

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Revealing compositional attributes of *Glossostemon bruguieri* Desf. root geographic origin and roasting impact via chemometric modeling of SPME-GC-MS and NMR metabolite profiles

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Farag, M.A., Afifi, S.M., Rasheed, D.M., Khattab, A.R. (2021). Revealing compositional attributes of *Glossostemon bruguieri* Desf. root geographic origin and roasting impact via chemometric modeling of SPME-GC-MS and NMR metabolite profiles. JOURNAL OF FOOD COMPOSITION AND ANALYSIS, 102, 1-13 [10.1016/j.jfca.2021.104073].

Availability:

This version is available at: <https://hdl.handle.net/11585/1011702> since: 2025-03-29

Published:

DOI: <http://doi.org/10.1016/j.jfca.2021.104073>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

Revealing compositional attributes of *Glossostemon bruguieri* Desf. root geographic origin and roasting impact via chemometric modeling of SPME-GC-MS and NMR metabolite profiles

Mohamed A. Farag^{a*}, Sherif M. Afifi^b, Dalia M. Rasheed^c, Amira R. Khattab^{d}**

^aPharmacognosy Department, College of Pharmacy, Cairo University, Kasr El Aini St., P.B. 11562, Cairo, Egypt;

^bPharmacognosy Department, Faculty of Pharmacy, University of Sadat City, Sadat City 32897, Egypt;

^cPharmacognosy Department, Faculty of Pharmacy, October 6 University, Cairo, Egypt;

^dPharmacognosy Department, College of Pharmacy, Arab Academy for Science, Technology and Maritime Transport, Alexandria 1029, Egypt

***Corresponding author email:** Mohamed.farag@pharma.cu.edu

****Dr_amira_khattab@aast.edu**

1
2
3
4 **Revealing compositional attributes of *Glossostemon bruguieri* Desf. root geographic origin**
5 **and roasting impact via chemometric modeling of SPME-GC-MS and NMR metabolite**
6 **profiles**
7

8 **Running head:** *Glossostemon bruguieri* metabolomics
9
10
11

12 **Abstract**

13 *Glossostemon bruguieri* Desf. (Moghat, Sterculiaceae), is a shrub well-known for its
14 thick long-tapering dark colored roots, grown in the Middle East and North Africa. Studies
15 on *Glossostemon* have focused on its nutritive value and health benefits attributed to its
16 phenolic content with no extensive analysis of secondary metabolome. NMR metabolite
17 fingerprinting of 12 *G. bruguieri* root specimens of different origins using 1D and 2D
18 experiments identified 12 major primary and secondary metabolites. Orthogonal partial least
19 squares-discriminant analysis (OPLS-DA) of NMR aromatic regions was competent in
20 discriminating between roots origin based on moghatin and pinoresinol levels. Quantification
21 of ¹H NMR spectra revealed enrichment of Siwa specimens in sugars *i.e.*, sucrose and inositol
22 (17.89 and 12.06 µg/mg, respectively), as well as some secondary metabolites *viz.* moghatin,
23 pinorsinol, tatakina and fraxetin (7.73, 4.55, 3.81 and 3.03 µg/mg, respectively), posing it for
24 future incorporation in nutraceuticals. However, much less is known regarding aroma and
25 roasting impact on root chemical composition. Volatile profiles of 15 unroasted and roasted
26 specimens were analyzed using headspace solid-phase microextraction coupled to mass
27 spectrometry. 100 Volatile constituents were identified dominated by aldehydes, ethers and
28 mono- and sesquiterpenes, whereas esters were abundant in roasted roots. OPLS-DA
29 identified 2-ethyl-6-methylpyrazine and 5-methyl-2-furfural as roasting descriptors.

30
31 **Keywords:** *Glossostemon bruguieri*; Metabolomics; NMR; Quantification; SPME-GC-MS
32
33

1. Introduction

Glossostemon bruguieri Desf. (Syn. *Dombeya arabica* Baker) is a monotypic shrub belonging to the cacao family, Sterculiaceae, that is currently merged into Malvaceae family (Ali, 2020). Malvaceae comprises about 244 genera and probably more than 4225 species and is characterized by mucilage cavities, robust fibers with a wide global distribution (Xu and Deng, 2017). *G. bruguieri* grows wild in Iran, Iraq, and naturalized in Egypt as Moghat for its edible thick elongated tapering dark-colored roots (Al-Snafi, 2019). The roots are collected after cultivation for 1-3 years, then peeled and sun-dried for *ca.* 4–7 days prior to its consumption (Gamel et al., 2010).

The traditional use of *G. bruguieri* root is attributed to its nutritional traits. It is considered as food supplement specially for nursing women as a general tonic and a demulcent drink promoting lactation owing to its richness in macronutrients *i.e.*, starch (50%), fibers (21-33%), mucilage (12-21%), pectin (7-9%) and protein (6-8%) (Gamel et al., 2010). Aside from its nutritive value, *G. bruguieri* roots were used in folk medicine as an aphrodisiac agent and to treat gout, lupus, bruises and rheumatoid arthritis (Al-Snafi, 2019), whereas the leaves were used to relief gastrointestinal spasm (Sher and Alyemeni, 2011). The plant also revealed several therapeutic merits *viz.* acaricidal, antidiabetic, diuretic and antineoplastic activities (Al-Snafi, 2019).

The pharmacological activities of *G. bruguieri* root are correlated to its bioactive components, *i.e.*, polysaccharides, proteins, phenols (13.5%), flavonoids (1.5%) and alkaloids (5.8%) in addition to minor constituents, *viz.*, steroids, minerals and cardiac glycosides (Ghareeb et al., 2014). The protein content was much higher in seeds (19.5%) compared to roots (4.5%) versus mucilage enrichment in roots detected at 29.6 and 15.8% in roots of two and one year-old plants, respectively (Ibrahim et al., 1997). Flavonoids represented major secondary metabolite class in *G. bruguieri* root including moghatin (biflavone), takakin (Meselhy, 2003), takakin 8-*O*- β -glucoside, isoscutellarein (Sharaf et al., 1998) and apigenin (Alwhibi et al., 2017). Other constituents detected in roots included estrone (Amin et al., 1969), squalene (Alwhibi et al., 2017), 2',4-dihydroxy-4'-

methoxychalcone (Sharaf et al., 1998), sesamin and anthraquinones (emodin, chrysophanol, methoxyemodin) (Meselhy, 2003).

The aforementioned studies indicated *G. bruguieri* as a potential functional food being enriched in dietary fibers, unsaturated fatty acids, minerals and biologically-active metabolites. However, the aroma profile of roots as major determinant for its hot beverages flavor and the impact of heating/roasting on their chemical composition has never been reported before. Solid phase microextraction (SPME) is a relatively modern simple, solvent free and sensitive approach for detection of aromatic compounds at low temperature, and hence was adopted for determining the true aroma of the roots (Rasheed et al., 2020).

Further, characterization of *G. bruguieri* root extract for quality control assessment was attempted using NMR based-metabolomics which encompasses two approaches in studying plant metabolome; metabolite profiling that is used to provide comprehensive structural information in complex unresolved matrix. However, metabolite fingerprinting employs a data matrix generated from the unassigned ¹H NMR spectra in conducting a multivariate analysis for the sake of comparing the overall metabolic composition among different specimens (Farak et al., 2018).

Further, ¹H NMR spectra provides quantitative estimation of the metabolites in crude extracts based on intra-molecular proton ratios for the sake of its standardization to be incorporated in nutraceuticals (Bharti and Roy, 2012; Farak et al., 2020a; Farak et al., 2020b).

To the best of our knowledge, this is the first study aiming for untargeted metabolite fingerprinting of *G. bruguieri* roots from different origins using a multiplex approach of NMR and GC-MS analyzed using chemometric tools. Considering the magnitude of spectral data generated by hyphenated techniques, chemometric approaches for untargeted sample classification have gained prominence with wide applications in quality control analysis (Afifi et al., 2021). Unsupervised viz., PCA (principal component analysis) and HCA (hierarchical clustering analysis) in addition to supervised viz., OPLS-DA (orthogonal partial least squares-discriminant analysis) approaches for data analysis proved to be a powerful tool to achieve reliable sample classifications (Farak et al., 2021a). The purpose of this research

was also extended to the application of metabolomics tools to standardize *G. bruguieri* extract using NMR for future incorporation in nutraceuticals.

2. Materials and Methods

2.1. Plant material

12 *Glossostemon bruguieri* (Desf.) root specimens belonging to different territories in Egypt (Siwa Oasis “GBSE”, Aswan “GBAE” and Rosetta “GBRE”) and Saudi Arabia “GBSA” were selected for this study and depicted in a map shown in **Suppl. Fig. S1**. Each specimen was analyzed using three biological replicates. Roasting of GBSE samples was then accomplished by wrapping powdered roots with aluminum foil then heating them at 180°C for 15-20 min on a heating mantle (Farag et al., 2021b). Voucher specimens are deposited at the College of Pharmacy Herbarium, Cairo University, Egypt.

2.2. Chemicals and reagents

Methanol-d₄ (99.80% D) and hexamethyldisiloxane (HMDS, 99%) were purchased from Deutero GmbH (Kastellaun, Germany). For SPME sampling, polydimethylsiloxane (PDMS) or divinylbenzene-carboxen-polydimethylsiloxane (DVB-CAR-PDMS) 1 cm long fibers of 50/30 µm were obtained from Supelco (Bellefonte, PA, USA). All standards, chemicals, and solvents were obtained from Sigma Aldrich (St. Louis, Mo, USA).

2.3. Extraction procedure and sample preparation for NMR analysis

The lyophilized root samples were separately ground with a mortar and pestle. The powder (120 mg) was mixed with 5 ml 100% MeOH using a Turrax mixer (11000 RPM) for five 20 s periods. To prevent heating, a time interval of 1 min separated each mixing period. To remove plant debris, extracts were then vortexed vigorously and centrifuged at 3000 g for 30 min. For NMR analysis, 4 ml were aliquoted using a syringe and the solvent was evaporated under a stream of nitrogen till dryness. Dried extracts were resuspended with 700 µl 100% methanol-d₄ containing HMDS (0.94 mM), which served as an internal chemical shift NMR standard. After centrifugation (13000 g for 1 min), the supernatant was transferred to a 5 mm NMR tube (Farag et al., 2018). All ¹H NMR spectra for multivariate data analysis were acquired consecutively within a 48-h time interval with samples prepared immediately prior to data acquisition. Repeated control experiments after 48 h showed no additional

variation. For each specimen, three biological replicates were provided and extracted in parallel under the same conditions.

2.4. NMR analysis

All spectra were performed on an Agilent VNMRS 600 NMR spectrometer operating at a ^1H NMR frequency of 599.83 MHz with the following parameters: digital resolution 0.367 Hz/point (32 K complex data points), relaxation delay = 23.7 s; pulse width (pw) = (90°), acquisition time = 2.7 s; number of transients = 160. Zero filling up to 128 K and an exponential window function with lb = 0.4 were used before Fourier transformation. The relaxation delay of 24 s is more than 5 times the longest T1 observed for HMDS. For the quantification to be optimal, full relaxation of the protons of targeted NMR signals and the internal standard was achieved. In addition, for NMR setting for 2D-NMR experiments, 45° pulse width was used with number of transients=16. 2D NMR spectra were recorded using standard CHEMPACK 4.1 pulse sequences (gDQCOSY, gHSQCAD, gHMBCAD) implemented in Varian VNMRJ 2.2C spectrometer software. The heteronuclear single quantum coherence spectroscopy (HSQC) experiment was adjusted to $^1J_{\text{CH}} = 146$ Hz with DEPT-like editing and ^{13}C -decoupling during acquisition time. The heteronuclear multiple bond correlation (HMBC) experiment was optimized for a long-range coupling of 8 Hz; a two-step $^1J_{\text{CH}}$ filter was utilized (130–165 Hz). For NMR analysis, samples were randomly allocated in the sequence run.

2.5. NMR Quantification

For metabolite quantification using NMR spectroscopy, the peak areas of selected proton signals related to the target compounds and the internal standard (HMDS) were integrated manually for all samples. The following equation was applied for absolute metabolite level calculations:

$$m_{\text{T}} = M_{\text{T}} \times \frac{I_{\text{T}}}{I_{\text{St}}} \times \frac{X_{\text{St}}}{X_{\text{T}}} \times C_{\text{St}} \times V_{\text{St}}$$

m_{T} : mass of the target compound [μg] in the solution used for ^1H NMR measurement.

M_{T} : molecular weight of the target compound [g/mol].

I_{T} : relative integral value of the ^1H NMR signal of the target compound.

147 I_{St} : relative integral value of the 1H NMR signal of the standard compound.
 148 X_{St} : number of protons belonging to the 1H NMR signal of the standard compound.
 149 X_T : number of protons belonging to the 1H NMR signal of the target compound.
 150 C_{St} : concentration of internal standard (HMDS) in the solution used for 1H NMR
 151 measurement [mmol/L].
 152 V_{St} : volume of solution used for 1H NMR measurement [mL].

153 **2.6. NMR data processing**

154 The 1H NMR spectra were automatically Fourier transformed to [.esp] files through
 155 ACD/NMR Manager lab version 10.0 software (Toronto, Canada). The spectra were
 156 referenced to internal HMDS at 0.062 ppm for 1H NMR and to internal CD_3OD signals at 49
 157 ppm for ^{13}C -NMR, respectively (Farag, Elsebai, & Khattab, 2018). Spectral intensities were
 158 lowered to integrated regions, referred to as buckets, of equal width (0.04 ppm) within the
 159 region of $\delta=10.94$ to 0.0 ppm. The regions between $\delta = 5.0-4.7$ ppm and $\delta = 3.26-3.20$ ppm
 160 belonging to residual water and methanol signals, respectively, were removed and scaling to
 161 HMDS was performed as described in (Farag et al., 2018) prior to multivariate analyses.

162 **2.7. Headspace SPME Volatiles analysis of roots**

163 The finely pulverized root specimens (20 mg) were placed in an SPME screw cap
 164 vials. The HS-SPME volatile analysis was then performed in triplicates as reported in (Farag,
 165 Hegazi, Dokhalahy, & Khattab, 2020). The analysis of volatile compounds was conducted on
 166 a Shimadzu GC-17A gas chromatogram equipped with DB-5 column (30 m x 0.25 mm i.d.
 167 x 0.25 μm film thickness; Supelco) and coupled to Shimadzu QP5050A mass spectrometer.
 168 Gas chromatography- mass spectrometry analysis was performed under identical conditions
 169 described by (Farag, El-Kersh, Rasheed, & Heiss, 2017).

170 **2.8. GC-MS data processing**

171 Volatile components were identified by comparing their Kovats indices (KI) relative
 172 to C6-C20 n-alkane series, as well as matching of the mass spectra obtained with the NIST
 173 and WILEY libraries and with standards when available. Before mass spectral matching,
 174 peaks were first deconvoluted through AMDIS software (www.amdis.net), and their
 175 abundance data were extracted using MS Dial tool

(<http://prime.psc.riken.jp/compms/msdial/main.html>) was utilized for peak abundance extraction from all GC-MS files .

2.9. Multivariate statistical analysis

All data were expressed as mean $\pm$ S.D. of three replicates in each group. Statistical analysis was carried out by one-way analysis of variance (ANOVA) using IBM SPSS Statistics, Version 23.0 (Armonk, NY: IBM Corp). Significance values between groups were at $P < 0.05$.

All variables were mean-centered and scaled to Pareto variance. Cross validation method is employed to assess the model quality by computing a set of parameters *i.e.* the number of significant components, R², Q², pCV-ANOVA, and the permutation test to ensure no model overfit (Triba et al., 2015)

Two data matrices were generated from NMR spectral intensities and GC-MS peaks abundance and initially subjected to an exploratory analysis *i.e.* PCA (principal component analysis) and HCA (hierarchical clustering analysis) followed by applying supervised pattern recognition *i.e.* OPLS-DA (orthogonal partial least squares-discriminant analysis) using SIMCA-P Version 13.0 (Umetrics, Umeå, Sweden). OPLS-DA was employed to model the metabolite differences among the different extracts. model performance was assessed by monitoring R² and Q² values, where R² expresses the goodness of model fit, whereas Q² indicates the degree of model predictability (Farrag et al., 2019).

3. Results and discussion

3.1. H-NMR Metabolite Fingerprinting of *G. bruguieri* (Desf.) root extract

NMR-based metabolite fingerprinting is a robust technique widely employed for the identification and quantification of metabolites in crude food extracts (Farag et al., 2021c), mostly supported by the application of 2D-NMR (two dimensional nuclear magnetic resonance) analysis that aids in structural elucidation. In the present study, NMR metabolite fingerprinting of *G. bruguieri* root extracts was carried out to provide both qualitative and quantitative determination of major metabolites. ¹H NMR analysis of *G. bruguieri* root extracts revealed the characteristic signals of sugars, aliphatic and fatty acids in the up-field

region ($\delta_H = 0.0$ to 6.0 ppm), whereas the down-field region ($\delta_H = 6.0$ to 9.0 ppm) revealed for its major secondary metabolites. Assignment of the 1H NMR signals of the identified metabolites was confirmed by various 2D NMR experiments as listed in **Table (1)**, structures are illustrated in **(Fig. 1)** with representative 1H NMR spectra given in **(Fig. 2)**.

Evaluation of root 1H NMR spectra revealed for the abundance of acidic metabolites and sugar derivatives, displaying dense signals in the up-field region ($\delta_H = 0.0$ to 6.0 ppm) **(Fig. 2A)**. This observation was in accordance with previous reports highlighting the enrichment of *G. bruguieri* roots with starch, mucilage, proteins and fats (Ibrahim et al., 1997). Anomeric protons of sugars can be observed in the region ($\delta_H = 3.5$ to 5.5 ppm) of the 1H NMR spectra, showing signals characteristic for sucrose (**N1**) at (δ_H 3.61 and 5.37 ppm), β -glucose (**N2**) (δ_H 4.45 ppm), and α -glucose (**N3**) (δ_H 5.08 ppm) as well as signals indicative of the sugar alcohol i.e., inositol (**N4**) at (δ_H 3.40 and 3.70 ppm). Identified organic acids included acetic (**N6**), succinic (**N7**) and fumaric (**N8**) acids were identified by singlet signals at δ_H 1.94, 2.50 and 6.57 ppm, respectively. However, quantitative NMR examination revealed *G. bruguieri* extract enrichment in sugars, viz. sucrose and glucose, and inositol compared to organic acids **(Table 2)** which imparts its sweet and slightly tart taste (PANGBORN, 1963).

Investigation of the down-field region ($\delta_H = 6.0 - 9.0$ ppm) of the 1H NMR spectra of *Glossostemon* root extracts revealed for its major secondary metabolites **(Fig. 2B)**, with signals characteristic for the coumarin fraxetin (**N9**), bioflavonoid moghatin (**N10**), flavone tatakina (**N11**) and the furano-lignan pinoresinol (**N12**), that were in agreement with previous reports of phytochemical constituents in moghat (El-Sayed et al., 2004; Meselhy, 2003). Fraxetin, a hydroxycoumarin with 6-methoxy substitution, was readily assigned by the characteristic doublets of its olefinic protons H-3 and H-4 at δ_H 6.17 and 7.83 ppm, respectively, exerting distinct correlation peaks in 1H - 1H COSY experiment, and cross peak correlations with the carbonyl carbon (C-2) at δ_C 162.9 ppm in 1H - ^{13}C HMBC spectra. The methoxy group substitution was detected by a sharp singlet signal at δ_H 3.87 ppm **(Fig. 2A)**, correlated to the carbon at δ_C 57.1 ppm in 1H - ^{13}C HSQC experiment, and in accordance with

232 previous reports (Liu et al., 2005). The flavonoid composition of family Sterculiaceae
 233 includes mostly flavones primarily apigenin and luteolin conjugates, flavans and oligomers
 234 thereof including dimers, trimers or even tetramers (Al Muqarrabun and Ahmat, 2015). ^1H
 235 NMR spectra of moghat roots displayed signals characteristic for the biflavonoid moghatin
 236 and the flavone 4'-methoxy-isoscutellargin (tatakin), previously isolated from the plant (El-
 237 Sayed et al., 2004; Meselhy, 2003). Moghatin is 3'''-hydroxy derivative of cupressuflavone,
 238 an asymmetric dimer composed of apigenin and luteolin moieties with C-8 and C-8''
 239 interflavonoid linkage (**Fig. 1**) (Meselhy, 2003). Signals of both (C) rings were detected as
 240 sharp singlets at δ_{H} 6.61 and 6.65 ppm for H-3 and H-3'', respectively exerting long range
 241 coupling with the carbonyl carbons (C-4, C-4'') at δ_{C} 185.0 ppm in ^1H - ^{13}C HMBC
 242 examination. Another two sharp singlets in moghatin were assigned to H-6 and H-6'' at δ_{H}
 243 6.28 and 6.29 ppm, respectively showing cross peak correlations to δ_{C} 105.2 ppm (C-8, C-
 244 8'') in ^1H - ^{13}C HMBC spectra. Absence of any ^1H NMR signals around δ_{H} 6.40 ppm, together
 245 with the observed downfield chemical shift of carbons (C-8, C-8'') compared to that of
 246 equivalent carbons in respective monomers, viz. luteolin and apigenin, affirmed for their
 247 involvement in the inter-flavonyl linkage (Meselhy, 2003), and supporting the dimeric form.
 248 The doublets at δ_{H} 6.90 and 8.01 ppm (dd, J = 8.8 and 2.2 Hz), exhibited correlation peaks in
 249 ^1H - ^1H COSY experiment, implying an AX coupling system attributed to H-3'/5' and H-2'/6',
 250 respectively of the 1',4'-disubstituted B- ring of apigenin moiety. B- ring protons of luteolin
 251 moiety were assigned to the signals at δ_{H} 7.60 (d, J = 2.2 Hz, H-2'''), δ_{H} 6.87 (d, J = 8.4 Hz,
 252 H-5''') and δ_{H} 7.42 (d, J = 8.4 Hz, H-6'''), coupled to δ_{C} 113.6, 116.1 and 120.1 ppm,
 253 respectively in ^1H - ^{13}C HSQC spectra. The flavone compound tatakin was likewise detected
 254 from its singlet signals at δ_{H} 6.61 and 6.28 ppm, assigned to H-3 and H-6, respectively which
 255 exerted cross peak correlations to δ_{C} 185 and 122.8 ppm attributed to C-4 and C-8,
 256 respectively in ^1H - ^{13}C HMBC experiments. Another AX spin system was observed at δ_{H} 7.06
 257 and 8.11 ppm (dd, J = 8.8 and 2.2 Hz) with obvious correlations in ^1H - ^1H COSY spectra. The
 258 sharp singlet signal at δ_{H} 3.66 ppm (**Fig. 2A**), correlated to the carbon at δ_{C} 54.1 confirmed
 259 the methoxy substitution at (C-4'), and previous ^1H NMR analysis of tatakin (Meselhy, 2003).

Previous chemical analysis of moghat roots reported the presence of furano-lignans namely sesamin (Meselhy, 2003), not detected in this study but rather its precursor pinoresinol (Pathak et al., 2019). Assignment of the aromatic protons resonating at δ_H 6.69, 6.73 and 6.77 ppm, as well as the methylene protons CH_2 -4/8 at δ_H 3.41 and 3.76 ppm assigned to δ_C 71.5 ppm in 1H - ^{13}C HSQC and δ_C 84.6 ppm in 1H - ^{13}C HMBC experiments, were in accordance to that reported for pinoresinol (Roy et al., 2002). Protons H-1/5 appeared as multiplets at δ_H 2.94 ppm as a result of coupling with α and β methylene protons (CH_2 -4/8) as confirmed by 1H - 1H COSY correlations.

3.2. Multivariate data analysis of 1H NMR fingerprinting dataset

Multivariate analysis of the 1H NMR -derived fingerprints was performed to pinpoint for differences in metabolites abundance in an untargeted manner.

3.2.1. Unsupervised PCA analysis of full range of 1H NMR dataset

Principal component analysis (PCA) was first performed for all root samples using the full region of 1H NMR (δ 10.0 to δ 0.0 ppm). The generated model (**Fig. 3**) accounted for a sample variance of 82.3%. A good reproducibility for the extraction method was also confirmed by the close grouping of each sample replicates in the score plot (**Fig. 3A**) hence, confirming the low technical and preparation method variability.

By examining the obtained PCA score plot (**Fig. 3A**), the two Egyptian *G. bruguieri* (Desf.) roots samples *i.e.* GBSE and GBRE obtained from *i.e.* Siwa and Rosetta, respectively were clustered apart from the other two samples *i.e.* GBAE and GBSA which were collected from Aswan, Egypt and Saudi Arabia (*i.e.* located at same latitudes of 24.0908° N and 24.68773° N, respectively) (<https://www.geodatos.net/en/coordinates/egypt/aswan>).

The most influential variables for such segregation were revealed from the corresponding PCA loading plot (**Fig. 3B**) assigned to 1H NMR signals appeared at δ_H 3.73 ppm for methoxy group of pinoresinol that was more abundant in GBSE specimens compared to the other specimens. HCA dendrogram (**Fig. 3C**) depicted a similar sample segregation pattern as that observed in the PCA model which showed an overlap between GBAE and GBSA specimens owing to their comparable metabolic makeup.

3.2.2. Unsupervised PCA analysis of aromatic region of ¹H NMR dataset

PCA analysis was then performed for all root samples using only the aromatic ¹H NMR region (δ 10.0 to δ 5.5 ppm) to reveal for distinctive secondary metabolites that could provide better sample classifications (**Suppl. Fig. S2**). The established model explained 87.6% of the variance (**Suppl. Fig. S2A**) which is comparable to the amount of explained variance (82.3%) when the full range of ¹H NMR was modeled (**Fig.3**); however, a different segregation profile was obtained. Differentiation between both GBAE and GBSA from the other samples was attained along PC1 as explained by the loading plot (**Suppl. Fig. S2B**) that revealed for the enrichment of GBAE and GBSA in moghatin and pinoresinol as evidenced from ¹H NMR signal at δ_H 6.69 assigned to CH-2'/2'' of pinoresinol and the other two signals at δ_H 6.87 and 8.01 corresponding to CH-5''' and CH-2'/6' of moghatin. While GBSE and GBRE samples were clustered along the negative side of PC1 and found to be most enriched in fraxetin as implied from its ¹H NMR signals at δ_H 7.83 ppm, respectively, which possessed the highest negative scores over PC1.

3.2.3. Supervised multivariate OPLS-DA analysis of ¹H NMR dataset

To obtain geographic origin descriptive markers of the four *G. bruguieri* (Desf.) root samples, partial least squares (PLS) regression analysis was primarily employed but no valid model was obtained (data not shown). Accordingly, orthogonal partial least squares-discriminant analysis (OPLS-DA) was adopted as it outperforms PLS by maximizing the separation between the modelled classes especially when low within-class variability exists (Wiklund et al., 2008). The four *G. bruguieri* (Desf.) root samples were modeled by OPLS-DA against one another using aromatic region data (δ 10.0-5.5 ppm) (**Fig.4**). The derived score plot (**Fig.4A**) showed a clear separation between each sample belonging to a particular territory with goodness of model fit “R²” of 0.978 and a prediction goodness parameter “Q²” of 0.939. RMSEC and RMSEcv were also computed during cross and found acceptable as not exceeding 0.127. The OPLS-DA score plot (**Fig.4A**) displayed a much better discrimination between the different *G. bruguieri* (Desf.) root samples.

As indicated by PCA model (**Suppl. Fig. S2**), the loading plot (**Fig.4B**) identified moghatin and pinoresinol as the discriminatory markers for both GBSA and GBAE from

other samples GBSE and GBRE mostly attributed to fraxetin abundance in GBSE and GBRE. Fraxetin is recognized for its strong antioxidant effect and in a recent study showed anti-inflammatory effect, anti-fibrotic and anti-apoptotic effects in acute pancreatitis (Balaha et al., 2021; Thuong et al., 2009). On the other hand, pinoresinol, being a phytoestrogen, it possessed anti-tumor activities against different types of breast cancer (López-Biedma et al., 2016b). The enrichment of these phytoconstituents poses moghat as a potential supplement for nursing women and to be considered for post menopause management.

3.2.4. Quantitative NMR of major *G. bruguieri* (Desf.) root metabolites

Quantitative NMR (qNMR) was employed to determine the absolute quantities of *G. bruguieri* metabolites based on the quantification protocol reported in (Farg et al., 2018; Rasheed et al., 2018), and for future standardization purposes in case to be incorporated in nutraceuticals. The absolute amounts of different metabolites ($\mu\text{g}/\text{mg}$) were further analyzed using one-way analysis of variance (ANOVA) to reveal for significant differences among different specimens. Differential relative metabolite levels were recorded in **Table 2**, where sucrose was the most abundant metabolite in all the samples with its level ranging from 7.7 to 17.9 $\mu\text{g}/\text{mg}$. The highest levels of sucrose were observed in GBSE and GBRE (*ca.* 17.9 and 17.0 $\mu\text{g}/\text{mg}$) at $P \leq 0.05$. GBSE samples were also more abundant in other sugars *i.e.* β - and α -anomers of glucose and the sugar alcohol *i.e.* inositol detected at a significantly high levels of 5, 3.9 and 12.1 $\mu\text{g}/\text{mg}$, respectively at $P \leq 0.05$.

OPLS-DA model (**Fig. 4**) revealed the abundance of fraxetin in the samples of GBSE and GBRE which both contained 3 $\mu\text{g}/\text{mg}$, as compared to slightly lower levels in GBAE and GBSA at 2.2 and 2.8 $\mu\text{g}/\text{mg}$, respectively. In contrast, tatakina was recorded at its highest level in GBSE (*ca.* 3.7), followed by GBSA, GBRE and GBAE at 3.6, 3.6 and 2.7 $\mu\text{g}/\text{mg}$, respectively. As revealed from PCA model (**Fig. 3**), the highest level of pinoresinol was also recorded in GBSE (*ca.* 4.55 $\mu\text{g}/\text{mg}$) followed by GBSA, GBRE and GBAE at 4.27, 4.23 and 3.17 $\mu\text{g}/\text{mg}$, respectively. Pinoresinol was reported to exhibit anti-tumor effect in breast cancer through its action on estrogen receptors (López-Biedma et al., 2016a) which is an adding value to its use by nursing women post-delivery.

GBSE and GBSA showed the highest moghatin level at 7.7 and 7.6 µg/mg, respectively. Hence, among examined specimens, GBSE sample obtained from Siwa, Egypt appeared the most enriched root sample in both primary and secondary metabolites and this highlighted its high nutraceutical value. Siwa is an isolated oasis in western Egypt desert and hence has been interbred less (Khalil et al., 2017), that account for the unique metabolite composition of its indigenous plants.

3.3. GC-MS peaks identification

Whilst NMR was powerful to identify major primary and secondary metabolites, no hint was made towards aroma compounds in *G. bruguieri*. Solid phase microextraction (SPME) coupled to GC-MS was implemented to assess from aroma profile of *G. bruguieri* (Desf.) roots belonging to different geographical origins as well as in response to roasting. GC-MS chromatograms (Fig. 5) revealed 100 volatiles (Table 3) identifying *G. bruguieri* root flavor to be dominated by volatile classes viz., aldehydes, esters, alcohols and ethers/oxides, with qualitative differences among samples from different locations.

ANOVA revealed significant differences in the percentile levels of aldehydes i.e. α -citral, (*E*)-cinnamaldehyde and 1-pentadecanal, aliphatic hydrocarbons viz. Tetradecane, 2,6,10-trimethylhexadecane and octadecane, ethers/oxides viz. cineol and anethole, ketones/lactones viz. 1-hydroxypropan-2-one and coumarin among the different root specimens with a confidence level of 95%.

GBSE_R contained furans i.e. 5-Methyl-2-furfural and pyrazines/pyrroles i.e. methylpyrazine, 2,6-dimethyl pyrazine, 2-ethyl-6-methylpyrazine, 2,3,5-trimethyl pyrazine and 3-ethyl-2,5-dimethylpyrazine at a significantly higher levels compared to other root specimens at $p < 0.05$. GBAE contained a significantly higher levels of monoterpenes viz. β -pinene, *o*- and *p*-cymenene and sesquiterpenes viz. (*Z*) and (*E*)- β -farnesene, cedrene and β -bisabolene than other root specimens at $p < 0.05$.

3.3.1. Esters

Esters represented the most abundant volatile class in roasted roots from Siwa, Egypt (GBSE_R) amounting for 51% followed by specimens from Saudi Arabia (GBSA) and Aswan (GBAE) at ca. 26 and 22%, respectively. Most of the detected esters were derived

from organic *i.e.*, acetic or fatty acids *i.e.*, palmitic acids with a direct relationship observed between the levels of acetic acid (1) and its corresponding esters *i.e.* Z-3-hexenol acetate (39), ethyl acetate (37) and acetonyl acetate (38) suggesting for their close biosynthetic relationship and/or activation of acetyl transferase (Lilly et al., 2006). The predominant ester *i.e.* Z-3-Hexenol acetate (39) in GBSE_R (38%) and GBAE (18%) imparts a walnut husk and banana notes exhibiting fruity and green flavor (Dhifi et al., 2005), which dropped to *ca.* 5% in unroasted *G. bruguieri* from Siwa (GBSE) and Rosetta (GBRE), respectively. Ethyl acetate (37) was the major ester in GBSA at 12.7% likely to impart a sour taste at high levels (Lyumugabe et al., 2014), versus to a fruity, pineapple odor as the case in *G. bruguieri* from Aswan (2.7%) (Chen et al., 2006). Fatty acyl esters were represented by methyl palmitate (44) at 4.5% in GBSE.

3.3.2. Aldehydes

Aldehydes were the major class in GBRE from Rosetta (49.8%) attributed to the abundance in (*E*)-cinnamaldehyde (25) at 46%. (*E*)-Cinnamaldehyde (25) imparts a cinnamon like flavor and is likely to contribute to the anti-microbial activity of *G. bruguieri* root (Firmino et al., 2018). In contrast, cumaldehyde (22) amounted for the predominant aldehyde in GBAE and GBSA at 6.0 and 3.2%, respectively, and to provide a strong, pungent, cumin-like aroma (Cui et al., 2019). 1-pentadecanal (27) was found exclusively at 3.4% in GBAE from Aswan suggestive to serve as a potential marker for this specimen with a floral and waxy note.

3.3.3. Ethers/oxides

Ethers/oxides amounted for 25.3% in GBSE from Siwa, and declined to 1% in its roasted counterpart (GBSE_R). The high levels of this class was also detected in GBAE and GBSA exemplified by (*E*)-anethole (48) and safrole (49). Anethole (48), the main fragrance of anise and sweet fennel, is characterized by sweet, distinct, anise-like flavor (Wang et al., 2020), whereas safrole (49) is typified by a 'candy type' aroma.

3.3.4. Alcohols

The highest level of alcohols was detected in GBSA from Saudi Arabia at 26% exemplified by phenethyl alcohol (13.4%). Phenethyl alcohol (7) is a fragrance ingredient with a floral, rose-like aroma (Li et al., 2014). 1-Tetradecanol (5) or myristyl alcohol, was the major alcohol in GBAE from Aswan at 5.1%.

3.3.5. Mono- /sesqui-terpene hydrocarbons

The highest content of monoterpene hydrocarbons was detected in GBAE (13.5%), followed by GBSA at 6.4%. β -Myrcene and β -pinene were the major monoterpenes in GBAE reaching ca. 4%. β -Myrcene (70) is a pleasant volatile with analgesic, antioxidant and anti-ulcer effects (Ciftci et al., 2014), whereas β -pinene (69) is common in essential oils with piney, spicy, and turpentine-like note.

Compared to monoterpenes, the highest sesquiterpenes level was detected in GBRE from Rosetta at 13.6% followed by GBSE and GBAE at 9.0 and 7.8%, respectively. β -Caryophyllene (95) was the dominant sesquiterpene in GBSE from Siwa at 8.5%, whereas calamenene (99) was found almost exclusively (7.1%) in GBRE.

3.3.6. Ketones /lactones

Ketones/lactones amounted for 8% in GBSE followed by its roasted counterpart (GBSE_R) at 5.7%. Coumarin (66) and α -isomethyl ionone (67) were the most abundant in GBSE at 3.2 and 2.0%, respectively. Coumarin (66) and α -isomethyl ionone (67) were reported as potentially hazardous compounds with limited applications as fragrant agents (Avonto et al., 2018).

3.3.7. Furans

Roasted GBSE_R from Siwa was enriched in furans at 8.4% exemplified by furfural (3.9%) and 5-methyl-2-furfural (3.0%). Furfural (54) provided a caramel-like, sweet aroma (Leffingwell and Alford, 2005) and is a common ingredient in roasted food *i.e.*, coffee, whereas 5-methyl-2-furfural (56) demonstrated bitter almond flavor (Pico et al., 2015).

3.3.8. Pyrazines/pyrroles/nitrogenous compounds

Being generated mainly from a Maillard reaction between sugars and amino acids, pyrazines also contribute substantially to aroma developed during roasting, heating and

toasting of foods (Scalone et al., 2015). Roasted GBSE_R surpassed other specimens being enriched with pyrazine/pyrrole at 15%. 2-Ethyl-6-methylpyrazine (83) was at the major form (5.4%) in GBSE_R reported in roasted peanuts and coffee (Li et al., 2018). In contrast, considerable amounts of methyl 2-methylaminobenzoate were detected in unroasted GBSE and GBSA at 10.1 and 6.1%, respectively, and to provide a floral- orange blossom note (Elmaci and Altug, 2005).

3.4.1 Unsupervised multivariate data analysis of headspace aroma GC-MS dataset

The GC-MS derived aroma profiles were initially modelled using PCA to reduce the high-dimensional dataset in a two-dimensional score plot and to discern the volatilome similarity/heterogeneity existing among specimens. The established PCA model (**Suppl. Fig. S3**) explained only 54% of the total variance ascribed by PC1 and PC2, and moreover did not provide a good separation between all specimens suggestive for the highly diversified aroma profile of the *G. bruguieri* (Desf.) roots (**Suppl. Fig. S3A**). Such heterogeneity can be attributed to either agricultural practices in different regions or variable drying/storage conditions employed in each territory. As shown in PCA score plot (**Suppl. Fig. S3A**), only one sample “GBRE” obtained from Rosetta, Egypt clustered apart from the remaining specimens due to its enrichment in (*E*)-cinnamaldehyde as revealed from PCA loading plot (**Suppl. Fig. S3B**).

A comparable clustering picture with highly overlapping specimens was also revealed from HCA dendrogram (**Suppl. Fig. S3C**). Accordingly, GC-MS profiles failed to capture the variabilities arising from the different geographical origin of roots samples when modeled using PCA unsupervised tools compared to NMR.

3.4.2 Supervised multivariate OPLS-DA analysis of headspace aroma GC-MS dataset in response to roasting

To determine the roasting impact on *G. bruguieri* (Desf.) aroma profile as analysed using GC-MS, OPLS-DA modeling of the aroma profiles of unroasted roots against the roasted one of the same origin was performed (**Fig. 6**). A maximum separation between the samples based on the roasting effect was attained by the developed model which showed a

high prediction power as validated by R² and Q² values of 0.999 and 0.994, respectively. The roasted root samples “GBSE_R” were grouped along the positive score of PC1 (**Fig.6A**) with their aroma profile found to be more enriched in pyrazines *i.e.* 2-ethyl-6-methylpyrazine and furan compound *i.e.* 5-methyl-2-furfural, which were posed as root roasting markers, as depicted in S-plot (**Fig.6B**). In contrast, ether *i.e.* anethol and oxide *i.e.* safrole and the nitrogenous compound *i.e.* methyl-2-methylamino benzoate were found more abundant in unroasted roots.

4. Conclusions

The current study presents the first detailed elucidation of the aroma profiles of roasted and unroasted *Glossostemon bruguieri* (Desf.) roots from four different geographical locations of Egypt and Saudi Arabia attributing to their organoleptic characteristics. Root roasting descriptors *i.e.* ethyl acetate, 2-ethyl-6-methylpyrazine and 5-methyl-2-furfural have been presented *via* the OPLS-DA modeling of aroma profiles of roasted versus unroasted root samples. The study also provides the first NMR based metabolomics approach to explore the metabolome variability among *G. bruguieri* roots and for standardization of their extracts. OPLS-DA analysis of NMR fingerprints (aromatic region) defined moghatin as the discriminatory metabolite for the Aswan, Egypt and Saudi Arabia over the other samples from Siwa and Rosetta, later were found more rich in fraxetin and pinoresinaol. Further, quantification of *G. bruguieri* (Desf.) root metabolites has been accomplished providing the exact differential relative metabolites difference among specimens and posing *G. bruguieri* samples from Siwa, Egypt to be most valuable for dietary supplements. It can also be incorporated in safe weight gain preparations without increasing blood glucose levels. In addition, its mucilage content should be further investigated as a functional food hydrocolloid in nutraceuticals using other tools such as HPLC and LC/MS. The multiplex approach targeting volatile, primary and secondary profiles, in parallel with multivariate data analysis could be applied to examine specimens from other geographical regions, and could help correlate biological effects of *G. bruguieri* to key metabolites using regression models such as OPLS-DA.

Head space solid phase microextraction technique was employed for the determination of trace volatiles in *G. bruguieri* samples, characterizing its genuine volatile profile as a raw or roasted herbal material. However, providing a score for the overall aroma would be better achieved using an olfactometer or an electronic nose for detection, which could be considered in the future. Further, extensive investigation of the secondary metabolome of *G. bruguieri* is still needed to be carried out ideally using liquid chromatography coupled to mass spectrometry (LC-MS) to rationalize more for its health effects and its utilization in nutraceuticals.

The current approach can be applicable for analyzing samples from further geographical origins as well as for exploiting factors that might influence the root metabolic makeup *i.e.*, seasonal variation, growth stage, storage and other processing conditions.

Acknowledgments: Prof. Mohamed A. Farag wishes to thank the Alexander von Humboldt Foundation, Germany for financial support.

Conflicts of Interest: The authors have declared that there is no conflict of interest.

References

- Afifi, S.M., El-Mahis, A., Heiss, A.G., Farag, M.A., (2021). Gas Chromatography–Mass Spectrometry-Based Classification of 12 Fennel (*Foeniculum vulgare* Miller) Varieties Based on Their Aroma Profiles and Estragole Levels as Analyzed Using Chemometric Tools. *ACS Omega* 6(8), 5775-5785.
- Al-Snafi, A.E., (2019). Medical importance of *Glossostemon bruguieri*–A review. *IOSR Journal of pharmacy* 9(5), 34-39.
- Al Muqarrabun, L.M.R., Ahmat, N., (2015). Medicinal uses, phytochemistry and pharmacology of family Sterculiaceae: A review. *European Journal of Medicinal Chemistry* 92, 514-530.
- Ali, Z.A.A., (2020). Taxonomic study of *Glossostemon bruguieri* Desf.(Malvaceae) in Iraq. *Plant Archives* 20(2), 926-929.
- Alwhibi, M.S., Khalil, M.I., Ibrahim, M.M., El-Gaaly, G.A., Sultan, A.S., (2017). Potential antitumor activity and apoptosis induction of *Glossostemon bruguieri* root extract against hepatocellular carcinoma cells. *Evidence-Based Complementary and Alternative Medicine* 2017.
- Amin, E.S., Awad, O., Abd El Samad, M., Iskander, M., (1969). Isolation of estrone from Moghat roots and from pollen grains of Egyptian date palm. *Phytochemistry* 8(1), 295-297.
- Avonto, C., Wang, M., Chittiboyina, A.G., Vukmanovic, S., Khan, I.A., (2018). Chemical stability and in chemico reactivity of 24 fragrance ingredients of concern for skin sensitization risk assessment. *Toxicology in Vitro* 46, 237-245.

- Balaha, M., Ahmed, N., Geddawy, A., Kandeel, S., (2021). Fraxetin prevented sodium fluoride-induced chronic pancreatitis in rats: Role of anti-inflammatory, antioxidant, antifibrotic and anti-apoptotic activities. *International Immunopharmacology* 93, 107372.
- Bharti, S.K., Roy, R., (2012). Quantitative ¹H NMR spectroscopy. *TrAC Trends in Analytical Chemistry* 35, 5-26.
- Chen, J.L., Yan, S., Feng, Z., Xiao, L., Hu, X.S., (2006). Changes in the volatile compounds and chemical and physical properties of Yali pear (*Pyrus bertschneideri* Reld) during storage. *Food Chemistry* 97(2), 248-255.
- Ciftci, O., Oztanir, M.N., Cetin, A., (2014). Neuroprotective effects of β -myrcene following global cerebral ischemia/reperfusion-mediated oxidative and neuronal damage in a C57BL/J6 mouse. *Neurochemical research* 39(9), 1717-1723.
- Cui, H., Siva, S., Lin, L., (2019). Ultrasound processed cuminaldehyde/2-hydroxypropyl- β -cyclodextrin inclusion complex: Preparation, characterization and antibacterial activity. *Ultrasonics sonochemistry* 56, 84-93.
- Dhifi, W., Angerosa, F., Serraiocco, A., Oumar, I., Hamrouni, I., Marzouk, B., (2005). Virgin olive oil aroma: Characterization of some Tunisian cultivars. *Food Chemistry* 93(4), 697-701.
- El-Sayed, N., Awaad, A., Mabry, T., (2004). Phytochemical studies and effect on urine volume of *Glossostemon bruguieri* Desf. constituents.
- Elmaci, Y., Altug, T., (2005). Flavor characterization of three mandarin cultivars (Satsuma, Bodrum, Clemantine) by using GC/MS and flavor profile analysis techniques. *Journal of Food Quality* 28(2), 163-170.
- Farag, M.A., Elsebai, M.F., Khattab, A.R., (2018). Metabolome based classification of artichoke leaf: A prospect for phyto-equivalency of its different leaf origins and commercial preparations. *Journal of Pharmaceutical and Biomedical Analysis* 158, 151-159.
- Farag, M.A., Fathi, D., Shamma, S., Shawkat, M.S.A., Shalabi, S.M., El Seedi, H.R., Afifi, S.M., (2021a). Comparative metabolome classification of desert truffles *Terfezia clavaryi* and *Terfezia boudieri* via its aroma and nutrients profile. *LWT* 142, 111046.
- Farag, M.A., Hegazi, N., Dokhalahy, E., Khattab, A.R., (2020a). Chemometrics based GC-MS aroma profiling for revealing freshness, origin and roasting indices in saffron spice and its adulteration. *Food Chemistry* 331, 127358.
- Farag, M.A., Khattab, A.R., Shamma, S., Afifi, S.M., (2021b). Profiling of Primary Metabolites and Volatile Determinants in Mahlab Cherry (*Prunus mahaleb* L.) Seeds in the Context of Its Different Varieties and Roasting as Analyzed Using Chemometric Tools. *Foods* 10(4), 728.
- Farag, M.A., Shakour, Z.T., Lübken, T., Frolov, A., Wessjohann, L.A., Mahrous, E., (2020b). Unraveling the metabolome composition and its implication for *Salvadora persica* L. use as dental brush via a multiplex approach of NMR and LC-MS metabolomics. *Journal of Pharmaceutical and Biomedical Analysis* 193, 113727.
- Farag, M.A., Sharaf El-Din, M.G., Selim, M.A., Owis, A.I., Abouzid, S.F., Porzel, A., Wessjohann, L.A., Otify, A., (2021c). Nuclear Magnetic Resonance Metabolomics Approach for the Analysis of Major Legume Sprouts Coupled to Chemometrics. *Molecules* 26(3), 761.
- Farrag, A.R.H., Abdallah, H.M., Khattab, A.R., Elshamy, A.I., El Gendy, A.E.-N.G., Mohamed, T.A., Farag, M.A., Efferth, T., Hegazy, M.-E.F., (2019). Antiulcer activity of *Cyperus alternifolius* in relation to its UPLC-MS metabolite fingerprint: A mechanistic study. *Phytomedicine* 62, 152970.
- Firmino, D.F., Cavalcante, T.T., Gomes, G.A., Firmino, N., Rosa, L.D., de Carvalho, M.G., Catunda Jr, F.E., (2018). Antibacterial and antibiofilm activities of *Cinnamomum* Sp. essential oil and cinnamaldehyde: Antimicrobial activities. *The Scientific World Journal* 2018.
- Gamel, T., El-Razek, A.A., Damir, A., (2010). Dried peeled roots of *glossostemon bruguieri* (moghat) as a potential functional food. *Journal of food processing and preservation* 34(1), 55-67.

- Ghareeb, D.A., El-Rashidy, F.H., El-Mallawany, S., (2014). Imbalanced diet deficient in calcium and vitamin D-induced juvenile osteopenia in rats; the potential therapeutic effect of Egyptian Moghat roots water extract (*Glossostemon bruguieri*). Iranian journal of pharmaceutical research: IJPR 13(2), 623.
- Ibrahim, N., El-Eraky, W., El-Gengaihi, S., Shalaby, A., (1997). Chemical and biological evaluation of proteins and mucilages from roots and seeds of *Glossostemon bruguieri* Desf.(Moghat). Plant Foods for Human Nutrition 50(1), 55-61.
- Khalil, M.N., Fekry, M.I., Farag, M.A., (2017). Metabolome based volatiles profiling in 13 date palm fruit varieties from Egypt via SPME GC-MS and chemometrics. Food Chem. 217, 171-181.
- Leffingwell, J.C., Alford, E., (2005). Volatile constituents of perique tobacco. Electronic Journal of Environmental, Agricultural and Food Chemistry 4(2), 899-915.
- Li, C., Sun, J., Li, T., Liu, S.-Q., Huang, D., (2014). Chemical and enzymatic synthesis of a library of 2-phenethyl esters and their sensory attributes. Food Chemistry 154, 205-210.
- Li, Y.-Y., Lu, Y.-Y., Lu, M., Wei, H.-Y., Chen, L., (2018). HPLC separation of 2-ethyl-5 (6)-methylpyrazine and its electroantennogram and alarm activities on fire ants (*Solenopsis invicta* Buren). Molecules 23(7), 1661.
- Lilly, M., Bauer, F.F., Lambrechts, M.G., Swiegers, J.H., Cozzolino, D., Pretorius, I.S., (2006). The effect of increased yeast alcohol acetyltransferase and esterase activity on the flavour profiles of wine and distillates. Yeast 23(9), 641-659.
- Liu, R., Sun, Q., Sun, A., Cui, J., (2005). Isolation and purification of coumarin compounds from *Cortex fraxinus* by high-speed counter-current chromatography. Journal of chromatography A 1072(2), 195-199.
- López-Biedma, A., Sánchez-Quesada, C., Beltrán, G., Delgado-Rodríguez, M., Gaforio, J.J., (2016a). Phytoestrogen (+)-pinoretinol exerts antitumor activity in breast cancer cells with different oestrogen receptor statuses. BMC Complementary and Alternative Medicine 16(1), 1-14.
- López-Biedma, A., Sánchez-Quesada, C., Beltrán, G., Delgado-Rodríguez, M., Gaforio, J.J., (2016b). Phytoestrogen (+)-pinoretinol exerts antitumor activity in breast cancer cells with different oestrogen receptor statuses. BMC Complementary and Alternative Medicine 16(1), 350.
- Lyumugabe, F., Uyisenga, J.P., Songa, E.B., Thonart, P., (2014). Production of traditional sorghum beer "Ikigage" using *Saccharomyces cerevisiae*, *Lactobacillus fermentum* and *Issatchenkia orientalis* as starter cultures. Food and Nutrition Sciences 2014.
- Meselhy, M.R., (2003). Constituents from Moghat, the roots of *Glossostemon bruguieri* (Desf.). Molecules 8(8), 614-621.
- PANGBORN, R.M., (1963). Relative Taste Intensities of Selected Sugars and Organic Acids a. Journal of Food Science 28(6), 726-733.
- Pathak, N., Bhaduri, A., Rai, A.K., (2019). Sesame: Bioactive compounds and health benefits. Bioactive molecules in food, 181-200.
- Pico, J., Bernal, J., Gómez, M., (2015). Wheat bread aroma compounds in crumb and crust: A review. Food Research International 75, 200-215.
- Rasheed, D.M., Porzel, A., Frolov, A., El Seedi, H.R., Wessjohann, L.A., Farag, M.A., (2018). Comparative analysis of *Hibiscus sabdariffa* (roselle) hot and cold extracts in respect to their potential for α -glucosidase inhibition. Food Chemistry 250, 236-244.
- Rasheed, D.M., Serag, A., Shakour, Z.T.A., Farag, M., (2020). Novel trends and applications of multidimensional chromatography in the analysis of food, cosmetics and medicine bearing essential oils. Talanta, 121710.
- Roy, S.C., Rana, K.K., Guin, C., (2002). Short and stereoselective total synthesis of furano lignans ($\pm$)-dihydrosesamin, ($\pm$)-laricresinol dimethyl ether, ($\pm$)-acuminatin methyl ether, ($\pm$)-sanshodiol methyl ether, ($\pm$)-

lariciresinol,(±)-acuminatin, and (±)-lariciresinol monomethyl ether and furofuran lignans (±)-sesamin,(±)-eudesmin,(±)-piperitol methyl ether,(±)-pinoresinol,(±)-piperitol, and (±)-pinoresinol monomethyl ether by radical cyclization of epoxides using a transition-metal radical source. *The Journal of organic chemistry* 67(10), 3242-3248.

Scalone, G.L.L., Cucu, T., De Kimpe, N., De Meulenaer, B., (2015). Influence of free amino acids, oligopeptides, and polypeptides on the formation of pyrazines in Maillard model systems. *Journal of agricultural and food chemistry* 63(22), 5364-5372.

Sharaf, M., El-Ansari, M., Saleh, N., (1998). A new flavonoid from the roots of *Glossostemon bruguieri*. *Fitoterapia* 69, 47-48.

Sher, H., Alyemeni, M., (2011). Pharmaceutically important plants used in traditional system of Arab medicine for the treatment of livestock ailments in the kingdom of Saudi Arabia. *African Journal of Biotechnology* 10(45), 9153-9159.

Thuong, P.T., Pokharel, Y.R., Lee, M.Y., Kim, S.K., Bae, K., Su, N.D., Oh, W.K., Kang, K.W., (2009). Dual anti-oxidative effects of fraxetin isolated from *Fraxinus rhynchophylla*. *Biological and Pharmaceutical Bulletin* 32(9), 1527-1532.

Triba, M.N., Le Moyec, L., Amathieu, R., Goossens, C., Bouchemal, N., Nahon, P., Rutledge, D.N., Savarin, P., (2015). PLS/OPLS models in metabolomics: the impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Molecular BioSystems* 11(1), 13-19.

Wang, C., Sun, J., Tao, Y., Fang, L., Zhou, J., Dai, M., Liu, M., Fang, Q., (2020). Biomass materials derived from anethole: conversion and application. *Polym. Chem.* 11(5), 954-963.

Wiklund S., Johansson E., Sjöström L., Mellerowicz E.J., Edlund U., Shockcor J.P., Gottfries J., Moritz T., Trygg J., (2008). Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Anal. Chem.* 80(1), 115-122.

Xu, Z., Deng, M., (2017). *Malvaceae, Identification and Control of Common Weeds: Volume 2*. Springer, pp. 717-735.

List of Tables and Figures

Table 1. Resonance assignments with chemical shifts of constituents identified in 600 MHz ^1H -NMR, HSQC and HMBC spectra of *G. bruguieri* (Desf.) root methanol extracts.

Primary Metabolites					
#	Metabolite	Assignment	$\delta\ ^1\text{H}$ (ppm)	$\delta\ ^{13}\text{C}$ (ppm)	HMBC correlations $\delta\ ^{13}\text{C}$ (ppm)
(A) Sugars and sugar alcohol					
N1	Sucrose*	CH-1 (α -D-Glucose)	5.37 (d, J = 3.9 Hz)	93.6	74.5 (C-4'), 105.3 (C-2')
		CH ₂ -1' (Fructose)	3.61 (d, J = 5.6 Hz)	64.1	105.3 (C-2'), 71.5 (C-2), 79.3 (C-3')
N2	β -D-Glucose*	CH-1	4.45 (d, J = 7.7 Hz)	97.4	n.d.
N3	α -D-Glucose*	CH-1	5.08 (d, J = 3.7 Hz)	92.7	73.1(C-3)
N4	Inositol	CH-1/3	3.40 (dd, J = 9.8, 3.7 Hz)	74.4	73.8 (C-5)
		CH-4/6	3.70 (m)	73.8	74.4 (C-1/3)
(B) Organic and fatty acids					
N5	Fatty acid	CH ₂ -2	2.33	n.d.	26.1 (C-3), 30.6 (C-4), 175.6 (C-1)
		CH ₂ -3	1.58 (m)	26.1	n.d.
		Inner CH ₂	1.28-1.37 (m)	29.2	29.2 (CH ₂)n
		CH- ω -2	2.01-2.05 (m)	24.1-24.5	29.2 (CH ₂)n, 131.4 (olefinic Cs)
		CH ₃ (ω)	0.98 (t, J = 6.9 Hz)	15.1	22.8, 30.0
N6	Acetic acid	CH ₃ -2	1.94 (s)	22.4	177.4 (C-1)
N7	Succinic acid	CH ₂ -1,2	2.50 (s)	31.7	31.6 (C-1/2)
N8	Fumaric acid	2 CH	6.57 (s)	135.3	170.9 (C-1)
Secondary Metabolites					
#	Metabolite	Assignment	$\delta\ ^1\text{H}$ (ppm)	$\delta\ ^{13}\text{C}$ (ppm)	HMBC correlations $\delta\ ^{13}\text{C}$ (ppm)
(C)- Coumarin					
N9	Fraxetin	CH-3	6.17 (d, J = 9.54 Hz)	118.5	162.9 (C-2)
		CH-4	7.83 (d, J = 9.54 Hz)	145.0	162.9 (C-2)
		CH-5	6.70 (br.s)	104.7	123.8 (C-10), 145.1 (C-4), 150.5 (C-7)
		OCH ₃	3.87 (s)	57.1	n.d.
(D)- Biflavonoid and flavone					

N10	Moghatin	CH-3"	6.61 (s)	102.9	105.9 (C-10"), 120.1 (C-6""), 167.1(C-2"), 185.0 (C-4")
		CH-3	6.65 (s)	103.0	105.9 (C-10), 167.1(C-2), 185.0 (C-4)
		CH-6	6.28 (s)	98.8	105.2 (C-8), 166.5 (C-5)
		CH-6"	6.29 (s)	98.0	105.2 (C-8"), 166.5 (C-5")
		CH-3'/5'	6.90 (dd, $J = 8.8, 2.2$ Hz)	117.0	123.9 (C-2'/6'), 162.8 (C-4')
		CH-2'/6'	8.01 (dd, $J = 8.8, 2.2$ Hz)	123.9	130.4 (C-1'), 162.8 (C-4'), 167.1 (C-2)
		CH-2'''	7.60 (d, $J = 2.2$ Hz)	113.6	120.1 (C-6''')
		CH-5'''	6.87 (d, $J=8.4$ Hz)	116.1	120.1 (C-6'''), 149.3 (C-4''')
		CH-6'''	7.42 (d, $J = 8.4$ Hz)	120.1	128.5 (C-1''')
N11	4'-Methoxy-isoscutellargin (Tatakin)	CH-3	6.61 (s)	102.9	166.1 (C-2), 185.0 (C-4)
		CH-6	6.28 (s)	98.0	122.8 (C-8)
		CH-3'/5'	7.06 (dd, $J = 8.8, 2.2$ Hz)	116.1	124.1 (C-1'), 163.2 (C-4')
		CH-2'/6'	8.11 (dd, $J = 8.8, 2.2$ Hz)	130.6	163.2 (C-4')
		OCH ₃	3.66 (s)	54.5	122.8 (C-8)
(E)- Furano-lignan					
N12	Pinoresinol	CH-1/5	2.94 (m)	53.2	53.2 (C-1/5)
		CH ₂ -4/8 α	3.41 (dd, $J= 9.9,6.2$ Hz)	71.5	53.2 (C-1/5), 84.6 (C-2)
		CH ₂ -4/8 β	3.76 (dd, $J= 9.5,6.9$ Hz)		
		CH-5'/5"	6.73 (d, $J=7.8$ Hz)	106.9	136.0 (C-1'/1"), 146.4 (C-4'/4")
		CH-6'/6"	6.77 (dd, $J= 8.4,2.2$ Hz)	117.5	106.9 (C-5'/5")
		CH-2'/2"	6.69 (d, $J=1.8$ Hz)	112.9	n.d.
		OCH ₃	3.73 (s)	55.8	n.d.

n.d., not detected. *denotes metabolites confirmed using standards.

Table 2: ¹H-NMR quantification of major primary and secondary metabolites in different samples of *G. bruguieri* (Desf.) methanol extracts (GBSE, GBAE, GBSA and GBRE). Values are expressed as µg/mg dry powder ± S.D (n=3), see experimental section. Chemical shifts used for metabolite quantification were determined in methanol-*d*₆ and expressed as relative values to HMDS (0.94 mM final concentration). Different letters indicate significant differences between artichoke specimens according to least significant difference analysis (LSD).

ID	Compound	Amount µg/mg dry matter (n=3)			
		GBSE ^(a)	GBAE ^(b)	GBSA ^(c)	GBRE ^(d)
N1	Sucrose	17.89±1.01 ^{b,c}	7.67±0.57 ^{a,c,d}	11.24±0.18 ^{a,b,d}	16.99±0.79 ^{b,c}
N2	β-Glucose	5.01±0.17 ^{b,c,d}	1.97±0.13 ^{a,c,d}	2.88±0.11 ^{a,b}	3.00±0.04 ^{a,b}
N3	α-Glucose	3.91±0.35 ^{b,c,d}	1.19±0.14 ^{a,c,d}	2.21±0.14 ^{a,b,d}	1.81±0.13 ^{a,b,c}
N4	Inositol	12.06±0.51 ^{b,c,d}	5.15±0.52 ^{a,c,d}	7.38±0.12 ^{a,b,d}	10.84±0.59 ^{a,b,c}
N6	Acetic acid	0.35±0.01 ^{b,c,d}	0.13±0.04 ^{a,d}	0.102±0.02 ^{a,d}	0.29±0.01 ^{a,b,c}
N7	Succinic acid	1.05±0.02 ^{b,c,d}	0.71±0.06 ^{a,d}	0.76±0.02 ^{a,d}	0.60±0.01 ^{a,b,c}
N9	Fraxetin	3.03±0.27 ^b	2.18±0.27 ^{a,c,d}	2.79±0.22 ^{a,b,d}	3.03±0.11 ^b
N10	Moghatin	7.73±0.36 ^b	5.68±0.42 ^{a,c,d}	7.62±0.26 ^b	7.36±0.26 ^b
N11	Tatakin	3.81±0.13 ^b	2.66±0.21 ^{a,c,d}	3.58±0.12 ^b	3.55±0.11 ^b
N12	Pinoresinol	4.55±0.15 ^b	3.17±0.26 ^{a,c,d}	4.27±0.14 ^b	4.23±0.13 ^b

*Results are average of 3 independent biological replicates ± std. deviation.

*(a–d): Different letters indicate significant differences between *G. bruguieri* specimens according to least significant difference analysis (P < .05; Tukey's test).

Table 3 Relative percentile of volatile components in *Glossostemon bruguieri* (Desf.) specimens using SPME-GC-MS (n=3)^{a-e}, results are presented as average $\pm$ (std. deviation). Sample codes are given in experimental section 2.1. *denotes metabolites confirmed using standards.

No.	Rt(min)	KI	Volatile	GBSE_R ^(a)	GBSE ^(b)	GBAE ^(c)	GBSA ^(d)	GBRE ^(e)
Organic acids								
1	2.1	450	Acetic acid	2.83 $\pm$ 3.60	0.01 $\pm$ 0.01	0.02 $\pm$ 0.02	0.05 $\pm$ 0.07	0.02 $\pm$ 0.02
2	6.947	979	Caproic acid	0.10 $\pm$ 0.09	0.22 $\pm$ 0.24	0.67 $\pm$ 0.45	1.14 $\pm$ 1.57	0.41 $\pm$ 0.07
			Total	2.93	0.23	0.69	1.19	0.43
Alcohols								
3	4.417	693	2,3-Butandiol	0.06 $\pm$ 0.03	2.21 $\pm$ 1.93	2.02 $\pm$ 0.58	1.73 $\pm$ 1.94	1.79 $\pm$ 0.76
4	5.263	780	3-(Z)-Hexenol*	0.13 $\pm$ 0.09 ^{c,e}	0.13 $\pm$ 0.15 ^{c,e}	1.14 $\pm$ 0.40 ^{a,b}	0.73 $\pm$ 0.67	1.11 $\pm$ 0.12 ^{a,b}
5	7.933	1075	1-Octanol	0.70 $\pm$ 0.88	0.09 $\pm$ 0.06	0.26 $\pm$ 0.04	3.11 $\pm$ 4.31	0.73 $\pm$ 0.14
6	8.282	1108	Linalool*	0.31 $\pm$ 0.18	2.72 $\pm$ 3.67	2.34 $\pm$ 0.27	2.66 $\pm$ 2.27	0.30 $\pm$ 0.09
7	8.471	1125	Phenethyl alcohol	0.01 $\pm$ 0.00	0.02 $\pm$ 0.01	0.01 $\pm$ 0.00	13.41 $\pm$ 18.96	0.00 $\pm$ 0.00
8	8.523	1130	Fenchol	0.18 $\pm$ 0.11	0.06 $\pm$ 0.01	0.03 $\pm$ 0.00	2.41 $\pm$ 3.36	0.04 $\pm$ 0.00
9	8.641	1140	1,2-Dihydrolinalool	0.04 $\pm$ 0.00	0.05 $\pm$ 0.05	0.03 $\pm$ 0.00	0.10 $\pm$ 0.08	0.01 $\pm$ 0.00
10	9.008	1174	1-Nonanol	0.08 $\pm$ 0.04	0.06 $\pm$ 0.04	0.07 $\pm$ 0.02	0.37 $\pm$ 0.36	0.15 $\pm$ 0.01
11	9.078	1181	dihydro- γ -Terpineol	0.10 $\pm$ 0.05 ^e	0.08 $\pm$ 0.06 ^e	0.18 $\pm$ 0.03 ^e	0.17 $\pm$ 0.09 ^e	0.45 $\pm$ 0.05 ^{a-d}
12	9.179	1190	4-Terpineol	0.02 $\pm$ 0.00 ^d	0.20 $\pm$ 0.26	0.22 $\pm$ 0.03	0.39 $\pm$ 0.20 ^a	0.10 $\pm$ 0.02
13	9.22	1194	<i>p</i> -Cymene-8-ol	0.03 $\pm$ 0.02	0.02 $\pm$ 0.01	0.03 $\pm$ 0.01	0.26 $\pm$ 0.34	0.15 $\pm$ 0.01
14	9.297	1201	α -Terpineol*	0.12 $\pm$ 0.09	0.40 $\pm$ 0.52	0.62 $\pm$ 0.09	0.49 $\pm$ 0.19	0.36 $\pm$ 0.05
15	14.069	1693	1-Tetradecanol	0.22 $\pm$ 0.18 ^c	0.21 $\pm$ 0.16 ^c	5.09 $\pm$ 2.90 ^c	0.20 $\pm$ 0.13 ^{a,b,d,e}	0.49 $\pm$ 0.10 ^c
			Total	2.00	6.27	12.05	26.04	5.68
Aldehydes								
16	4.547	707	Hexanal	0.55 $\pm$ 0.49	0.14 $\pm$ 0.04	0.61 $\pm$ 0.13	0.67 $\pm$ 0.48	0.23 $\pm$ 0.01
17	6.668	947	Benzaldehyde*	0.01 $\pm$ 0.01 ^d	0.00 $\pm$ 0.00 ^d	0.01 $\pm$ 0.00 ^d	0.44 $\pm$ 0.37 ^{a,b,c,e}	0.00 $\pm$ 0.00 ^d
18	7.667	1050	Phenyl acetaldehyde	0.03 $\pm$ 0.02 ^{b,c,d,e}	0.35 $\pm$ 0.09 ^{a,c,e}	0.75 $\pm$ 0.15 ^{a,b,c,d}	0.22 $\pm$ 0.02 ^{a,c,e}	1.28 $\pm$ 0.02 ^{a-d}

19	8.324	1112	Nonanal*	0.60 ± 0.46	0.23 ± 0.05	0.14 ± 0.02	0.12 ± 0.08	0.47 ± 0.15
20	9.378	1210	Decanal	0.44 ± 0.33	0.35 ± 0.09	0.18 ± 0.03	0.29 ± 0.03	0.29 ± 0.03
21	9.765	1252	Z-Citral	0.02 ± 0.01 ^{b,d}	0.09 ± 0.02 ^{a,e}	0.02 ± 0.00 ^d	0.12 ± 0.06 ^{a,c,e}	0.01 ± 0.00 ^{b,d}
22	9.814	1256	Cumaldehyde	0.18 ± 0.03 ^c	1.63 ± 1.69 ^c	5.98 ± 0.95 ^{a,b,e}	3.21 ± 2.70	0.82 ± 0.17 ^c
23	9.943	1270	p-Anisaldehyde	0.09 ± 0.04 ^b	0.79 ± 0.42 ^{a,e}	0.48 ± 0.08	0.33 ± 0.12	0.11 ± 0.01 ^b
24	10.039	1280	α-Citral*	0.04 ± 0.03 ^d	0.11 ± 0.10 ^d	0.09 ± 0.02 ^d	0.44 ± 0.29 ^{a,b,c,e}	0.06 ± 0.00 ^d
25	10.103	1287	(E)-Cinnamaldehyde*	8.35 ± 3.85 ^e	11.18 ± 10.30 ^e	1.02 ± 0.55 ^e	1.26 ± 1.02 ^e	46.15 ± 7.33 ^{a-d}
26	11.257	1408	Dodecanal	0.04 ± 0.04 ^e	0.01 ± 0.00 ^e	0.12 ± 0.03 ^e	0.09 ± 0.07 ^e	0.31 ± 0.08 ^{a-d}
27	14.383	1728	1-Pentadecanal	0.02 ± 0.02 ^c	0.02 ± 0.01 ^c	3.37 ± 0.49 ^{a,b,d,e}	0.04 ± 0.03 ^c	0.01 ± 0.00 ^c
Total				10.36	14.90	12.78	7.24	49.76
Aliphatic/aromatic hydrocarbon								
28	8.076	1089	Unknown	0.02 ± 0.02	3.41 ± 4.16	2.85 ± 0.71	0.52 ± 0.26	0.31 ± 0.08
29	9.444	1217	2,6-Dimethylundecane	0.01 ± 0.00 ^c	0.06 ± 0.04 ^c	0.02 ± 0.00 ^{a,b,e}	0.35 ± 0.48	0.00 ± 0.00 ^c
30	10.517	1330	4,6-Dimethyldodecane	0.02 ± 0.01	0.15 ± 0.18 ^c	0.59 ± 0.12 ^b	0.29 ± 0.26	0.05 ± 0.01
31	11.18	1401	Tetradecane	0.19 ± 0.19 ^c	0.06 ± 0.06 ^c	0.37 ± 0.06 ^{a,b,e,d}	0.26 ± 0.17 ^c	0.22 ± 0.00 ^c
32	13.212	1599	Hexadecane*	0.01 ± 0.01 ^e	0.02 ± 0.01 ^e	1.25 ± 0.17 ^e	0.02 ± 0.01 ^e	0.06 ± 0.00 ^{a-d}
33	14.229	1711	Heptadecane*	0.04 ± 0.02	0.06 ± 0.04	0.04 ± 0.01	0.03 ± 0.02	0.45 ± 0.01
34	14.296	1718	2,6,10-Trimethylhexadecane	0.09 ± 0.06 ^b	0.19 ± 0.22 ^{a,c,d,e}	0.06 ± 0.01 ^b	0.05 ± 0.04 ^b	0.12 ± 0.01 ^b
35	15.034	1800	Octadecane*	0.05 ± 0.02 ^b	0.59 ± 0.34 ^{a,c,d,e}	0.05 ± 0.03 ^b	0.05 ± 0.02 ^b	0.14 ± 0.02 ^b
36	16.95	2000	Eicosane	0.15 ± 0.11	1.08 ± 0.72	0.02 ± 0.00	0.02 ± 0.01	0.05 ± 0.02
Total				0.58	5.62	5.24	1.59	1.39
Esters								
37	2.355	478	Ethyl acetate	9.83 ± 4.81	3.08 ± 2.28	2.65 ± 3.64	12.70 ± 17.89	4.45 ± 4.00
38	5.445	800	Acetonyl acetate	2.29 ± 1.70	0.14 ± 0.06 ^e	0.10 ± 0.05 ^e	0.96 ± 1.22	4.11 ± 2.52 ^{b,c}
39	7.208	1009	(Z)-3-Hexenol acetate*	38.01 ± 11.04 ^{b,c,d,e}	4.57 ± 4.03 ^a	17.68 ± 8.92 ^a	8.74 ± 7.73 ^a	4.85 ± 3.46 ^a
40	9.437	1215	Caprylyl acetate	0.03 ± 0.02	0.07 ± 0.04	0.04 ± 0.01	0.77 ± 1.06	0.01 ± 0.00
41	9.541	1227	Methyl nonanoate	0.03 ± 0.01 ^d	0.05 ± 0.04 ^d	0.11 ± 0.03	0.23 ± 0.14 ^{a,b}	0.07 ± 0.04
42	9.845	1259	Linalyl acetate*	0.09 ± 0.07	0.35 ± 0.44	0.24 ± 0.03	0.55 ± 0.36	0.03 ± 0.00
43	10.808	1362	α-Terpinyol acetate	0.02 ± 0.01	0.84 ± 1.06	0.58 ± 0.13	0.91 ± 0.99	0.08 ± 0.02
44	16.173	1919	Methyl palmitate	0.07 ± 0.03 ^b	4.47 ± 3.41 ^{a,c,d,e}	0.16 ± 0.01 ^b	0.49 ± 0.44 ^b	0.17 ± 0.08 ^b

45	16.523	1955	Isopropyl palmitate	0.63 ± 0.36	1.08 ± 1.12	0.03 ± 0.01	0.29 ± 0.24	0.09 ± 0.04
			Total	51.00	14.64	21.57	25.64	13.85
			Ethers/oxides					
46	7.569	1042	Cineol*	0.06 ± 0.03 ^c	0.27 ± 0.31 ^c	0.70 ± 0.09 ^{a,b,d,e}	0.24 ± 0.11 ^c	0.24 ± 0.01 ^c
47	9.362	1208	Estragole*	0.10 ± 0.05	2.82 ± 1.99	2.65 ± 0.31	1.82 ± 1.94	0.13 ± 0.04
48	10.231	1300	Anethole*	0.67 ± 0.43 ^{b,c}	13.35 ± 5.35 ^{a,d,e}	8.36 ± 1.08 ^{a,e}	4.75 ± 3.42 ^b	0.23 ± 0.05 ^{b,c}
49	10.283	1306	Safrole	0.40 ± 0.21 ^b	8.62 ± 2.06 ^{a,e}	4.40 ± 0.82	5.82 ± 4.87	0.05 ± 0.01 ^b
50	10.839	1365	Eugenol*	0.03 ± 0.01	0.13 ± 0.12	0.08 ± 0.02	0.17 ± 0.10	0.03 ± 0.01
51	10.861	1367	4-Acetylanisole	0.03 ± 0.02 ^d	0.05 ± 0.02	0.05 ± 0.01	0.09 ± 0.03 ^{a,e}	0.03 ± 0.00 ^d
52	12.463	1525	Myristicin	0.02 ± 0.01	0.01 ± 0.00	0.49 ± 0.35	0.70 ± 0.97	0.12 ± 0.01
			Total	1.32	25.26	16.73	13.60	0.82
			Furans					
53	4.624	714	Dihydro-2-methyl-3-furanone	0.57 ± 0.32 ^{b,d}	0.04 ± 0.03 ^a	0.29 ± 0.01	0.20 ± 0.10 ^a	0.20 ± 0.01
54	4.924	743	Furfural	3.94 ± 2.86 ^e	1.46 ± 0.76	0.80 ± 0.05	0.94 ± 0.99	0.09 ± 0.04 ^a
55	6.042	872	Acetylfuran	0.25 ± 0.12 ^{d,e}	0.28 ± 0.12 ^{d,e}	0.11 ± 0.03	0.01 ± 0.01 ^{a,b}	0.02 ± 0.01 ^{a,b}
56	6.72	953	5-Methyl-2-furfural	3.01 ± 1.39 ^{b,c,d,e}	0.11 ± 0.06 ^a	0.47 ± 0.11 ^a	0.17 ± 0.11 ^a	0.13 ± 0.01 ^a
57	7.037	991	2-Pentylfuran	0.29 ± 0.17	0.05 ± 0.01	0.04 ± 0.00	0.36 ± 0.37	0.03 ± 0.01
58	8.045	1086	(E)-Furan linalool oxide (linaloloxide I)	0.31 ± 0.38	0.26 ± 0.21	0.00 ± 0.00	0.11 ± 0.08	0.00 ± 0.00
			Total	8.36	2.22	1.71	1.79	0.48
			Ketones/lactones					
59	2.851	528	1-Hydroxypropan-2-one	1.94 ± 1.11 ^{b,c,d,e}	0.28 ± 0.16 ^a	0.23 ± 0.08 ^a	0.24 ± 0.20 ^a	0.14 ± 0.11 ^a
60	3.109	555	2,3-Pentadione	0.04 ± 0.04	0.02 ± 0.01	0.00 ± 0.00	0.07 ± 0.08	0.03 ± 0.02
61	5.673	827	4-Cyclopenten-1,3-dione	0.21 ± 0.14 ^b	1.63 ± 1.05 ^{a,d}	0.99 ± 0.35	0.31 ± 0.27 ^b	0.92 ± 0.15
62	6.051	872	γ-Butyrolactone	0.73 ± 0.26	0.08 ± 0.04	0.14 ± 0.04	0.10 ± 0.07	0.06 ± 0.01
63	7.4	1026	3,6-Heptanedione	0.08 ± 0.05	0.03 ± 0.02	0.03 ± 0.01	0.16 ± 0.20	0.15 ± 0.01
64	7.802	1063	γ-Caprolactone	1.77 ± 1.23	0.09 ± 0.07	0.87 ± 0.19	0.21 ± 0.14	0.37 ± 0.02
65	8.973	1170	Isomenthone	0.17 ± 0.12	0.52 ± 0.35	0.06 ± 0.02	1.20 ± 1.49	0.57 ± 0.07
66	11.717	1453	Coumarin	0.55 ± 0.41 ^{b,c,d,e}	3.16 ± 3.01 ^a	0.09 ± 0.02 ^a	0.07 ± 0.05 ^a	2.23 ± 0.38 ^a
67	12.06	1487	α-Isomethyl ionone	0.12 ± 0.11 ^{b,d}	2.03 ± 1.68 ^a	1.26 ± 0.48	1.74 ± 1.36 ^a	0.10 ± 0.00

68	14.117	1698	Turmerone	0.11 ± 0.09	0.10 ± 0.08	0.08 ± 0.01	0.03 ± 0.02	0.13 ± 0.12
			Total	5.74	7.94	3.74	4.12	4.71
			Monoterpene hydrocarbons					
69	6.918	976	<i>β</i> -Pinene*	0.11 ± 0.05 ^c	0.12 ± 0.03 ^c	3.94 ± 1.20 ^{a,b,e}	1.88 ± 2.47	0.29 ± 0.00 ^c
70	7.049	992	<i>β</i> -Myrcene*	0.36 ± 0.10	0.09 ± 0.02 ^c	4.24 ± 2.54 ^b	0.43 ± 0.22	3.66 ± 3.43
71	7.283	1016	<i>δ</i> -3-Carene	0.11 ± 0.09	0.12 ± 0.03	0.14 ± 0.02	0.16 ± 0.08	0.39 ± 0.38
72	7.467	1033	<i>β</i> -Cymene	0.22 ± 0.09	0.34 ± 0.39	0.26 ± 0.04	0.41 ± 0.11	0.24 ± 0.04
73	7.471	1033	<i>o</i> -Cymene	0.07 ± 0.02 ^c	0.66 ± 0.85 ^c	1.83 ± 0.25 ^{a,b,d,e}	0.56 ± 0.21 ^c	0.06 ± 0.01 ^d
74	7.566	1041	Limonene*	0.10 ± 0.04 ^c	0.58 ± 0.68	1.04 ± 0.04 ^a	0.64 ± 0.32	0.27 ± 0.01
75	7.706	1054	(<i>Z</i>)-Ocimene	0.05 ± 0.03	0.15 ± 0.18	0.15 ± 0.03	0.20 ± 0.17	0.04 ± 0.00
76	7.852	1068	<i>γ</i> -Terpinene	0.11 ± 0.06	0.14 ± 0.11	0.88 ± 0.22	1.73 ± 2.06	0.26 ± 0.02
77	8.204	1100	<i>p</i> -Cymenene	0.04 ± 0.02 ^{c,d}	0.08 ± 0.01 ^{c,d}	1.03 ± 0.20 ^{a,b,c,e}	0.35 ± 0.15 ^{a,b,c}	0.22 ± 0.04 ^c
			Total	1.17	2.26	13.52	6.36	5.44
			Nitrogenous compounds					
78	11.395	1422	Methyl 2-methylaminobenzoate	0.36 ± 0.18 ^b	10.11 ± 5.35 ^{a,c,e}	2.37 ± 0.68 ^b	6.09 ± 4.77	0.03 ± 0.01 ^b
			Total	0.36	10.11	2.37	6.09	0.03
			Pyrazines/pyrroles					
79	4.878	740	Methylpyrazine	2.39 ± 0.90 ^{b,c,d,e}	0.05 ± 0.03 ^a	0.11 ± 0.08 ^a	0.18 ± 0.07 ^a	0.02 ± 0.00 ^a
80	6.032	870	2,6-Dimethyl pyrazine	1.79 ± 0.86 ^{b,c,d,e}	0.01 ± 0.01 ^a	0.00 ± 0.00 ^a	0.01 ± 0.00 ^a	0.13 ± 0.01 ^a
81	6.079	876	Ethylpyrazine	1.50 ± 1.10	0.08 ± 0.06	0.20 ± 0.02	1.24 ± 0.76	1.14 ± 0.08
82	6.112	880	2,3-Dimethylpyrazine	1.18 ± 0.38	0.02 ± 0.02	0.08 ± 0.04	1.12 ± 1.50	0.02 ± 0.01
83	7.069	995	2-Ethyl-6-methylpyrazine	5.39 ± 3.53 ^{b,c,d}	1.06 ± 0.27 ^a	0.79 ± 0.33 ^a	0.86 ± 0.85 ^a	2.15 ± 0.94
84	7.154	1003	2,3,5-Trimethyl pyrazine	0.72 ± 0.21 ^{b,c,d,e}	0.08 ± 0.02 ^a	0.17 ± 0.06 ^a	0.09 ± 0.03 ^a	0.11 ± 0.02 ^a
85	8.068	1088	3-Ethyl-2,5-dimethylpyrazine	0.43 ± 0.09 ^{b,c,d,e}	0.07 ± 0.02 ^a	0.05 ± 0.02 ^a	0.07 ± 0.04 ^a	0.04 ± 0.01 ^a
86	8.153	1096	2,3-Diethylpyrazine	0.15 ± 0.02	0.00 ± 0.01	0.00 ± 0.00	0.25 ± 0.35	0.00 ± 0.00
87	8.181	1097	Tetramethylpyrazine	0.07 ± 0.03	0.03 ± 0.00	0.03 ± 0.00	0.05 ± 0.03	0.02 ± 0.00
88	8.965	1171	2-Methyl-3,5-diethylpyrazine	1.07 ± 1.14	0.13 ± 0.02	0.28 ± 0.05	0.20 ± 0.09	0.10 ± 0.05
89	7.864	1069	2-Acetyl-1H-pyrrole	0.15 ± 0.13 ^b	0.00 ± 0.00 ^a	0.02 ± 0.00	0.05 ± 0.05	0.05 ± 0.00
90	9.188	1191	1-Furfurylpyrrole	0.14 ± 0.05	0.03 ± 0.03	0.04 ± 0.01	0.16 ± 0.17	0.02 ± 0.00

			Total	14.98	1.56	1.77	4.27	3.82
			Sesquiterpene hydrocarbons					
91	10.972	1379	Farnesan	0.05 ± 0.04	0.01 ± 0.01	0.02 ± 0.00	0.01 ± 0.00	0.05 ± 0.02
92	11.13	1396	α -Ylangene	0.45 ± 0.35 ^e	0.10 ± 0.07 ^e	0.26 ± 0.05 ^e	0.18 ± 0.08 ^e	3.04 ± 0.04 ^{a-d}
93	11.497	1431	(<i>E</i>)- β -Farnesene	0.01 ± 0.01 ^c	0.01 ± 0.00 ^c	2.19 ± 0.31 ^{a,b,d,e}	0.03 ± 0.02 ^c	0.14 ± 0.01 ^c
94	11.548	1437	Cedrene	0.02 ± 0.00 ^c	0.04 ± 0.01 ^c	0.80 ± 0.15 ^{a,b,d,e}	0.26 ± 0.17 ^b	0.05 ± 0.00 ^c
95	11.618	1444	β -Caryophyllene*	0.07 ± 0.03 ^b	8.51 ± 7.34 ^{a,d}	1.79 ± 0.26	0.38 ± 0.29 ^b	0.16 ± 0.03
96	11.741	1455	(<i>Z</i>)- β -Farnesene	0.06 ± 0.04 ^c	0.09 ± 0.05 ^c	0.76 ± 0.09 ^{a,b,d,e}	0.19 ± 0.18 ^c	0.06 ± 0.00 ^c
97	12.299	1510	β -Bisabolene	0.02 ± 0.02 ^{c,e}	0.07 ± 0.06 ^c	1.28 ± 0.18 ^{a,b,d,e}	0.26 ± 0.22 ^c	0.38 ± 0.01 ^{a,c}
98	12.504	1530	δ -Cadinene	0.06 ± 0.06	0.01 ± 0.00	0.02 ± 0.00	0.47 ± 0.61	0.02 ± 0.00
99	12.556	1535	Calamenene	0.36 ± 0.32 ^e	0.10 ± 0.03 ^{c,e}	0.56 ± 0.09 ^{b,d,e}	0.10 ± 0.06 ^{c,e}	7.06 ± 0.31 ^{a-d}
100	12.802	1559	β -Calacorene	0.10 ± 0.06 ^e	0.06 ± 0.03 ^e	0.15 ± 0.03 ^e	0.19 ± 0.13 ^e	2.63 ± 0.10 ^{a-d}
			Total	1.21	8.99	7.83	2.07	13.59

Different letters (a-e) indicate significant differences between *G. bruguieri* root specimens according to least significant difference (LSD) analysis ($p < 0.05$; Tukey's test).

Fig. 1: Structure of the major secondary metabolites detected in methanol extracts of the four *G. bruguieri* root specimens. Carbon numbering system for each compound is based on analogy rather than IUPAC rules. Metabolite numbers follow those listed in **Table (1)** for their identification using 1D- and 2D-NMR.

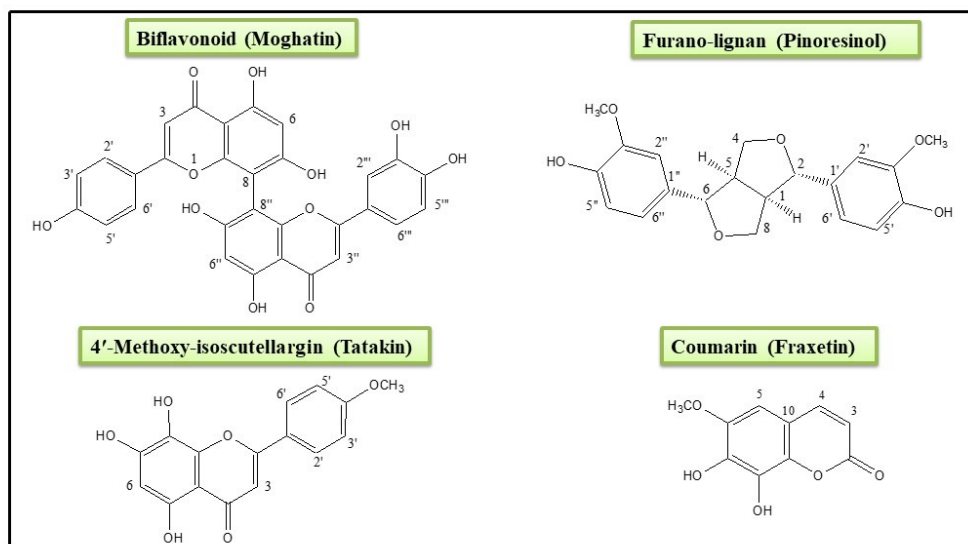


Fig. 2: ^1H -NMR spectrum of *G. bruguieri* (Desf.) root methanol extracts (GBSE, GBAE, GBSA and GBRE) showing characteristic signals for major primary metabolites in the spectral range δ 1.25-5.5 ppm (**A**); displaying the signals of: (**N1**) α -Sucrose, (**N2**) β -D-Glucose, (**N3**) α -D-Glucose, (**N4**) Inositol, (**N5**) Fatty acids, (**N6**) Acetic acid, (**N7**) Succinic acid; and secondary metabolites in the range δ 6.16-8.18 ppm (**B**), displaying the signals of: (**N8**) Fumaric acid, (**N9**) Fraxetin, (**N10**) Moghatin, (**N11**) 4'-Methoxy-isoscutellargin (Tatakin) (**N12**) Pinoresinol. Signal numbers follow those listed in **Table (1)**.

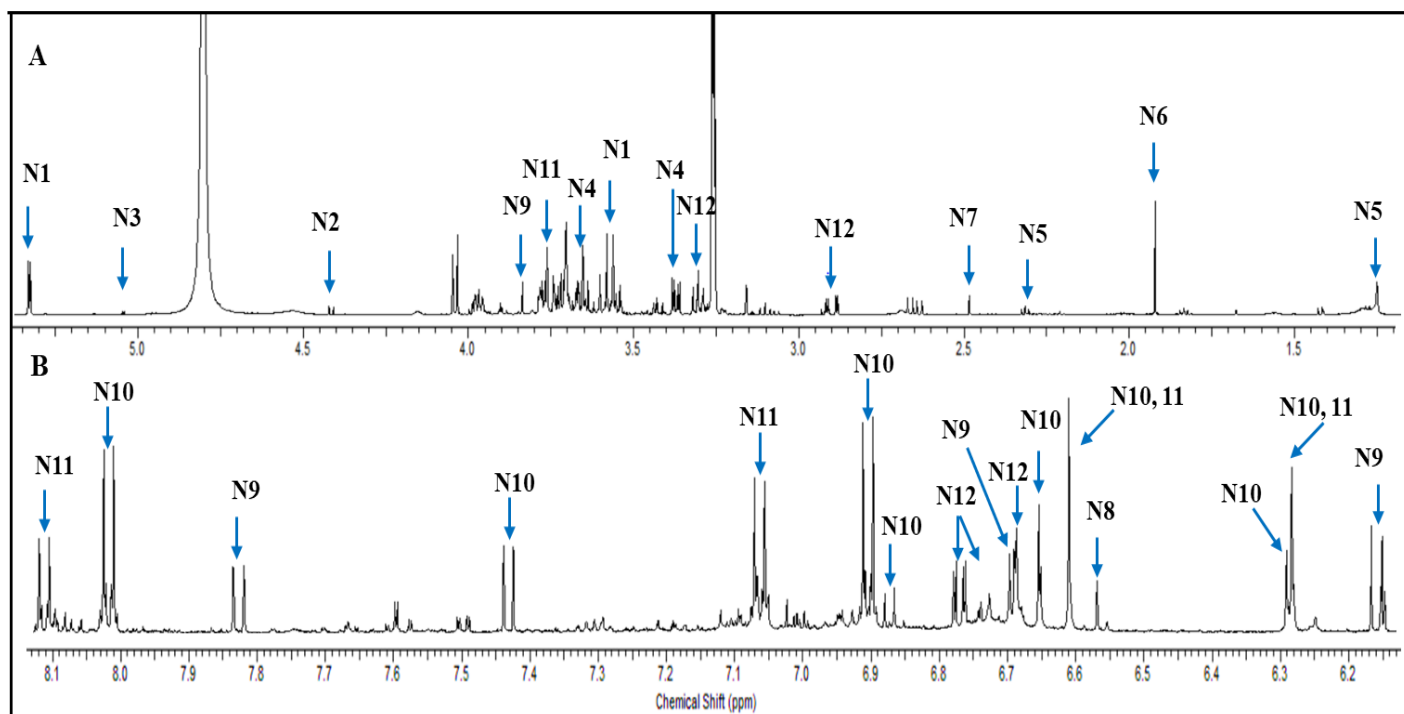


Fig. 3: ^1H -NMR (10.0-0.0 ppm) peak-based principal component analysis of the four studied *G. bruguieri* (Desf.) specimens (n=3). The clusters are located at distinct positions in two-dimensional space described by two vectors of principal component PC1 (0.616) and PC2 (0.207). **(A)** Score plot of PC1 vs. PC2 scores. **(B)** Loading plot for PC1 & PC2 contributing ^1H -NMR signals and their assignments, with each metabolite denoted by its chemical shift (ppm). It should be noted that ellipses do not denote statistical significance but are rather added for better visibility of clusters discussed. For sample codes, refer to experimental section 2.1.

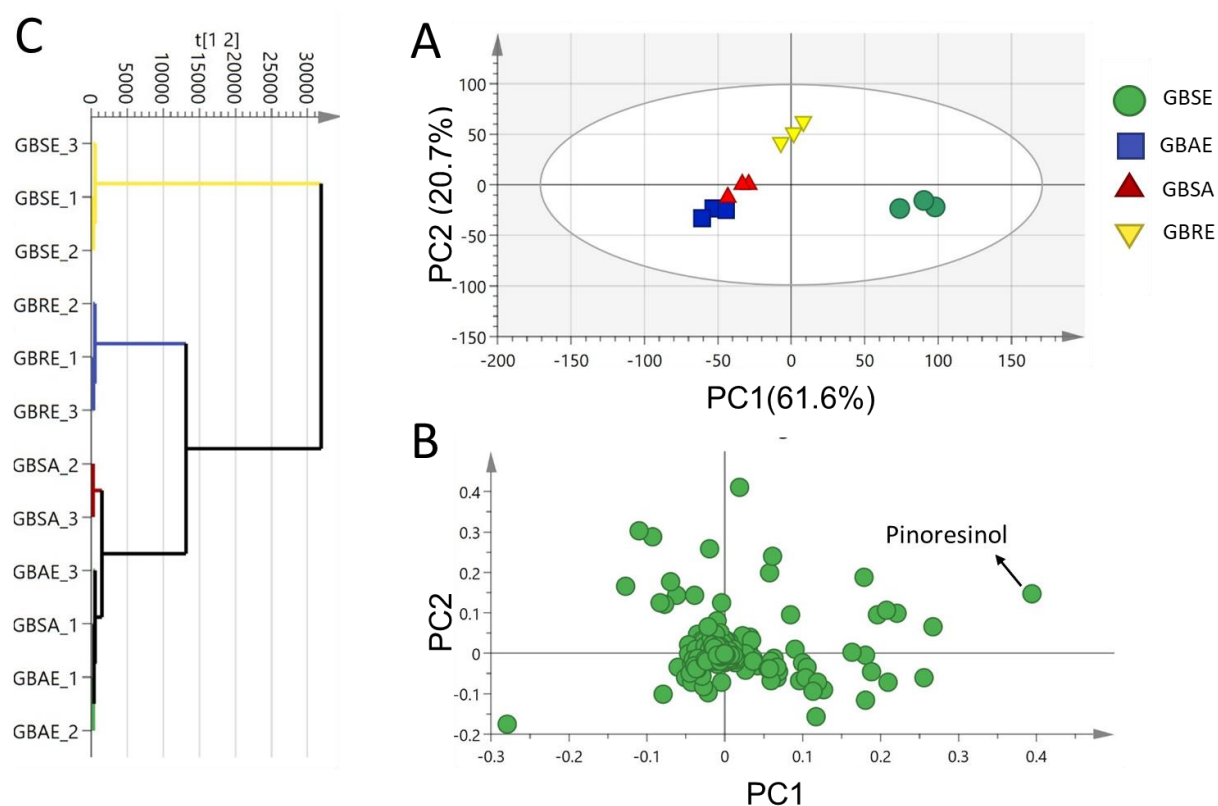


Fig. 4: ^1H -NMR aromatic region (10.0-5.5 ppm) based orthogonal projection to latent structures-discriminant analysis (OPLS-DA) of the four *G. bruguieri* (Desf.) specimens modelled against one another (Δ) ($n=3$). (A) OPLS-DA score plot (B) loading plot derived from samples modelled against each other. Peak numbering follow those listed in **Table (1)** for metabolite identification using 1D and 2D-NMR.

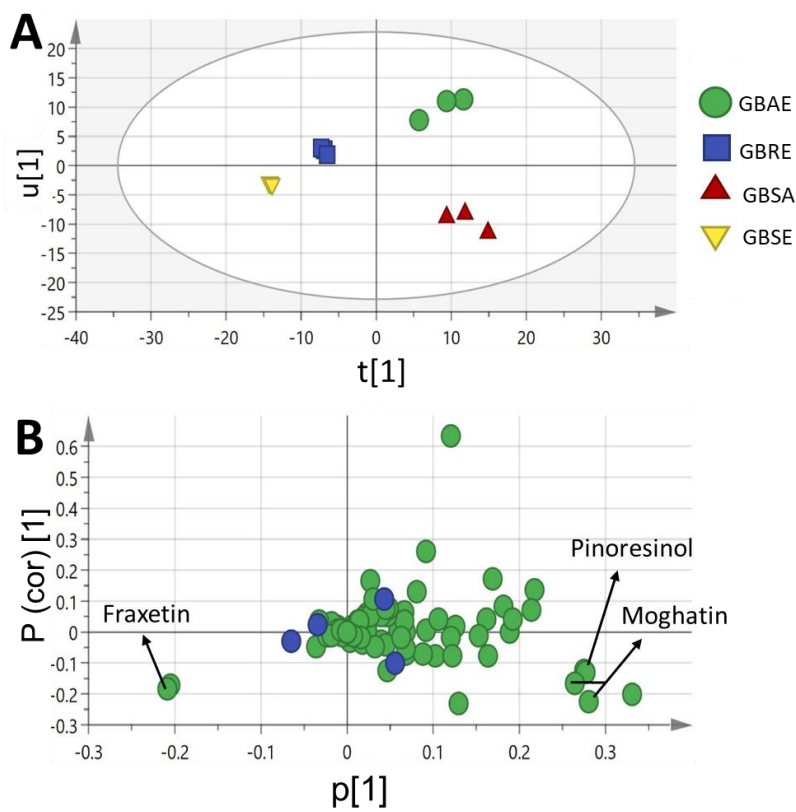


Fig. 5: SPME-GC-MS chromatogram of *G. bruguieri* (Desf.), (A) roasted from Siwa, (B) unroasted from Siwa, (C) unroasted from Aswan. The corresponding metabolites names for each peak follow that listed in **Table (3)**.

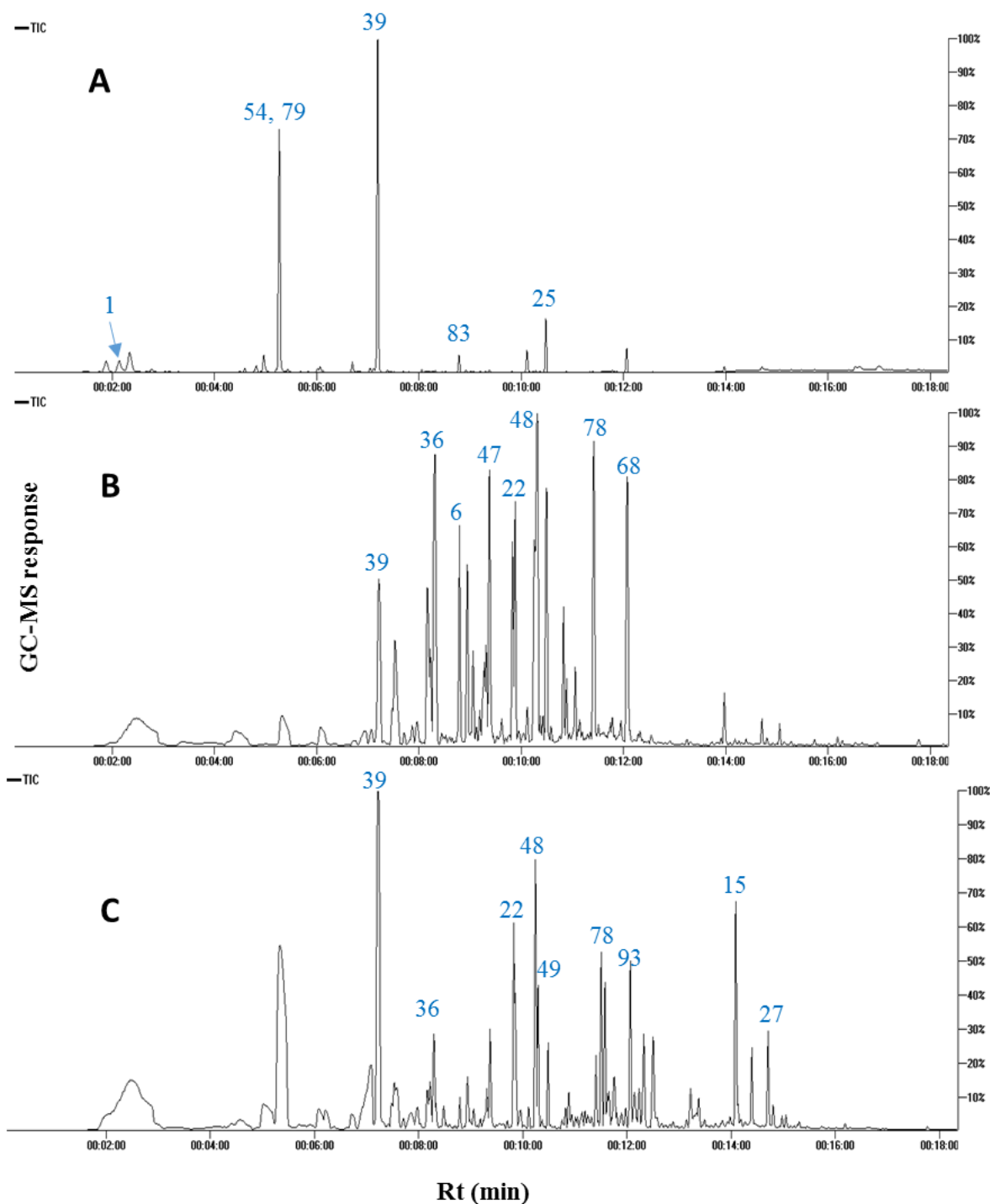


Fig. 6: OPLS score plot derived from modelling the sundried *G. bruguieri* (Desf.) specimens “GBSE” as one class group versus roasted ones “GBSE_R” (A). The S-plot (B) shows the covariance $p[1]$ against the correlation $p(\text{cor})[1]$ of the variables of the discriminating component of the OPLS-DA model. Cut-off values of $P < 0.05$ were used; selected variables are highlighted in the S-plot with identifications are discussed in text.

