal projection to latent structures (OPLS) analysis for the chemical variation of 12 *C. sylvestris* populations and its correlation with their cytotoxic activity against 3 cell lines. **a** OPLS analysis for the correlation of 12 *C. sylvestris* populations and their cytotoxic effect against HCT-116 cells; **b** OPLS analysis for the correlation of 12 *C. sylvestris* populations and their cytotoxic effect against PC-3 cells; **c** OPLS analysis for the correlation of 12 *C. sylvestris* populations and their cytotoxic effect against SF-295 cells; **d** OPLS analysis built from the average data of the cytotoxic activity of 12 *C. sylvestris* populations against 3 cell lines and their chemical profiles (n = 36).

Moreover, the results indicated that samples from *C. sylvestris* var. *sylvestris* collected at Atlantic Forest had cytotoxic activities comparable with that of caseargrewiin F (**Fig. 7S**, Supporting Information). Additionally, their chemical profiles were the richest in clerodane diterpenoids, as showed by the quantitative analysis. The groups showing the lowest cytotoxic activity comprised samples of *C. sylvestris* var. *lingua* collected at Cerrado areas or ecotones between Cerrado/Caatinga and Cerrado/Atlantic Forest. The quantitative analysis for these samples showed undetectable concentrations of clerodane diterpenoids, which might be correlated with their low cytotoxic activity (**Fig. 7S**, Supporting Information).

Conversely, the antioxidant assay showed that the samples from *C. sylvestris* var. *lingua* possess higher antioxidant potential than those of *C. sylvestris* var. *sylvestris* (**Fig. 8S**, Supporting Information). The chemical profiling and the quantitative analyses confirmed the glycosylated flavonoids as the main specialized metabolites found in *C. sylvestris* var. *lingua*. The OPLS analysis ($Q^2 = 0.92$, $p < 0.01$) further confirmed narcissin and other glycosylated flavonoids as the main metabolites responsible for the antioxidant activity variation (▶ **Fig. 6**, **Fig. 6SE**, Supporting Information). In line with these results, the individual linear regression analysis of each detected flavonoid showed narcissin and isorhamnetin-3-*O*-α-L-rhamnopyranosyl-(1 → 2)-α-L-arabinopyranoside to have a good correlation ($R^2 > 0.30$) with the antioxidant activity variation in *C. sylvestris* groups (▶ **Table 3**).

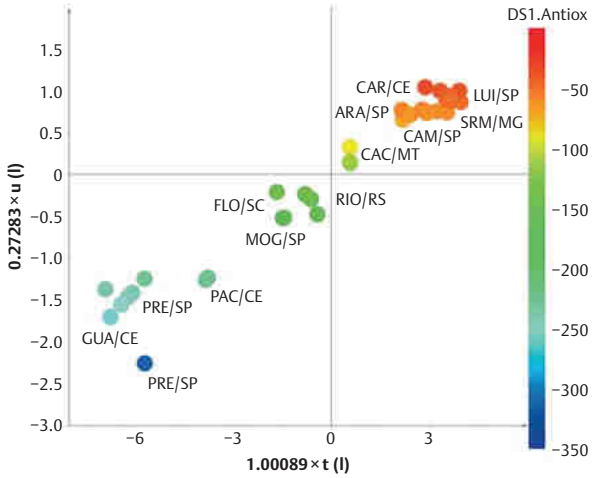
As a conclusion, regarding a previous report concerning the prevalence of *C. sylvestris* var. *sylvestris* in Atlantic Forest areas and *C. sylvestris* var. *lingua* in Cerrado areas, a systematic study extending this knowledge to other biomes or ecosystems, by means

Electronic reprint for personal use

▶ **Table 3** Correlation coefficients (R^2) between the average individual metabolites' content variation and their biological activities corresponded to the 12 *C. sylvestris* lyophilized extracts (n = 3).

Cytotoxicity activity				Antioxidant activity	
Compound	**HCT-116**	**PC3**	**SF-295**	**Compound**	**DPPH•**
D-1	0.16	0.37	0.24	F1	0.18
D-2	0.30	0.64	0.42	F2	0.32
D-3	0.16	0.40	0.26	F3	0.14
D-4	0.15	0.27	0.10	F4	0.04
D-5	0.27	0.58	0.37	F5	0.01
D-6	0.12	0.28	0.18	F6	0.17
D-7	0.00	0.00	0.00	F7	0.14
D-8	0.22	0.33	0.26	F8	0.29
D-9	0.01	< 0.001	0.04	F9	0.22
D-10	0.21	0.52	0.34	F10	0.32
D-11	0.1	0.24	0.17	F11	0.28
D-12	0.22	0.25	0.31	F12	0.05
D-13	0.19	0.44	0.29	F13	0.37
D-14	0.12	0.28	0.18		
D-15	0.24	0.70	0.43		
D-16	0.12	0.28	0.18		
D-17	0.12	0.28	0.18		
D-18	0.001	0.10	0.18		
D-19	0.31	0.62	0.46		
D-20	0.12	0.28	0.18		
D-21	0.08	0.30	0.28		
D-22	0.12	0.28	0.18		
D-23	0.12	0.28	0.18		
D-24	0.13	0.36	0.25		
D-25	0.12	0.35	0.24		

F-1 to F-3, unidentified flavonoid; F-4, isorhamnetin-3-*O*-trihexoside; F-5 and F-6, unidentified flavonoid; F-7, quercetin-3-*O*-rutinoside (rutin); F-8, isorhamnetin-3-*O*-neohesperidoside; F-9, unidentified flavonoid; F-10, isorhamnetin-3-*O*-rutinoside (narcissin); F-11, quercetin-3-*O*-hexoside; F-12, isorhamnetin-3-*O*-hexoside and kaempferol-3-*O*-dihexoside; F-13, isorhamnetin-3-*O*-α-L-rhamnopyranosyl-(1 → 2)-α-L-arabinopyranoside; D-1, unidentified diterpenoid; D-2, casearin L/A; D-3 to D-12, unidentified diterpenoid; D-13, casearin M/P/D; D-14, unidentified diterpenoid; D-15, casearin H; D-16, caseargewiin F; D-17, unidentified diterpenoid; D-18, casearin S; D-19 to D-21, unidentified diterpenoid; D-22, casearin X/I (derivative); D-23, casearin J; D-24, casearin X/I (derivative); D-25, casearin C



▶ **Fig. 6** Orthogonal projection to latent structures (OPLS) analysis for the chemical variation of 12 *C. sylvestris* populations and its correlation with their antioxidant activity assay.

of precise collection work, variety assigment, and detailed chemical study was still pending. Therefore, in the present study, the secondary metabolites composition from different populations of *C. sylvestris* was correlated to their varieties and geographic occurence at infraspecific level. The results confirmed the prevalence of *C. sylvestris* var. *lingua* in Cerrado areas, which correlates to the chemical profile rich in glycosylated flavonoids. Among them, narcissin is the most common metabolites for this variety. Their distinctive chemical composition determined their lower cytotoxic activity and higher antioxidant potential. On the other hand, *C. sylvestris* var. *sylvestris* have a closer association with Atlantic Forest areas and chemical composition rich in clerodane diterpenoids. Casearins D, S, and X and caseargrewiin F were correlated to their higher cytotoxic activity. Therefore, considering the diverse medicinal properties and traditional uses of *C. sylvestris* in the folkloric medicine, the certification of its medicinal properties through its chemical profiling is an urgent matter for the quality control of such material. These results highlight the need for and significance of the inclusion of geographical origin effects and the chemical profile in the quality control parameters for the medicinal use of plant materials used in traditional medicine. To do so, an extended sampling in the same geographical locations would provide an external validation for the present data in order to construct predictive models, which would help in the authentication of these botanical materials by metabolomics means.

Materials and Methods

Chemicals and materials

All solvents used for the sample preparation and chromatographic analyses were chromatographic grade. Acetonitrile, ethanol, and isopropanol were purchased from Merck. Formic acid was purchased from Sigma. Ultrapure water was obtained using a Millipore Milli-Q water purification system (Millipore). Butyl gallate (> 98.0%) was purchased from Sigma. Narcissin (> 95.0%), caseargrewiin F, and casearin X (> 90.0%) were previously isolated, purified, and identified in our laboratory (**Figs. 9S–23S**, Supporting Information). Caserin D and casearin S standards (> 90.0%) were graciously provided by Prof. Dr. André Gonzaga dos Santos from the Faculty of Pharmaceutical Sciences of Araraquara.

Electronic reprint for personal use

Plant material

The leaves of 61 trees from 12 populations of *C. sylvestris* (5 to 6 trees of each population) were collected between June and September of 2012 and 2013 across the Brazilian territory covering Cerrado (savannah), Atlantic Forest and Atlantic Forest remnant areas, Pampa, Pantanal, and ecotones (▶ **Table 1**) [28]. Voucher specimens of all samples were analyzed at the Agronomic Institute of Campinas (São Paulo State, Brazil) by Dr. Roseli B. Torres, who is the specialist on *Casearia* genus. After identity confirmation and variety assignment, vouchers No. IAC 55839 and No. IAC 55840 were deposited for *C. sylvestris* var. *lingua* and for *C. sylvestris* var. *sylvestris*, respectively. After collection, the leaves were immediately dried in an oven with air circulation at 40 °C, ground in an analytical mill with the aid of liquid nitrogen, and stored at room temperature until sample preparation.

Sample extraction and preparation

For the infraspecific chemical variability analysis, the dried leaves of each tree were individually extracted. Briefly, 50 mg of each powdered sample were extracted with 1.0 mL of water/ethanol/isopropanol (50 : 30 : 20 v/v/v) containing butyl gallate (0.5 mg/mL) as internal standard. The samples were ultra-sonicated at room temperature for 30 min and subsequently centrifuged at 5000 *g* during 5 min. Finally, 0.7 mL of the supernatant was filtered through a 0.22 µm Nylon membrane directly to 1.5 mL vials and immediately subjected to UHPLC-DAD analysis [21].

For biological activity assays, 1.0 g of the dried and powdered leaves of each tree was combined according to its population. The mixed powders of each population resulted in at least 5.0 g of a composed sample, corresponded to the 12 sampled populations. The mixtures were individually and exhaustively extracted with 20 mL of water/ethanol/isopropanol (50 : 30 : 20 v/v/v) under ultra-sonication during 30 min (3 ×). Each extract was filtered, evaporated, and lyophilized. The dry extracts were simultaneously subjected to cytotoxic and antioxidant activities assays with 3 replicates (n = 36). For chemical characterization, 15 mg of each lyophilized extract was dissolved in 1.0 mL of water/ethanol/isopropanol (50 : 30 : 20 v/v/v) containing butyl gallate (0.5 mg/mL) and subjected to chromatographic analysis for qualitative and quantitative profiling with 3 replicates (n = 36).

UHPLC-DAD parameters

The chromatographic analyses were carried out by UHPLC-DAD (Ultimate 3000RS, Dionex). The separation of the analytes was performed on a C-18 chromatographic column (Kinetex, 2.6 µm, 150 × 2.1 mm, Phenomenex). The mobile phase consisted in a mix of (A) purified water and (B) acetonitrile under the following gradient elution: 10–25 % of B from 0 to 15 min, 25–90 % of B until 35 min, 90 % of B until 40 min, and returning to the initial conditions within 2 min. The flow rate, oven temperature, and injection volume were set at 400 µL/min, 35 °C, and 2 µL, respectively. Spectral data was collected within 45 min over 200–800 nm of the absorption spectrum [21].

Quantitative analysis of glycosylated flavonoids and clerodane diterpenoids

The quantitative analysis of flavonoids and clerodane diterpenoids of the lyophilized extracts obtained from each one of the 12 *C. sylvestris* populations was performed by UHPLC-DAD. Narcissin (254 nm) and caseargrewiin F (235 nm) were selected to represent glycosylated flavonoids and clerodane diterpenoids, respectively. The solutions of narcissin and caseargrewiin F were prepared using water/ethanol/isopropanol (50 : 30 : 20 v/v/v), followed by serial dilutions to obtain concentrations from 10.0 µg/mL to 800.0 µg/mL for narcissin and 5.0 µg/mL to 400 µg/mL for caseargrewiin F. All the calibration solutions contained butyl gallate (0.5 mg/mL) as internal standard. The obtained linear correlations coefficients were: 0.999 for narcissin ($y = 0.002\,x + 0.014$) and 0.999 for caseargrewiin F ($y = 0.007\,x + 0.041$). For calculations, each peak corresponding to glycosylated flavonoids or to clerodane diterpenoids was assigned according to its characteristic UV spectra. The peak areas of all metabolites were normalized to that of the internal standard. Finally, the compound area/IS ratios were used for the quantitative analysis, being the results expressed in µg/mg of each assigned compound present in the lyophilized extracts on dry weight basis.

Compound identification/annotation

Identification of the main peaks was achieved by HRMS, MS/MS, and comparison of the target peak retention time and UV spectra with those of previously isolated standards. In detail, a representative sample containing a mixture of each one of the 12 *C. sylvestris* lyophilized extracts at concentration of 15 mg/mL was subjected to liquid chromatography coupled to mass spectrometry. The chromatographic conditions were the same described for the UHPLC-DAD analysis. The separation and identification were carried out in a LC-DAD-ESI-MS/MS (Shimadzu) coupled to a micrOTOF-Q II mass spectrometer (Bruker Daltonics). The spectrometer was equipped with an ESI source and a quadrupole-time of flight analyzer (qTOF, Bruker Daltonics). The quadrupole worked in negative mode for flavonoids and in positive mode for diterpenes (MS range of *m/z* 50–1300). The equipment was internally calibrated with trifluoroacetic acid during each run. The MS parameters were established as follows: nebulizer gas pressure, 5.0 Bar; dry gas flow, 10.5 L/min; capillary voltage, 3600 V; end plate offset: −450 V; ion source temperature, 220 °C; spectra rate acquisition, 2 spectra/s. Auto MS/MS fragmentation was carried out for the 4 most intense ions per spectrum, and it was performed applying a gradient of collision-induced dissociation energy from 20 to 65 eV according to the parent mass. In addition, the precursor ion was released after the acquisition of MS/MS spectra. All MS data was analyzed with the Bruker Compass DataAnalysis 4.3 software (Bruker Daltonics) [12]. The fragmentation pattern was defined as described in another study [29].

Evaluation of cytotoxic activity

The 12 *C. sylvestris* lyophilized extracts, caseargrewiin F, and narcissin were analyzed for their cytotoxic potential against prostate (PC-3), colon (HCT-116), and glioblastoma (SF-295) human tumor cell lines by MTT assay. Cancer lines were obtained from National Cancer Institute (USA) and maintained under sterile conditions.

Electronic reprint for personal use

Cells were plated in 96-well plates (0.1×10^6 cells/mL for PC3 and SF-295 lines, and 0.7×10^5 cells/mL for HCT-116) and incubated for 24 h to allow cell adhesion (Shel Lab CO_2 Incubator). Then, samples diluted in sterile DMSO were added to each well at a final concentration of 50 µg/mL. Negative control and the cells treated with samples were exposed to the same DMSO percentage in the well (0.1 %). After 69 h of incubation at 37 °C, 5 % CO_2, the supernatant was replaced by 150 µL fresh medium containing 10 % MTT and incubated for additional 3 h. The formazan product was finally dissolved in 150 µL DMSO, and the absorbance was measured at 595 nm (DTX 880 Multimode Detector, Beckman Coulter). Results were expressed as percentage of growth inhibition.

Evaluation of antioxidant activity

The radical scavenging capacity of the 12 *C. sylvestris* lyophilized extracts and a positive control (ascorbic acid) was evaluated by the 2,2-diphenyl-1-picrylhydrazyl (DPPH·) method. The lyophilized extracts were prepared at a concentration of 2 mg/mL and diluted in ethanol to get dilutions at 0.4, 4.0, 10.0, 20.0, 40.0, 70.0, and 100.0 µg/mL. In 96-well plates containing aliquots of 215 µL of a DPPH·ethanolic solution (0.1 mM), 35 µL each sample solution was added at different concentrations, in 3 replicates. The DPPH·ethanolic solution (215 µL; 0.1 mM) mixed with 35 µL of ethanol was used as blank. Plates were incubated in the dark for 30 min, and the absorbance values were recorded at 515 nm. The percentage of scavenged DPPH·was calculated accordingly the following equation: % DPPH·scavenging = $[Abs_{515\,nm}$ (blank) – $Abs_{515\,nm}$ (sample)/$Abs_{515\,nm}$ (blank)] × 100. The radical scavenging capacity was expressed in terms of EC_{50} (amount of antioxidant necessary to decrease the initial concentration of DPPH·by 50 %), which was calculated graphically using the analytical curves in the linear range by plotting the samples results by the corresponding scavenging effect.

Data analysis

PCA was performed using Pirouette software package (v.4 rev.1, Infometrix). For that, chromatographic data were exported on ASCII format, and the data matrices were organized with the help of OriginPro 8 (OriginLab). The chromatograms were plotted in 254 nm (from 0 to 25 min) and 235 nm (from 25 to 45 min) [21]. Peak alignment was performed using Matlab v.7.5 (MathWorks Inc.) with an implemented algorithm (www.models.kvl.dk/source/DTW_COW/index.asp). For projections, the final data set was normalized against the internal standard and mean-centered. OPLS and OPLS-DA were performed using SIMCA P software (v.14.1, Umetrics). For geographical origin effects, the variables of the X-matrix consisted of the same used for the first PCA analysis. For the biological activity correlation by OPLS, the X variables were composed by the targeted quantitative data of the assigned compounds. The biological assays results–cytotoxic and antioxidant activity–were set as Y-variables. All the models were scaled by the UV scaling method. The models were validated by cross-validation with a permutation (100 permutations) and CV-ANOVA tests. The models were considered as validated when the $Q^2 \geq 0.40$ and $p < 0.05$. The regressions for individual metabolites were constructed by plotting the metabolite content of each population in the X-axis and the activity degree in the Y-axis in Microsoft Excel 2013 using a linear regression model. The correlation coefficients (R^2 values) of each model were used as correlation parameter. When $R^2 > 0.30$, it was considered a good correlation between the metabolite content variation and the biological activity variation.

Supporting Information

Information about standards structural elucidation (UV, mass spectra, and NMR spectrum), discriminant annotations, biological assays results, and quantitative chemical analysis are available as Supporting Information.

Contributors' Statement

Conception and design of the work: P. C. P. Bueno, A. J. Cavalheiro, C. Pessoa, R. B. Torres, P. M. P. Ferreira, F. M. V. Pereira; data collection: P. C. P. Bueno, N. B. Anhesine, M. S. Giffoni, C. Pessoa, R. B. Torres, P. M. P. Ferreira, R. W. R. Sousa; analysis and interpretation of the data: P. C. P. Bueno, L. F. S. Abarca, M. S. Giffoni, A. J. Cavalheiro, C. Pessoa, R. B. Torres, F. M. V. Pereira, R. W. R. Sousa; Statistical analysis: P. C. P. Bueno, L. F. S. Abarca, N. B. Anhesine, M. S. Giffoni, A. J. Cavalheiro, P. M. P. Ferreira, F. M. V. Pereira, R. W. R. Sousa; drafting the manuscript: P. C. P. Bueno, L. F. S. Abarca, N. B. Anhesine, M. S. Giffoni, R. W. R. Sousa; critical revision of the manuscript: P. C. P. Bueno, L. F. S. Abarca, N. B. Anhesine, A. J. Cavalheiro, C. Pessoa, R. B. Torres, P. M. P. Ferreira, F. M. V. Pereira.

Acknowledgements

The present study was supported by São Paulo Research Foundation (FAPESP), CIBFar/CEPID (#2013/07600-3), SISBIOTA (#2010/52327-5), FAPESP (#2014/50945-4, #2018/18212-8, and #2019/01102-8), the Brazilian Coordination for the Improvement of Personnel in Higher Education (CAPES) (#88887136426/2017/00), and the Brazilian Council for Scientific and Technological Development (CNPq) (#563311/2010-0, #465571/2014-0 and 307328/2019-8). P. C. P. B. received a Ph.D. fellowship from FAPESP (processes #2011/21440-3, #2012/21921-4, and #2016/20295-3) and N. B. A. was supported by a scientific initiation scholarship from FAPESP (#2014/02738-0).

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Croteau R, Kutchan TM, Lewis NG. Natural Products (secondary Metabolites–chapter 24). In: Buchanan B, Gruissem W, Jones R, eds. Biochemistry & molecular Biology of Plants. Rockville: American Society of Plant Physiologists; 2000: 1367

[2] Dewick PM. Medicinal natural Products: A biosynthetic Approach, 3rd edition. Chichester: Wiley; 2009

[3] Weckwerth W, Loureiro ME, Wenzel K, Fiehn O. Differential metabolic networks unravel the effects of silent plant phenotypes. PNAS 2004; 101: 7809–7814

[4] Morgenthal K, Weckwerth W, Steuer R. Metabolomic networks in plants: transitions from pattern recognition to biological interpretation. Biosystems 2006; 83: 108–117

Electronic reprint for personal use

[5] Moore BD, Andrew RL, Külheim C, Foley WJ. Explaining intraspecific diversity in plant secondary metabolites in an ecological context. New Phytol 2014; 201: 733–750

[6] Sharma A, Cardoso-Taketa A, Choi YH, Verpoorte R, Villareal ML. A comparison on the metabolic profiling of the Mexican anxiolytic and sedative plant *Galphimia glauca* four years later. J Ethnopharmacol 2012; 141: 964–974

[7] Cavallari MM, Gimenes MA, Billot C, Torres RB, Cavalheiro AJ, Bouvet JM. Population genetic relationships between *Casearia sylvestris* (Salicaceae) varieties occurring sympatrically and allopatrically in different ecosystems in south-east Brazil. Ann Bot 2010; 106: 627–636

[8] Sleumer HO. Flora Neotropica (Monograph Number 22–Flacourtinaceae). New York: The New York Botanic Garden; 1980

[9] Marquete R, Mansano VF. O gênero *Casearia* no Brasil. Rev Biol Trop 2016; 13: 69–249

[10] Klein RM, Sleumer HO. Flacourtiaceae. In: Reitz R, ed. Flora Ilustrada Catarinense. Itajaí: Herbário Barbosa Rodrigues; 1984: 96

[11] Ferreira PMP, Costa-Lotufo LV, Moraes M, Barros FWA, Martins AMA, Cavalheiro AJ, Bolzani VS, Santos AG, Pessoa C. Folk uses and pharmacological properties of *Casearia sylvestris*: a medicinal review. An Acad Bras Ciênc 2011; 83: 1373–1384

[12] Anhesine NB, Bueno PCP, Torres RB, Lopes NP, Cavalheiro AJ. Non-polar and polar chemical profiling of six *Casearia* species (Salicaceae). Biochem Syst Ecol 2019; 87: 103954

[13] Marmitt DJ, Bitencourt S, Silva AC, Rempel C, Goettert MI. Scientific production of plant species included in the Brazilian national list of medicinal plants of interest to the unified health system (RENISUS) from 2010 to 2013. J Chem Pharm Sci 2016; 82: 123–132

[14] Carvalho PRF, Furlan M, Young MCM, Kingston DGI, Bolzani VS. Acetylated DNA-damage clerodane diterpenes from *Casearia sylvestris*. Phytochemistry 1998; 49: 1659–1662

[15] Santos AG, Ferreira PMP, Vieira Junior GM, Perez CC, Tininis AG, Silva GH, Bolzani VS, Costa-Lotufo LV, Pessoa CO, Cavalheiro AJ. Casearin X, its degradation product and other clerodane diterpenes from leaves of *Casearia sylvestris*: evaluation of cytotoxicity against normal and tumor human cells. Chem Biodivers 2007; 7: 205–215

[16] Carvalho ES, Santos AG, Cavalheiro AJ. Identificação de diterpenos clerodânicos em diferentes órgãos de *Casearia sylvestris* Sw. Rev Ciênc Farm Básica Apl 2009; 30: 277–284

[17] Wang W, Ali Z, Li X, Smillie TA, Guo D, Khan IA. New clerodane diterpenoids from *Casearia sylvestris*. Fitoterapia 2009; 80: 404–407

[18] Da Silva SL, Chaar JS, Yano T. Chemotherapeutic potential of two gallic acid derivative compounds from leaves of *C. sylvestris* Sw (Flacourtiaceae). Eur J Pharmacol 2009; 608: 76–83

[19] Da Silva SL, Calgarotto AK, Chaar J, Marangoni S. Isolation and characterization of ellagic acid derivatives isolated from *Casearia sylvestris* Sw. aqueous extract with anti-Pla2 activity. Toxicon 2008; 52: 655–666

[20] Bueno PCP, Passareli F, Anhesine NB, Torres RB, Cavalheiro AJ. Flavonoids from *Casearia sylvestris* Sw. variety *lingua* (Salicaceae). Biochem Syst Ecol 2016; 68: 23–26

[21] Bueno PCP, Pereira FMV, Torres RB, Cavalheiro AJ. Development of a comprehensive method for analysing clerodane-type diterpenes and phenolic compounds from *Casearia sylvestris* Swartz (Salicaceae) based on ultra-high-performance liquid chromatography combined with chemometric tools. J Sep Sci 2015; 38: 1649–1656

[22] Ferreira PMP, Militao GCG, Lima DJB, Costa NDJ, Machado KC, Santos AG, Cavalheiro AJ, Bolzani VS, Silva DHS, Pessoa CO. Morphological and biochemical alterations activated by antitumor clerodane diterpenes. Chem Biol Interact 2014; 222: 112–125

[23] Xia L, Guo Q, Tu P, Chai X. The genus *Casearia*: a phytochemical and pharmacological overview. Phytochem Rev 2014; 14: 99–135

[24] Itokawa H, Totsuka N, Takeya K, Watanabe K, Obata E. Antitumor principles from *Casearia sylvestris* Sw. (Flacourtiaceae), structure elucidation of new clerodane diterpenes by 2D NMR spectroscopy. Chem Pharm Bull 1998; 36: 1585–1588

[25] Santos AG, Ferreira PMP, Vieira Júnior GM, Perez CC, Tininis AG, Silva GH, Bolzani VS, Costa-Lotufo LV, Pessoa C, Cavalheiro AJ. Casearin X, its degradation product and other clerodane diterpenes from leaves of *Casearia sylvestris*: evaluation of cytotoxicity against normal and tumour human cells. Chem Biodivers 2010; 7: 205–215

[26] Napolitano DR, Mineo JR, Souza MA, Paula JE, Espindola LS, Espindola FS. Down-modulation of nitric oxide production in murine macrophages treated with crude plant extracts from the brazilian cerrado. J Ethnopharmacol 2005; 99: 37–41

[27] Wolfender J, Nuzillard J, Van der Hooft JJJ, Renault J, Bertrand S. Accelerating metabolite identification in natural product research: toward an ideal combination of liquid chromatography–high-resolution tandem mass spectrometry and NMR profiling, in slico databases, and chemometrics. Anal Chem 2019; 91: 704–742

[28] Ribeiro SM, Fratucelli EDO, Bueno PCP, Castro MKV, Francisco AA, Cavalheiro AJ, Klein MI. Antimicrobial and antibiofilm activities of *Casearia sylvestris* extracts from distinct brazilian biomes against *Streptococcus mutans* and *Candida albicans*. BMC Complement Altern Med 2019; 19: 308

[29] Danuello A, Castro RC, Pilon AC, Bueno PCP, Pivatto M, Vieira Junior GM, Carvalho FQ, Oda FB, Perez CJ, Lopes NP, Santos AG, Ifa DR, Cavalheiro AJ. Fragmentation study of clerodane diterpenes from *Casearia* species by tandem mass spectrometry (QToF and Ion Trap). Rapid Commun Mass Spectrom 2020; 34: e8781

[30] Mayr E. Systematics and the Origin of Species, from the Viewpoint of a Zoologist. Cambridge: Harvard University Press; 2009

[31] Biswal B, Joshi PN, Raval MK, Biswal UC. Photosynthesis, a global sensor of environmental stress in green plants: stress signalling and adaptation. Curr Sci 2011; 101: 47–56

[32] Givnish TJ. Adaptation to sun and shade: a whole-plant perspective. Aust J Plant Physiol 1988; 15: 63–92

[33] Ruban AV. Plants in light. Commun Intreg Biol 2009; 1: 50–55

[34] Qureshi MI, Qadir S, Zolla L. Proteomics-based dissection of stress-responsive pathways in plants. J Plant Physiol 2007; 164: 1239–1260

[35] Nijveldt RJ, Nood E, Hoorn DEC, Boelens PG, Norren K, Leeuwen PAM. Flavonoids: a review of probable mechanisms of action and potential application. Am J Clin Nutr 2001; 74: 418–425

Electronic reprint for personal use

Supplementary material

Infraspecific Chemical Variability and Biological Activity of *Casearia sylvestris* from Different Brazilian Biomes

Paula Carolina Pires Bueno[1], Luis Francisco Salomé Abarca[2], Naira Buzzo Anhesine[1], Maíra Silva Giffoni[1], Fabiola Manhas Verbi Pereira[1,3], Roseli Buzaneli Torres[4], Rayran Walter Ramos de Sousa[5], Paulo Michel Pinheiro Ferreira[5], Claudia Pessoa[6], Alberto José Cavalheiro[1]

Affiliation

[1] Institute of Chemistry, Department of Biochemistry and Organic Chemistry, São Paulo State University (UNESP), Araraquara/SP, Brazil

[2] Natural Products Laboratory, Institute of Biology, Leiden University, Leiden, The Netherlands

[3] National Institute of Alternative Technologies for Detection Toxicological Assessment and Removal of Micropollutants and Radioactive Substances (INCT-DATREM) Araraquara, São Paulo State, Brazil

[4] Agronomic Institute of Campinas, Herbarium IAC, Campinas/SP, Brazil

[5] Department of Biophysics and Physiology, Laboratory of Experimental Cancerology, Federal University of Piauí, Teresina/PI, Brazil

[6] Department of Physiology and Pharmacology, Faculty of Medicine, Federal University of Ceará, Fortaleza/CE, Brazil

Correspondence

Prof. Dr. Alberto José Cavalheiro

Institute of Chemistry

Department of Biochemistry and Organic Chemistry

São Paulo State University (UNESP)

CEP 14800-060

Araraquara

SP

Brazil

Phone: + 55 16 33019667

Fax number: + 55 16 33222308

albjcava@iq.unesp.br

 © 2020. Thieme. All rights reserved.

Table of Contents

Fig. 1S. Orthogonal projection to latent structures discriminant analysis (OPLS-DA) of 12 populations of *Casearia sylvestris* using 3 classes (*C. sylvestris* var. *sylvestris*, *C. sylvestris* var. *lingua*, and individuals showing intermediary morphology) (n= 61).

Fig. 2S. Fragmentation spectrum of casearin D.

Fig. 3S. Fragmentation spectrum of caseargrewiin F.

Fig. 4S. Fragmentation spectrum of casearin S.

Fig. 5S. Fragmentation spectrum of casearin X.

Table 1S. Quantitative analysis of flavonoids (F-1 to F-13[a]) and clerodane diterpenoids (D-1 to D-25[b]) of 12 *C. sylvestris* lyophilized extracts. Each compound/class was assigned according to their UV spectra. Putative annotations were performed as described in the main text. Results are expressed in μg/mg (n = 3).

Table 2S. Antioxidant and cytotoxic activities of 12 *C. sylvestris* lyophilized extracts.

Fig. 6S. Variable importance for the projection (VIP) plots showing the correlation between the variation of the secondary metabolites and biological activities for (**A**) cytotoxic activity on HCT-116 cells; (**B**) cytotoxic activity on PC-3 cells; (**C**) cytotoxic activity on SF-295 cells; (**D**) average cytotoxic activity; (**E**) antioxidant activity.

Fig. 7S. Average cytotoxic activity of the 12 *Casearia sylvestris* lyophilized extracts (n = 3). The bars represent ± 2 standard deviation. PAC/CE, Pacoti/CE; FLO/SC, Florianópolis/SC; PRE/SP, Presidente Venceslau/SP; GUA/CE, Guaramiranga/CE; CAM/SP, Campinas/SP; MOG/SP, Mogi-Guaçu/SP; RIO/RS, Rio Grande/RS;

CAC/MT, Cáceres/MT; CAR/CE, Cariri/CE; LUI/SP, Luis Antônio/SP; SRM/MG, São Roque de Minas/MG; ARA/SP, Araraquara/SP.

Fig. 8S. Average antioxidant activity of the 12 *Casearia. sylvestris* lyophilized extracts (n = 3). The bars represent ± 2 standard deviation. PAC/CE, Pacoti/CE; FLO/SC, Florianópolis/SC; PRE/SP, Presidente Venceslau/SP; GUA/CE, Guaramiranga/CE; CAM/SP, Campinas/SP; MOG/SP, Mogi-Guaçu/SP; RIO/RS, Rio Grande/RS; CAC/MT, Cáceres/MT; CAR/CE, Cariri/CE; LUI/SP, Luis Antônio/SP; SRM/MG, São Roque de Minas/MG; ARA/SP, Araraquara/SP.

Fig. 9S. UV spectrum of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

Fig. 10S. Mass spectrum (MS/MS) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

Fig. 11S. ^{1}H NMR spectrum (presat, 600 MHz, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

Fig. 12S. ^{13}C NMR spectrum (150 MHz, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

Fig. 13S. TOCSY 1D NMR spectrum (irradiating the signal at 5.43 ppm, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

Fig. 14S. TOCSY 1D NMR spectrum (irradiating the signal at 4.42 ppm, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

Fig. 15S. TOCSY 1D NMR spectrum (irradiating the signal at 1.04 ppm, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

Fig. 16S. HMBC NMR spectrum (expansion at anomeric protons region, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

Fig. 17S. HSQC NMR spectrum (expansion at anomeric protons region, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

Fig. 18S. UV spectrum of caseargrewiin F.

Fig. 19S. Mass spectrum (HRMS) of caseargrewiin F ($[M+K]^+$ = *m/z* 543; $[M+Na]^+$ = *m/z* 527).

Fig. 20S. ^{13}C NMR spectrum (150 MHz, MeOH-d_4) of caseargrewiin F.

Fig. 21S. UV spectrum of casearin X.

Fig. 22S. Mass spectrum (HRMS) of casearin X.

Fig. 23S. ^{13}C NMR spectrum (150 MHz, MeOH-d_4) of casearin X.

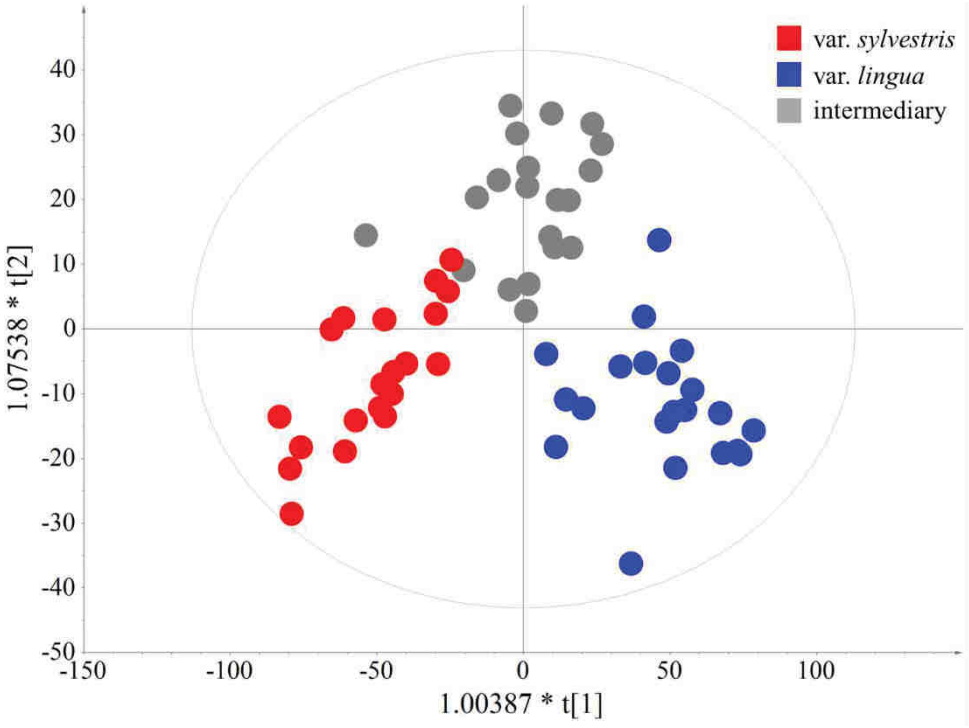


Fig. 1S. Orthogonal projection to latent structures discriminant analysis (OPLS-DA) of 12 populations of *Casearia sylvestris* using 3 classes (*C. sylvestris* var. *sylvestris*, *C. sylvestris* var. *lingua*, and individuals showing intermediary morphology) (n = 61).

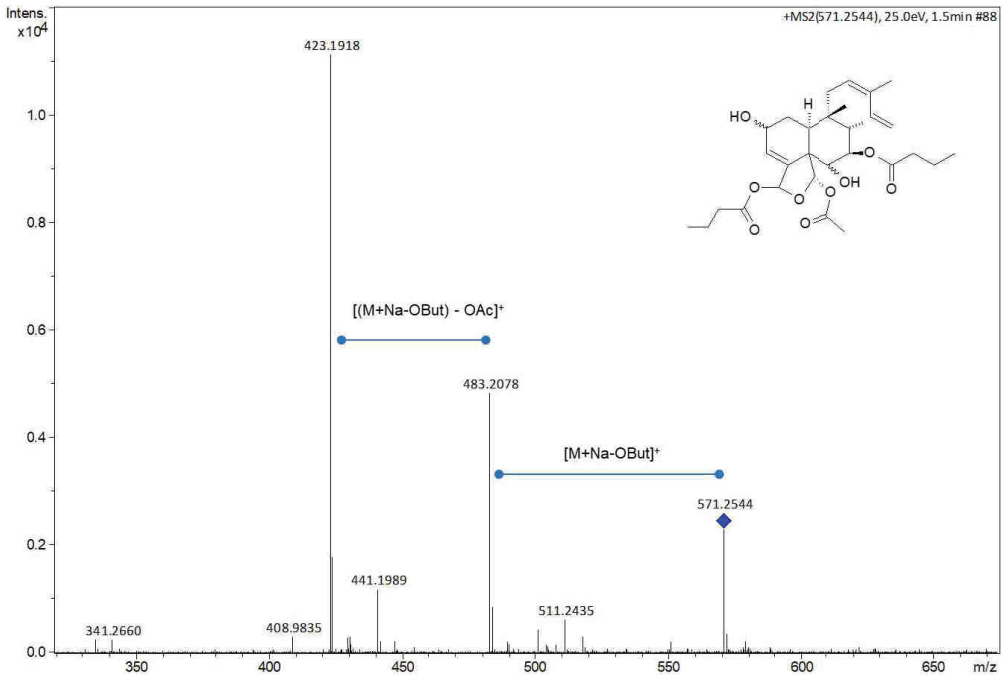


Fig. 2S. Fragmentation spectrum of casearin D.

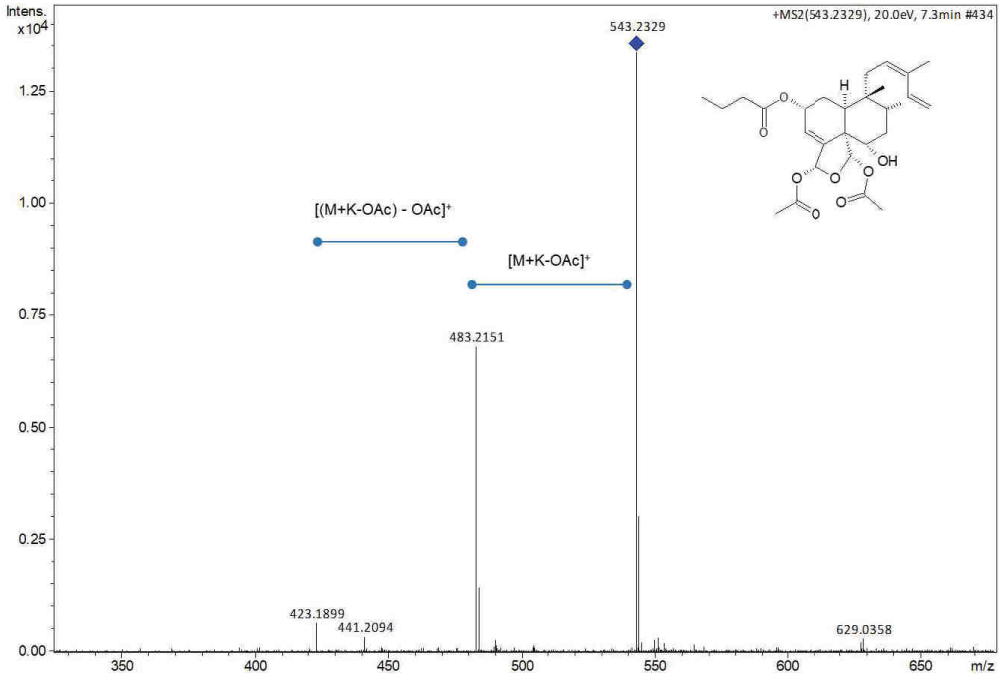


Fig. 3S. Fragmentation spectrum of caseargrewiin F.

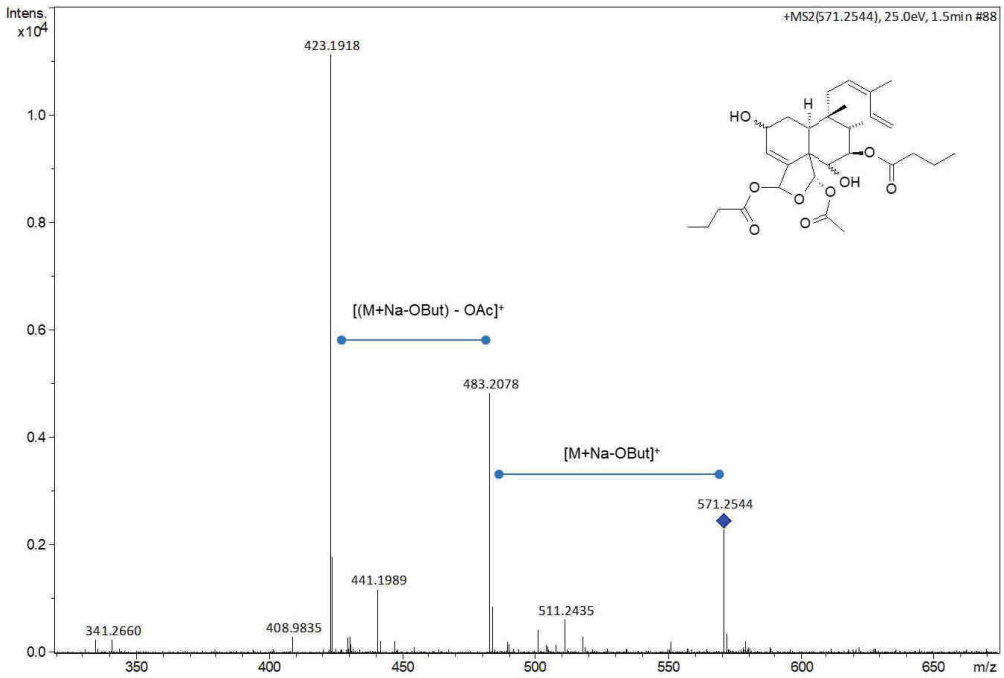


Fig. 4S. Fragmentation spectrum of casearin S.

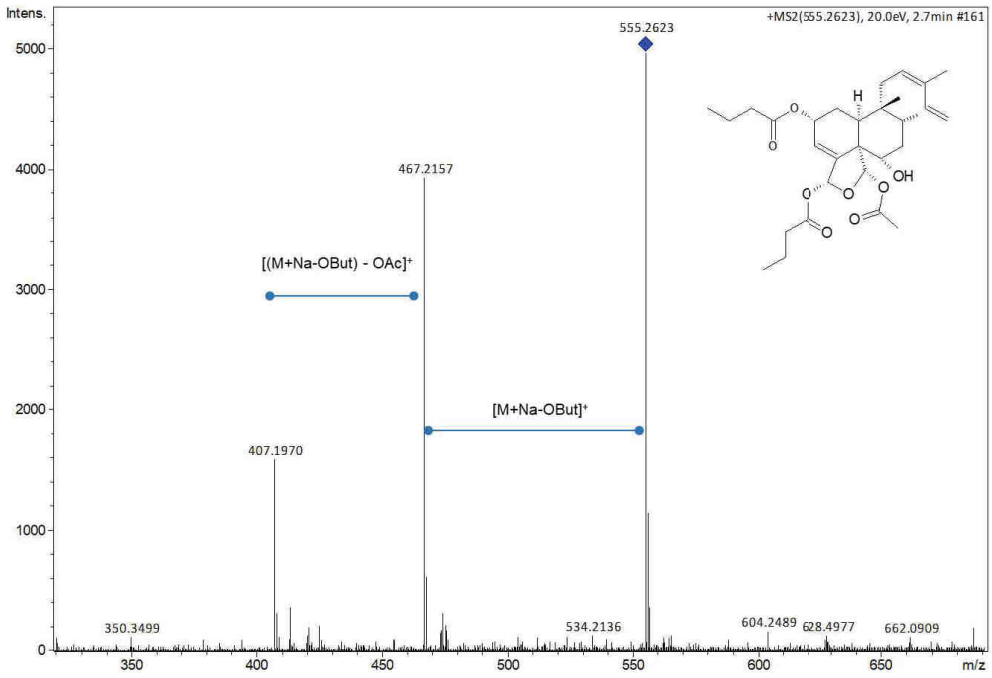


Fig. 5S. Fragmentation spectrum of casearin X.

Table 1S. Quantitative analysis of flavonoids (F-1 to F-13[a]) and clerodane diterpenoids (D-1 to D-25[b]) in the 12 *C. sylvestris* lyophilized extracts. Each compound/class was assigned according to their UV spectra. Putative annotations were performed as described in the main text. Results are expressed in µg/mg (n=3).

Compound	PAC/CE[c]	FLO/SC[d]	PRE/SP[e]	GUA/CE[f]	CAM/SP[g]	MOG/SP[h]	RIO/RS[i]	CAC/MT[j]	CAR/CE[k]	LUI/SP[l]	SRM/MG[m]	ARA/SP[n]
F-1	0.45	0.00	1.45	0.00	3.40	0.00	0.00	6.95	1.44	3.03	3.44	2.31
F-2	0.00	0.00	1.48	0.00	4.13	0.00	0.00	3.67	1.28	2.99	2.65	2.79
F-3	1.17	0.95	1.82	0.00	3.75	0.00	0.00	0.00	1.76	1.58	1.94	2.52
F-4	0.14	0.52	1.82	0.53	3.75	0.95	0.00	10.21	0.71	1.58	0.77	2.71
F-5	1.26	2.94	0.93	1.07	1.29	1.67	8.63	0.80	2.61	3.36	3.27	1.24
F-6	2.58	5.33	4.15	2.69	6.31	4.42	14.89	1.00	8.26	11.82	11.33	5.12
F-7	2.60	1.87	2.13	1.82	3.58	3.81	7.96	15.23	7.31	7.68	5.53	2.43
F-8	0.00	1.43	0.00	0.00	0.00	1.42	2.66	1.88	2.75	3.60	2.68	0.00
F-9	0.00	0.00	1.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F-10	1.77	0.00	0.00	0.00	0.00	4.99	0.00	2.44	19.30	10.07	11.00	0.00
F-11	0.00	1.52	0.00	0.00	0.00	1.33	3.12	2.15	8.41	4.15	4.26	0.00
F-12	0.25	0.90	0.66	0.35	0.70	0.72	2.28	2.65	1.53	2.04	0.00	0.47
F-13	0.00	0.00	0.00	0.00	0.00	3.87	0.00	4.42	5.01	6.65	3.04	0.00
D-1	0.05	0.00	0.70	0.08	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.04
D-2	1.15	0.15	2.01	0.31	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.31
D-3	0.41	0.00	0.02	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03
D-4	0.02	0.00	0.27	0.00	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00
D-5	0.31	0.00	0.33	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.06
D-6	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

D-7	0.00	0.51	0.98	0.49	0.07	0.00	0.11	0.00	0.00	0.00	0.00	0.21
D-8	0.00	0.00	0.00	0.00	0.00	2.22	0.00	0.00	0.00	0.00	0.00	0.00
D-9	1.05	0.00	0.16	0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.13
D-10	0.18	0.21	0.22	1.34	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.43
D-11	0.00	0.44	1.63	1.55	0.46	0.00	0.91	0.00	0.00	0.00	0.00	0.80
D-12	0.02	0.25	1.27	0.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33
D-13	1.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-14	0.25	0.00	0.35	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17
D-15	0.78	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-16	0.00	0.00	0.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-17	1.15	0.00	0.00	1.29	0.00	0.00	0.00	1.25	0.00	0.00	0.00	0.42
D-18	3.18	2.50	6.45	7.71	0.01	0.11	0.76	0.00	0.00	0.00	0.00	2.53
D-19	1.71	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01
D-20	0.55	0.00	0.00	0.35	0.00	0.00	0.00	0.19	0.00	0.00	0.00	0.11
D-21	1.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-22	1.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-23	1.10	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01
D-24	1.19	0.00	0.00	1.74	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.57
D-25	1.31	0.00	0.34	3.01	0.05	0.00	0.00	0.00	0.00	0.00	0.00	1.04

[a]F-1 to F-3, unidentified flavonoid; F-4, isorhamnetin-3-*O*-trihexoside; F-5 to F-6, unidentified flavonoid; F-7, quercetin-3-*O*-rutinoside (rutin); F-8, isorhamnetin-3-*O*-neohesperidoside; F-9, unidentified flavonoid; F-10, isorhamnetin-3-*O*-rutinoside (narcissin); F-11, quercetin-3-*O*-hexoside; F-12, isorhamnetin-3-*O*-hexoside and kaempferol-3-*O*-dihexoside; F-13, isorhamnetin-3-*O*-*α*-L-rhamnopyranosyl-(1→2)-*α*-L-arabinopyranoside; [b]D-1, unidentified diterpenoid; D-2, casearin L/A; D-3 to D-12,

unidentified diterpenoid; D-13, casearin M/P/D; D-14, unidentified diterpenoid; D-15, casearin H; D-16, caseargewiin F; D-17, unidentified diterpenoid; D-18, casearin S; D-19 to D-21, unidentified diterpenoid; D-22, casearin X/I (derivative); D-23, casearin J; D-24, casearin X/I (derivative); D-25, casearin C. [c]PAC/CE, Pacoti/CE; [d]FLO/SC, Florianópolis/SC; [e]PRE/SP, Presidente Venceslau/SP; ;[f]GUA/CE, Guaramiranga/CE [g]CAM/SP, Campinas/SP; [h]MOG/SP, Mogi-Guaçu/SP; [i]RIO/RS, Rio Grande/RS; [j]CAC/MT, Cáceres/MT; [k]CAR/CE, Cariri/CE; [l]LUI/SP, Luis Antônio/SP; [m]SRM/MG, São Roque de Minas/MG; [n]ARA/SP, Araraquara/SP.

Table 2S. Antioxidant and cytotoxic activities results of the 12 *C. sylvestris* lyophilized extracts.

Samples	Antioxidant activity (µg/mL, n = 3)		Cytotoxicity* (GI%, n = 3, samples at 50 µg/mL)					
	IC_{50}	SD	HCT-116	SD	PC3	SD	SF-295	SD
PAC/CE[a]	228.10	1.23	97.86	0.19	99.50	0.40	99.48	1.03
FLO/SC[b]	165.40	1.99	73.96	1.29	41.93	3.87	46.48	8.48
PRE/SP[c]	244.70	3.45	98.54	0.00	99.44	0.32	99.74	0.07
GUA/CE[d]	253.20	14.20	96.63	1.80	88.65	1.19	100.00	0.15
CAM/SP[e]	60.18	5.81	94.45	1.16	46.51	6.72	67.64	0.96
MOG/SP[f]	58.26	2.95	70.77	14.04	47.58	2.37	25.89	6.04
RIO/RS[g]	143.60	3.81	70.63	5.73	22.71	1.34	65.65	4.64
CAC/MT[h]	105.10	9.06	15.31	4.64	12.53	1.82	45.69	11.20
CAR/CE[i]	41.27	6.65	48.23	6.25	35.56	3.56	17.81	5.23
LUI/SP[j]	49.78	8.82	27.47	10.50	23.32	0.32	33.19	5.90
SEM/MG[k]	46.89	1.18	25.69	3.73	25.61	4.82	-1.37	10.24
ARA/SP[l]	50.75	14.71	12.35	6.89	23.15	2.92	-5.17	1.03
Ascorbic acid	22.72	0.13	-	-	-	-	-	-
Narcissin	-	-	0.28	1.67	1.24	3.40	44.34	0.59
Caseagrewiin F	-	-	103.03	0.00	100.00	0.00	97.13	0.06

*Data are presented as growth inhibition percentage (% GI) for human HCT-116 (colon), PC-3 (prostate), and SF-295 (glioblastoma) cancer cells. Caseargrewiin F was used as positive control. Experiments were performed in 3 replicates. [a]PAC/CE,

 © 2020. Thieme. All rights reserved.

Pacoti/CE; [b]FLO/SC, Florianópolis/SC; [c]PRE/SP, Presidente Venceslau/SP; [d]GUA/CE, Guaramiranga/CE; [e]CAM/SP, Campinas/SP; [f]MOG/SP, Mogi-Guaçu/SP; [g]RIO/RS, Rio Grande/RS; [h]CAC/MT, Cáceres/MT; [i]CAR/CE, Cariri/CE; [j]LUI/SP, Luis Antônio/SP; [k]SRM/MG, São Roque de Minas/MG; [l]ARA/SP, Araraquara/SP.

 © 2020. Thieme. All rights reserved.

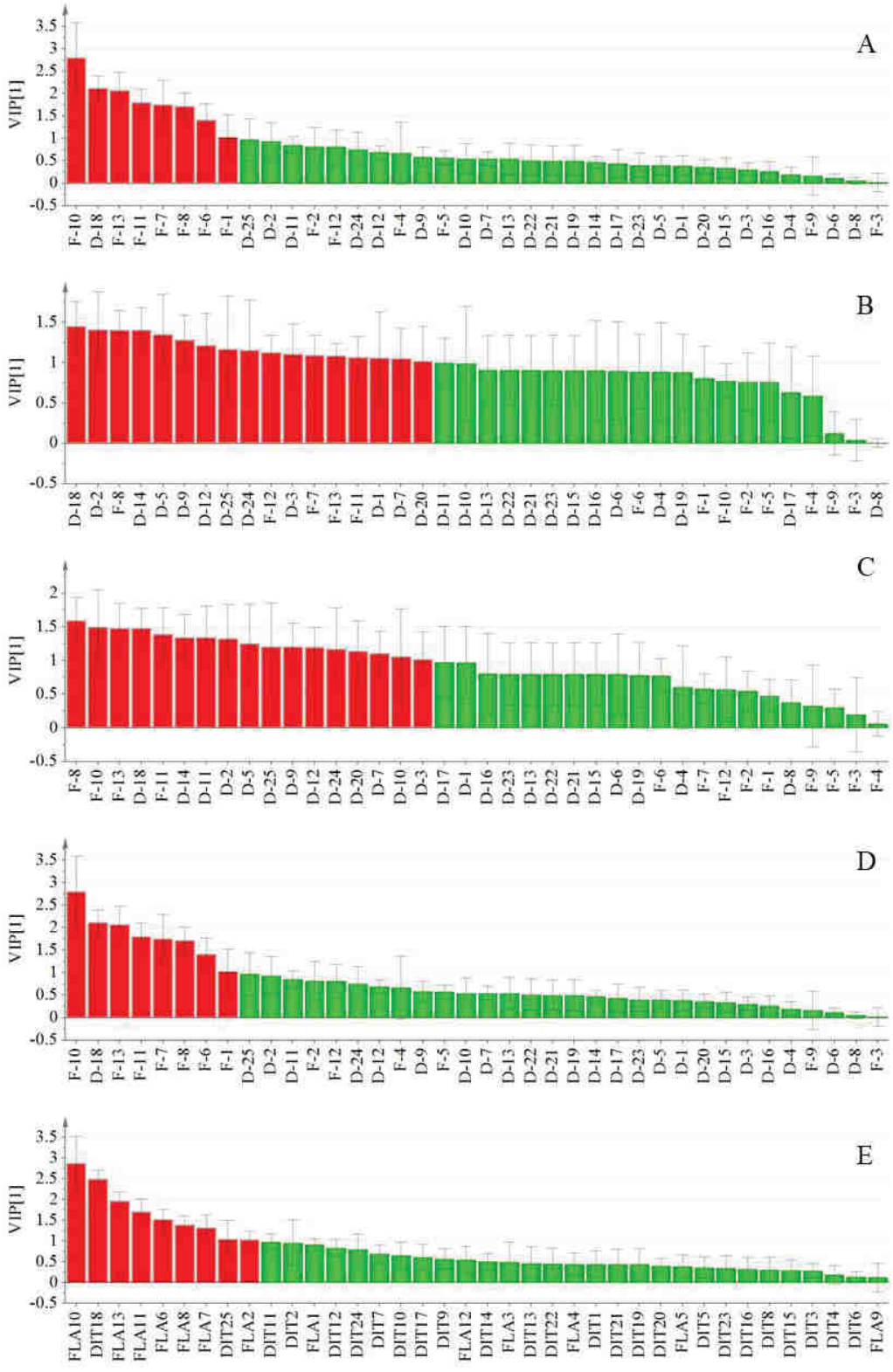


Fig. 6S. Variable importance for the projection (VIP) plots showing the correlation between the variation of the secondary metabolites and biological activities for (**A**)

cytotoxic activity on HCT-116 cells; (**B**) cytotoxic activity on PC-3 cells; (**C**) cytotoxic activity on SF-295 cells; (**D**) average cytotoxic activity; (**E**) antioxidant activity.

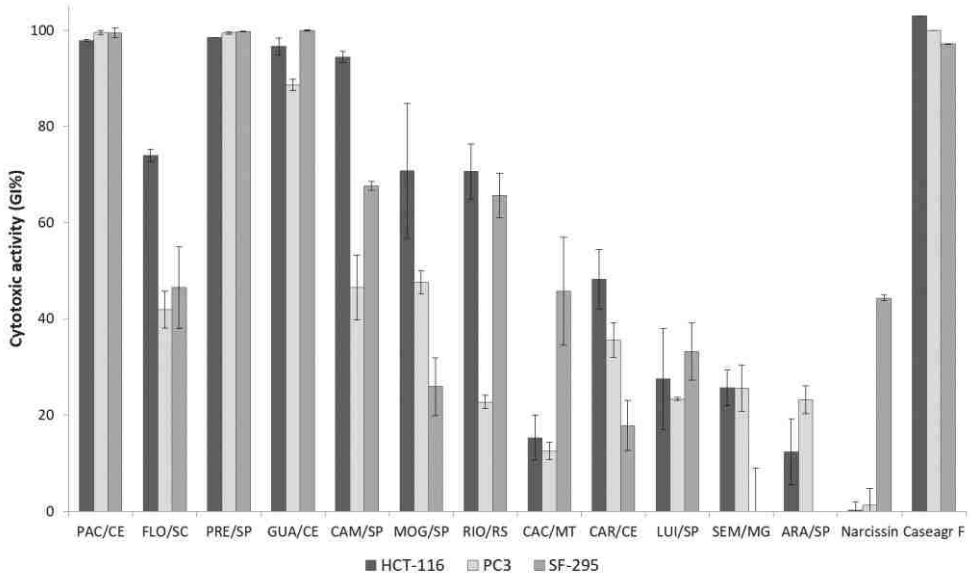


Fig. 7S. Average cytotoxic activity of the 12 *Casearia sylvestris* lyophilized extracts (n = 3). The bars represent ± 2 standard deviation. PAC/CE, Pacoti/CE; FLO/SC, Florianópolis/SC; PRE/SP, Presidente Venceslau/SP; GUA/CE, Guaramiranga/CE; CAM/SP, Campinas/SP; MOG/SP, Mogi-Guaçu/SP; RIO/RS, Rio Grande/RS; CAC/MT, Cáceres/MT; CAR/CE, Cariri/CE; LUI/SP, Luis Antônio/SP; SRM/MG, São Roque de Minas/MG; ARA/SP, Araraquara/SP.

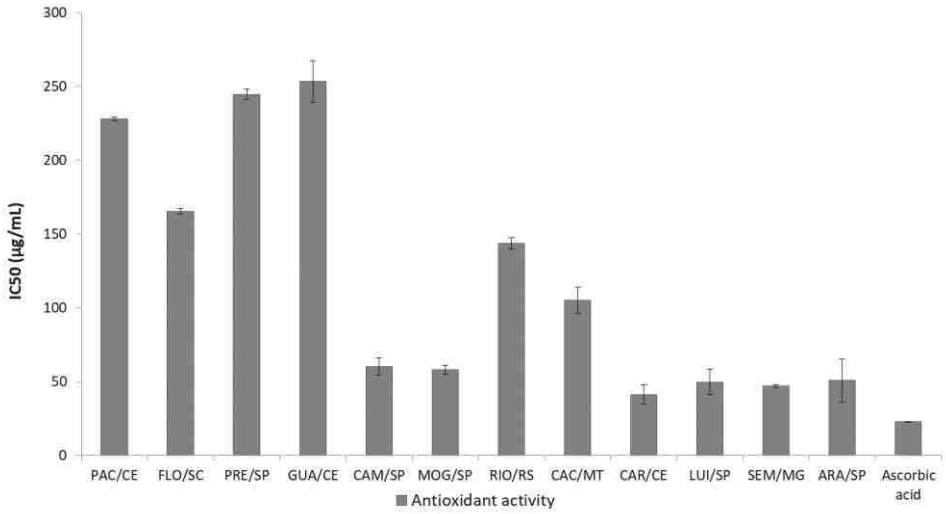


Fig. 8S. Average antioxidant activity of the 12 *Casearia sylvestris* lyophilized extracts (n = 3). The bars represent ± 2 standard deviation. PAC/CE, Pacoti/CE; FLO/SC, Florianópolis/SC; PRE/SP, Presidente Venceslau/SP; GUA/CE, Guaramiranga/CE; CAM/SP, Campinas/SP; MOG/SP, Mogi-Guaçu/SP; RIO/RS, Rio Grande/RS; CAC/MT, Cáceres/MT; CAR/CE, Cariri/CE; LUI/SP, Luis Antônio/SP; SRM/MG, São Roque de Minas/MG; ARA/SP, Araraquara/SP.

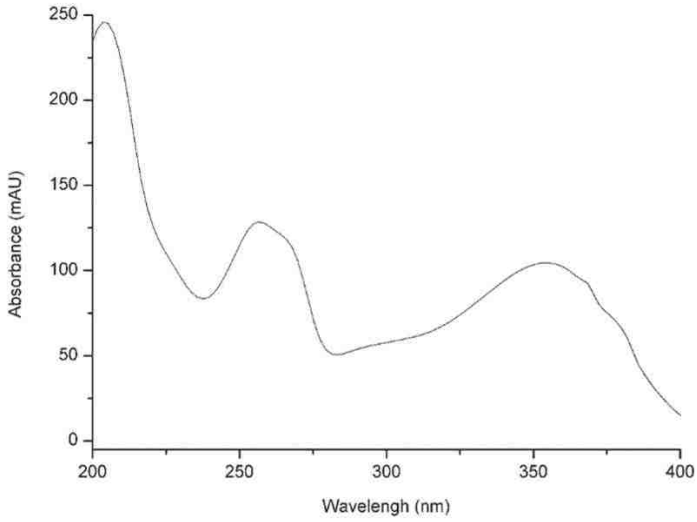


Fig. 9S. UV spectrum of isorhamnetin-*3-O-α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

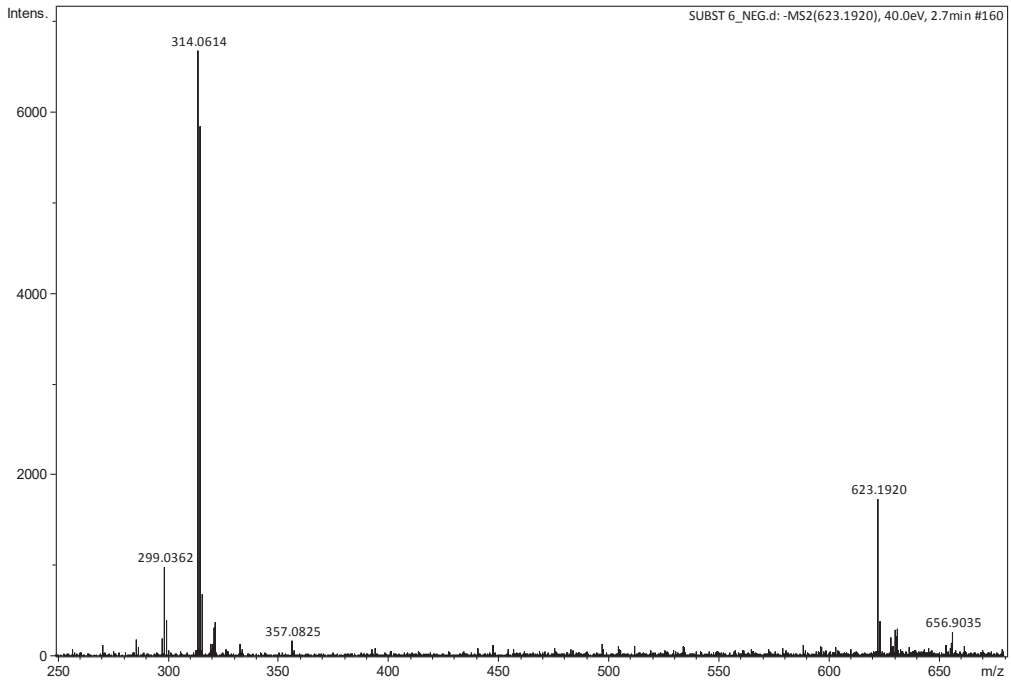


Fig. 10S. Mass spectrum (MS/MS) of isorhamnetin-*3-O-α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

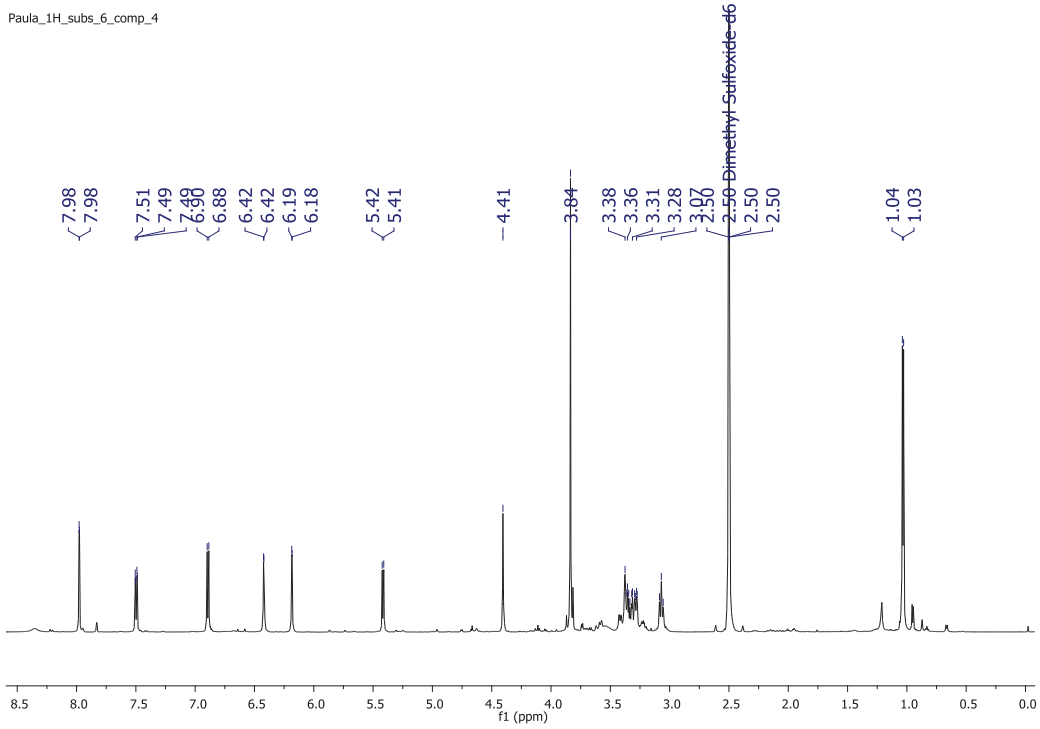


Fig. 11S. ^{1}H NMR spectrum (presat, 600 MHz, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

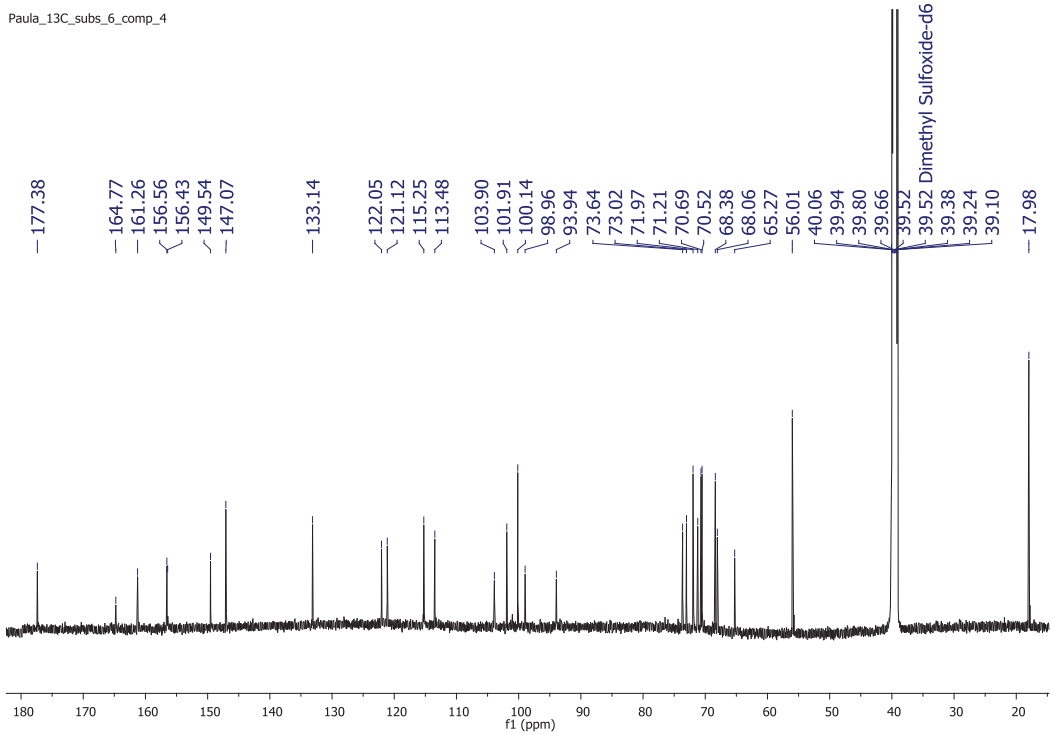


Fig. 12S. ^{13}C NMR spectrum (150 MHz, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

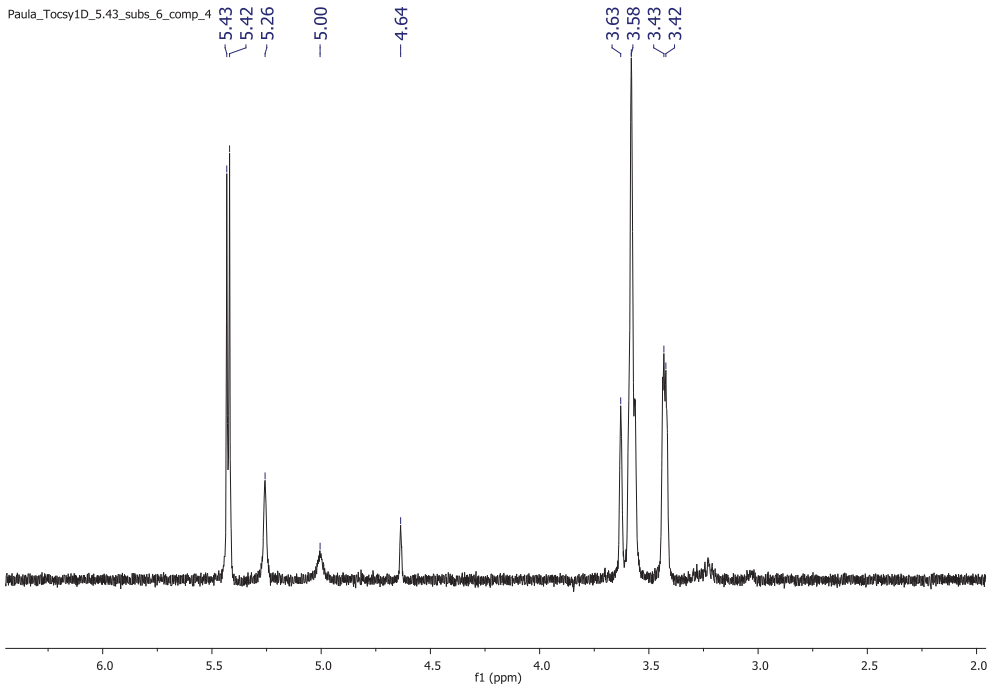


Fig. 13S. TOCSY 1D NMR spectrum (irradiating the signal at 5.43 ppm, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

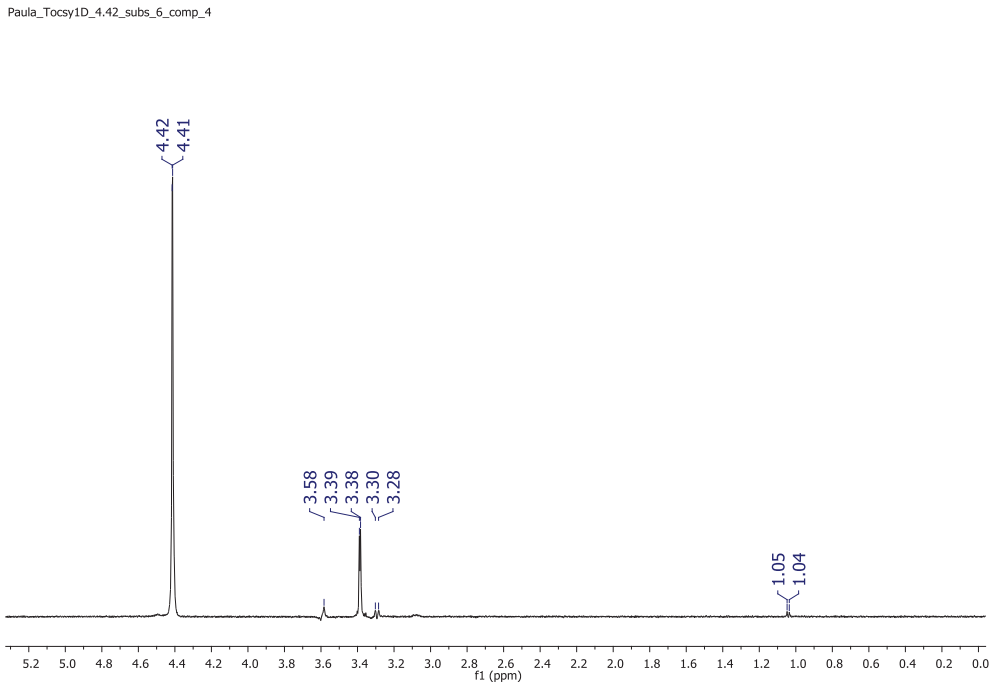


Fig. 14S. TOCSY 1D NMR spectrum (irradiating the signal at 4.42 ppm, $[CD_3]_2SO$) of isorhamnetin-*3-O-α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

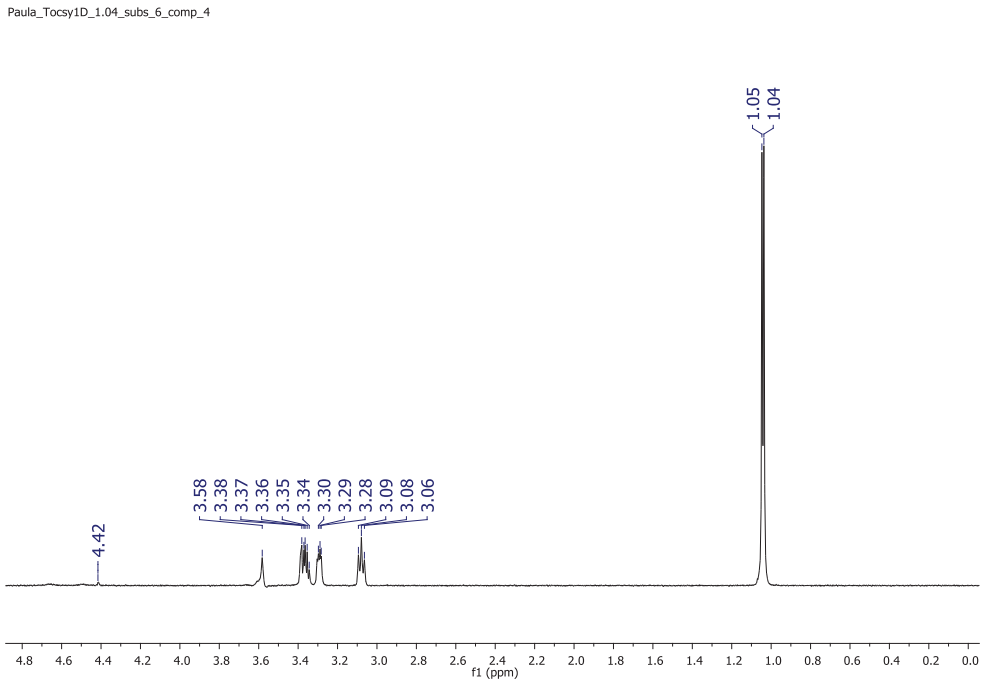


Fig. 15S. TOCSY 1D NMR spectrum (irradiating the signal at 1.04 ppm, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

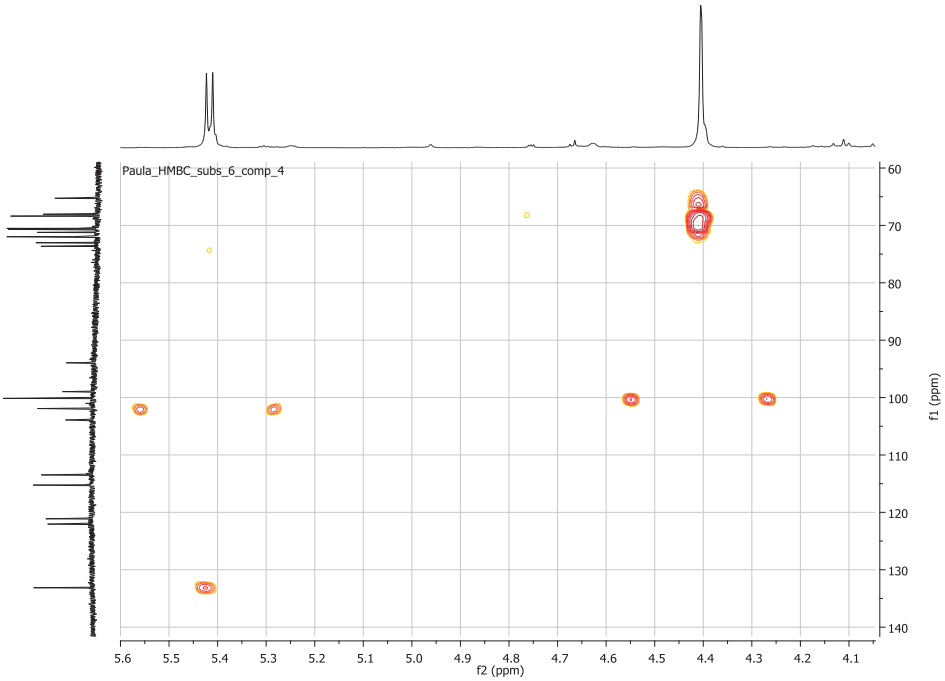


Fig. 16S. HMBC NMR spectrum (expansion at anomeric protons region, $[CD_3]_2SO$) of isorhamnetin-*3*-*O*-*α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

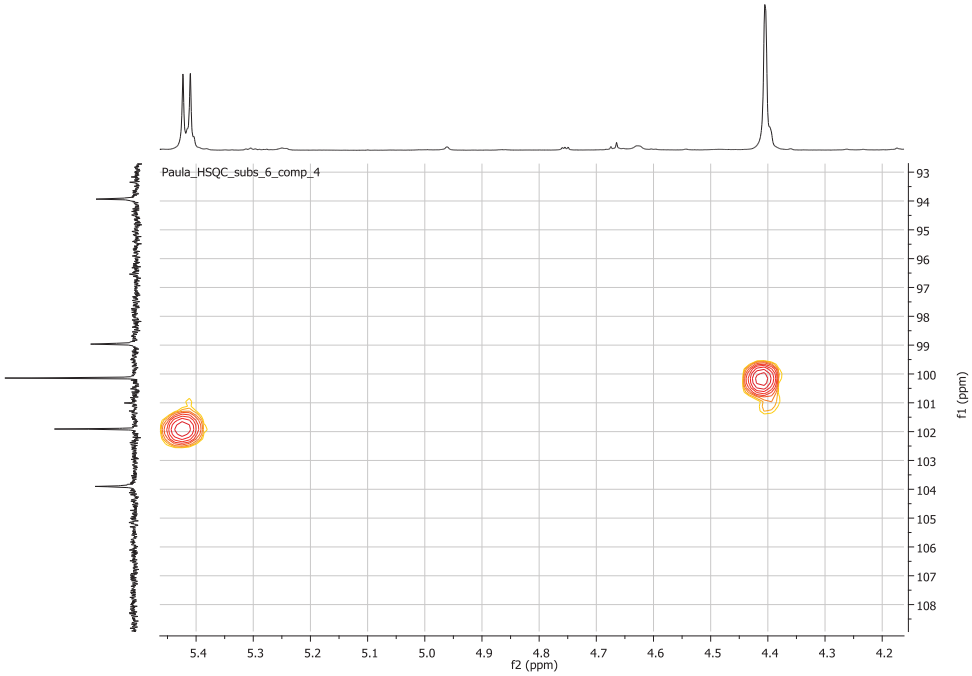


Fig. 17S. HSQC NMR spectrum (expansion at anomeric protons region, $[CD_3]_2SO$) of isorhamnetin-*3-O-α*-L-rhamnopyranosyl-(1→6)-*β*-D-glucopyranoside (narcissin).

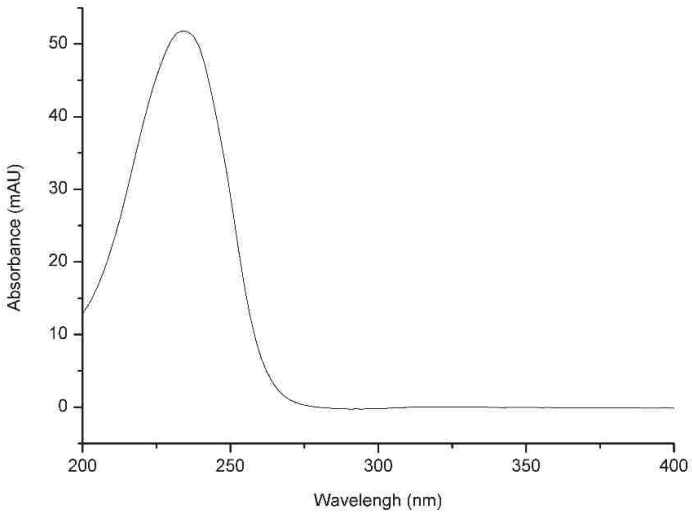


Fig. 18S. UV spectrum of caseargrewiin F.

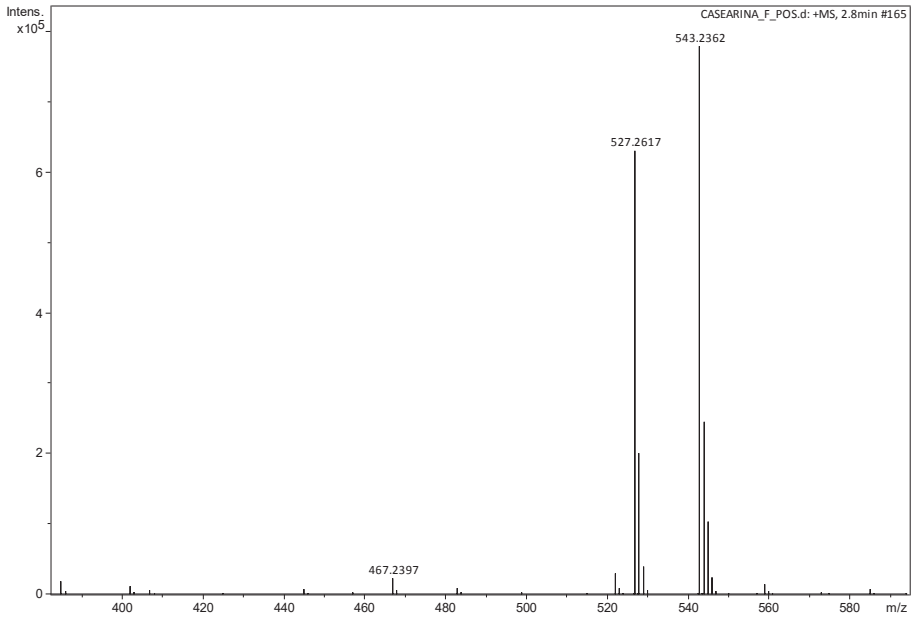


Fig. 19S. Mass spectrum (HRMS) of caseargrewiin F ($[M+K]^+$ = *m/z* 543; $[M+Na]^+$ = *m/z* 527).

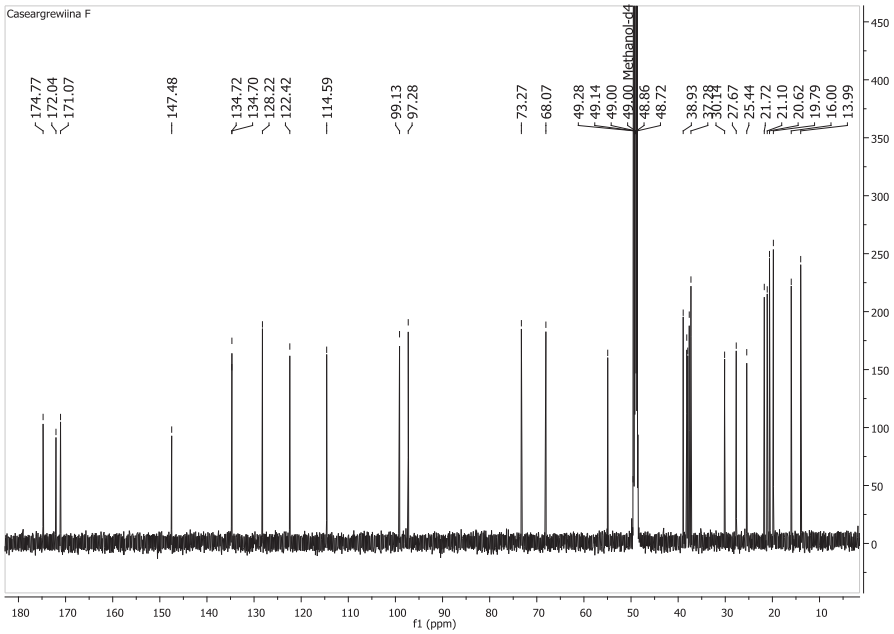


Fig. 20S. ^{13}C NMR spectrum (150 MHz, MeOH-d_4) of caseargrewiin F.

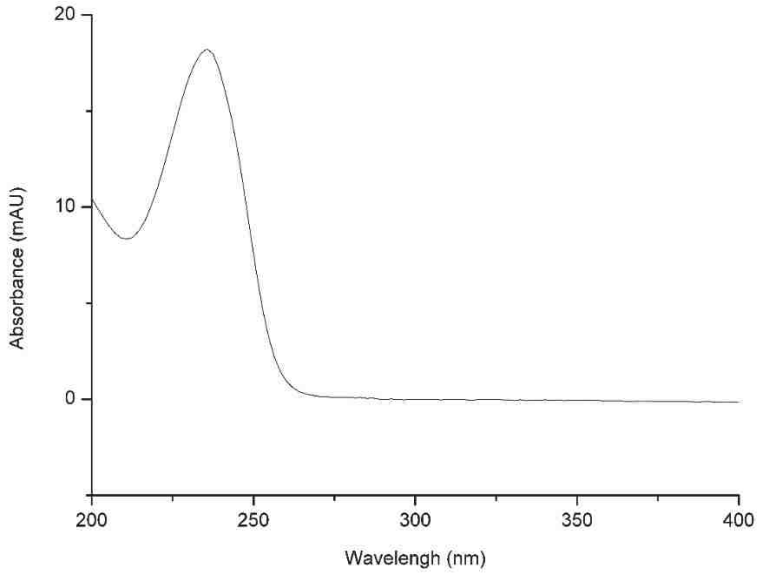


Fig. 21S. UV spectrum of casearin X.

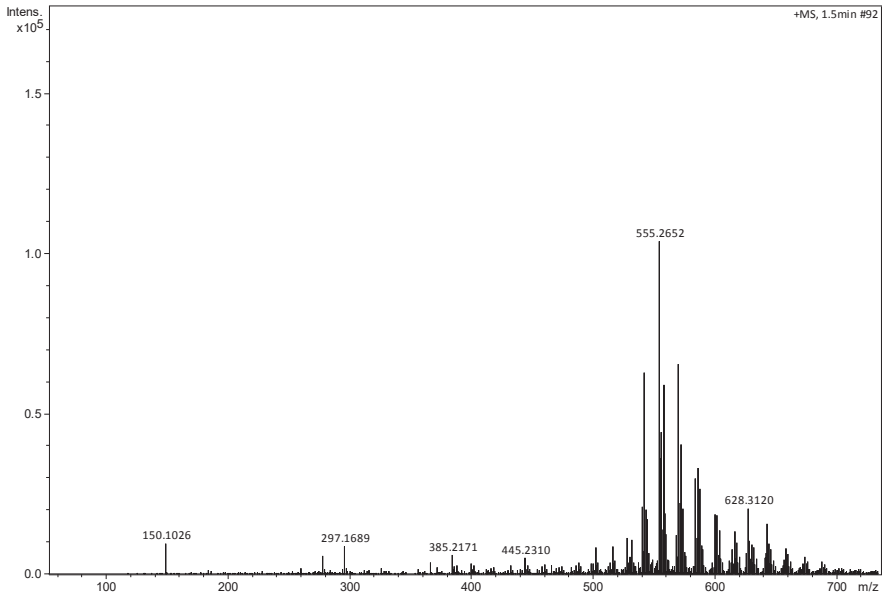


Fig. 22S. Mass spectrum (HRMS) of casearin X.

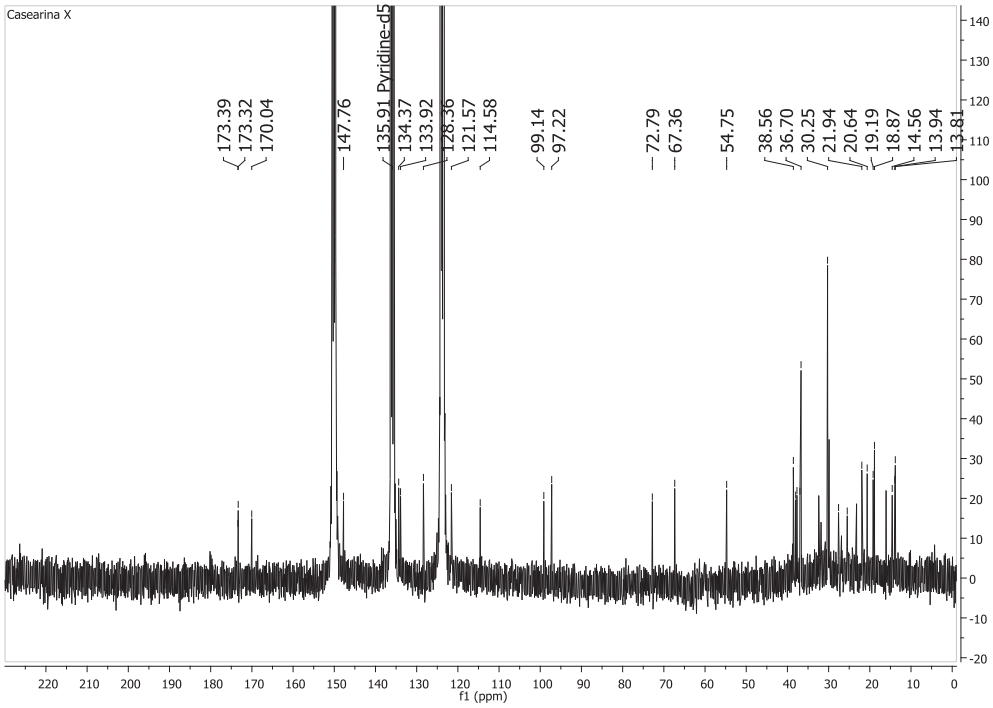


Fig. 23S. ^{13}C NMR spectrum (150 MHz, MeOH-d_4) of casearin X.