Hindawi
Journal of Chemistry
Volume 2023, Article ID 9463366, 8 pages
https://doi.org/10.1155/2023/9463366

Research Article

Careyarboside A: A New Flavan-3-ol Glycoside from the Stems of *Careya arborea* (Lecythidaceae)

Yeon Woo Jeong,[1] Jung A. Lee,[1] Eun-Kyung Ahn,[1] Jae-Shin Kang,[2] Hye Woo Byun,[2] H. E. Keo Omaliss,[3] Chun Whan Choi,[1] and Seong Su Hong [1]

[1]*Bio-Center, Gyeonggido Business & Science Accelerator, Suwon 16229, Republic of Korea*
[2]*Biological Genetic Resources Utilization Division, National Institute of Biological Resources (NIBR), Incheon 22689, Republic of Korea*
[3]*Forestry Administration, Ministry of Agriculture Forestry and Fisheries, Phnom Penh, Cambodia*

Correspondence should be addressed to Seong Su Hong; bestgene@gbsa.or.kr

Received 8 February 2023; Revised 25 August 2023; Accepted 12 September 2023; Published 25 September 2023

Academic Editor: Michele Benedetti

A new flavan-3-ol glycoside, compound **1** (careyarboside A), was isolated from the ethanol extract of the stems of *Careya arborea*, together with six known compounds. Of these, three phenolic glycosides (leonuriside A, picraquassioside D, and careyarboside B) have never been previously isolated from genus *Careya* and family Lecythidaceae. Their chemical structures were identified using spectroscopic data, specifically HRESIMS, 1D and 2D NMR data, including ^{1}H-^{1}H COSY, HSQC, and HMBC. In addition, the chemical structures of the known compounds were identified through spectral comparisons with those of published values.

1. Introduction

Careya arborea Roxb. (Lecythidaceae), commonly known as wild guava in English and kumbhi in Hindi, is a large evergreen tree found in deciduous forests and grasslands throughout Sri Lanka, India, and the Malay Peninsula up to an altitude of 1500 meters [1]. The dried stem bark of this plant has been widely used as traditional medicine for the treatment of tumors, inflammation, epileptic fits, skin disease, dysentery with bloody stools, dyspepsia, ulcer, and chronic ear pain, as well as an anthelmintic, astringent, and antidote to snake venom [1–3]. *C. arborea* has a wide range of pharmacological activities, such as antitumor, antioxidant, antihyperlipidemic, antibacterial, antimicrobial, hepatoprotective, anti-inflammation, analgesic, gastroprotective, antiasthmatic, antifertility, and antidiarrheal effects [4–14]. The pharmacognostical properties are mainly determined by the various natural compounds in *C. arborea*, such as triterpenoids, phenolics, flavonoids, sterols, coumarin, saponins, and tannins [1, 15–19].

In our continuous efforts to discover more bioactive components, a new flavan-3-ol glycoside, careyarboside A (**1**) was isolated from the stems of *C. arborea*, along with six known compounds, (+)-gallocatechin (**2**), (−)-epigallocatechin (**3**), leonuriside A (**4**), picraquassioside D (**5**), 2,3-dihydroxyphenylethanol-1-*O*-*β*-D-glucopyranoside (**6**), and gallic acid (**7**). Here, we described the extraction, isolation, and structural characterization of compounds **1**–**7**.

2. Materials and Methods

2.1. General Experimental Procedures. A JASCO P-2000 polarimeter was used to determine the optical rotation (Tokyo, Japan). Using a Bruker Ascend III 700 spectrometer, NMR spectra were recorded in methanol-d_4 and pyridine-d_5 at room temperature (Bruker BioSpin GmbH., Rheinstetten, Germany). ESI-MS spectra were acquired on an Agilent 6130 series quadrupole LC/MS System (Agilent Technologies, CA, USA). HRESIMS spectra were recorded by a Triple TOF 5600+ mass spectrometer (AB SCIEX, MA, USA). Open column chromatography was performed using Diaion

HP-20 adsorbent resin (Mitsubishi Chemical Corp, Tokyo, Japan) and Lichroprep RP-18 (40–63 μM, Merck, NJ, USA). MPLC was performed with a CombiFlash Rf flash chromatography system (Teledyne ISCO Inc., NE, USA), and the separations were carried out using a Redi*Sep*® Rf Silica and Gold® C_{18} column. TLC was performed using DC-Fertigfolien ALUGRAM® SIL G/UV$_{254+366}$ (0.2 mm, Macherey-Nagel GmbH & Co. KG, Düren, Germany) plates, and spots were visualized by a 10% vanillin-sulfuric acid reagent (10% mixture of conc. H_2SO_4 soln. and 1% vanillin in EtOH). All chemicals and solvents were of analytical quality and were utilized without further purification.

2.2. Plant Material. In September 2015, stems of *C. arborea* were taken from Cambodia's Seima Protection Forest for the present investigation. Botanical identification was made by one of the authors (Dr. H. E. Keo Omaliss), and a voucher specimen (CA101) of the plant was placed at the laboratory of natural products research team, Gyeonggido Business & Science Accelerator.

2.3. Extraction and Isolation. The shade-dried stems of *C. arborea* (2.2 kg) were macerated with 50% aqueous EtOH at room temperature. Following the evaporation of the solvent under reduced pressure, the residue (165 g) was dissolved in water (5 L) and sequentially partitioned with *n*-hexane, CH_2Cl_2, EtOAc, and *n*-BuOH sat.H_2O to yield extracts dry weighing 10, 17, 14, and 29 g, respectively. Part of the *n*-BuOH-soluble fraction was chromatographed over a Diaion HP-20 resin and eluted with a water-MeOH stepwise gradient solvent system (1 : 0–0 : 1) to give six fractions (CA101A$_1$–CA101A$_6$). Of these, fraction CA101A$_2$ (1.7 g) was fractionated on medium pressure liquid chromatography (MPLC) over an ODS column (RediSep® Rf silica gold 240 g, 100 mL/min, 50 min, detection at 210 and 254 nm) by eluting with a gradient mixture of 5–50% acetonitrile in H_2O to afford 12 subfractions (CA101B$_1$–CA101B$_{12}$). CA101B$_2$ (250 mg) was subjected to Sephadex LH-20 column chromatography (MeOH, isocratic) to obtain pure compounds **2** (25.8 mg), **6** (11.3 mg), and **7** (8.4 mg). Subfraction CA101B$_7$ (76.1 mg) was subjected to the abovementioned MPLC system (Silica gel, 12 g Gold, CH_2Cl_2–MeOH–H_2O, 10 : 3:1 to 7 : 3 : 1, v/v, 18 mL/min, 55 min, detection at 210 and 254 nm) resulting in the isolation of compounds **3** (16.6 mg, t_R = 9.5 min), **4** (4.6 mg, t_R = 14.8 min), and **5** (5.5 mg, t_R = 19.5 min). Further purification of subfraction CA101B$_{10}$ (145.5 mg) was chromatographed on a silica gel column and eluted with a CH_2Cl_2–MeOH–H_2O stepwise gradient system (10 : 3 : 1, 7 : 3 : 1, 6 : 4 : 1) to give compounds **1** (26.6 mg) and **3** (36.2 mg). The isolation process used in the present study is summarized in Scheme S1 (Supplementary material).

2.4. Acid Hydrolysis. A solution of compound **1** (2 mg) in 2 M HCl/MeOH (4 : 1, 5 mL) was heated in a water bath for 6 h with the temperature at 80°C. The reaction mixture was cooled, then diluted with water (15 mL), neutralized with $NaHCO_3$, and extracted with EtOAc (3 × 15 mL). For sugar analysis, the aqueous layer was concentrated to an adequate level (1 mL) and analyzed by TLC (silica gel) using the solvent system $CHCl_3$/MeOH/H_2O (70 : 30 : 1). Using a vanillin-sulfuric acid spray followed by heating, a spot was detected. The R_f value of glucose was 0.25. Thus, the sugar component was identified as glucose. The optical rotation of the water solution was determined to be $[\alpha]_D^{25}$ + 26.8° (c = 0.035, H_2O) for compound **1**. Positive optical rotation indicated that this sugar is D-configuration (lit. $[\alpha]_D^{25}$ + 52.5°) [20]. On the other hand, the aglycone migrated into the EtOAc layer. The combined EtOAc-soluble extracts were evaporated to dryness to yield the aglycone. For further purification, the aglycone was subjected to flash column chromatography on RP-18 and eluted with an MeCN – H_2O gradient system (10 : 90 to 60 : 40, v/v) to yield aglycone of compound **1**. The optical rotation value of aglycone was $[\alpha]_D^{25}$ – 47.1° (c = 0.25, CH_3OH).

2.5. Apparatus and Chromatographic Conditions. On a Waters alliance e2695 (Waters Co., Milford, MA, USA) system with a 2998 PDA detector and column heater/cooler with passive pre-heater, HPLC analysis was conducted. Using a YMC-Triart C18 column (250 × 4.6 mmI.D., 5 μm particle size), separation was accomplished (YMC Co., Ltd., Japan). The mobile phase consisted of water–TFA (99.95 : 0.05; v/v) (solvent A) and acetonitrile (solvent B). The elution was carried out utilizing the following gradient mobile phase: initial 95 : 5 (A : B v/v); 40 min 60 : 40 (A : B v/v). The column temperature was set at 25°C, and the volume of injection was 10 μL. Empower™ 3 software was responsible for all operations, data acquisition, and analysis (Waters Co., Milford, MA, USA).

3. Results and Discussion

*3.1. Isolation and Purification of Compounds **1**–**7**.* Dried stems of *C. arborea* were crushed and extracted with 50% EtOH at room temperature by rotary evaporation to obtain the crude EtOH extract. The EtOH extract was sequentially employed in the solvent partition process with four solvents, *n*-hexane, methylene chloride (CH_2Cl_2), EtOAc, and aqueous-saturated *n*-butanol (BuOH), which yielded four main solvent fractions with increasing polarity. Chromatographic purification of the BuOH-soluble fractions yielded the new flavan-3-ol glycoside (**1**), along with compounds **2**–**7** (Figure 1).

*3.2. Spectroscopic Data of Compounds **1**–**7**. Careyarboside A* (**1**) White amorphous powder; $[\alpha]_D^{25}$ – 24.4° (c = 0.12, CH_3OH), $[\alpha]_D^{25}$ – 47.1° (c = 0.25, CH_3OH, aglycone); UV (MeOH) λ_{max} nm: 205, 272; ^{1}H (700 MHz) and ^{13}C NMR (175 MHz), see Tables 1–3; ESIMS *m/z* 321 $[M + H - 162]^+$, 483 $[M + H]^+$, 505 $[M + Na]^+$, 481 $[M - H]^-$; ESIMS *m/z*. HRESIMS (positive mode) *m/z* 483.1504 $[M + H]^+$ (calcd for $C_{22}H_{27}O_{12}$, 483.1503).

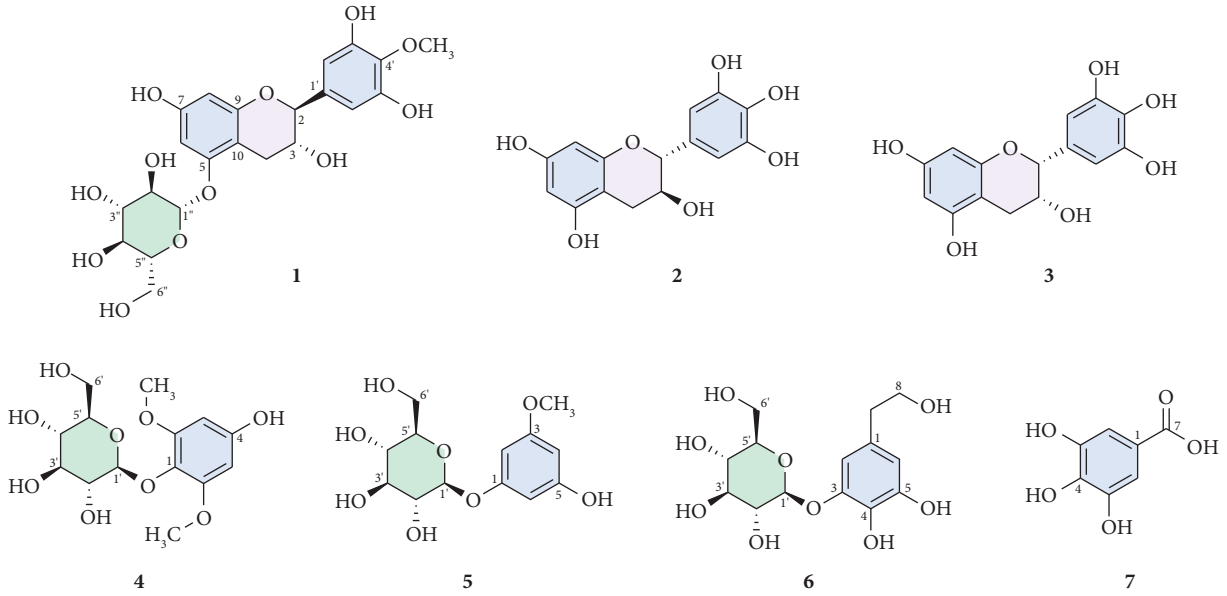

Figure 1: Chemical structures of compounds **1**–**7** identified from *C. arborea*.

Table 1: ^{1}H (700 MHz) data of compounds **1**–**3** in pyridine-d_5 (δ in ppm)[a].

Position	**1**	**2**	**3**
	δ (J in Hz)		
2	5.18 d (7.0)	5.29 d (7.0)	5.38 brs
3	4.55 m	4.68 m	4.76 m
4	3.50 dd (16.1, 4.9) 3.32 dd (16.1, 7.7)	3.67 dd (16.1, 5.6) 3.36 dd (16.1, 7.7)	3.54 dd (16.1, 3.5) 3.40 dd (16.1, 4.9)
6	7.10 d (2.1)	6.71 d (2.1)	6.70 d (2.1)
8	6.72 d (2.1)	6.66 d (2.1)	6.66 d (2.1)
2′, 6′	7.18 s	7.29 s	7.48 s
1″	5.64 d (7.7)		
2″	4.34[b]		
3″	4.34[b]		
4″	4.41 t (9.1)		
5″	3.96 m		
6″	4.39 dd (11.2, 4.9)		
OCH_3	4.00 s		

[a]Assignments confirmed by ^{1}H-^{1}H COSY, HSQC, and HMBC experiments. [b]Overlapped signals are reported without designating multiplicity.

(+)-Gallocatechin (**2**). White amorphous powder; $[\alpha]_D^{25}$ + 10.8° (c = 0.2, CH_3OH); UV (MeOH) λ_{max} nm: 205, 269; ^{1}H (700 MHz) and ^{13}C NMR (175 MHz), see Tables 1 and 3; ESIMS m/z 305 $[M - H]^-$.

(−)-Epigallocatechin (**3**). White amorphous powder; $[\alpha]_D^{25}$ − 36.0° (c = 0.2, CH_3OH); UV (MeOH) λ_{max} nm: 206, 270; ^{1}H (700 MHz) and ^{13}C NMR (175 MHz), see Tables 1 and 3; ESIMS m/z 305 $[M - H]^-$.

Leonuriside A (**4**). White amorphous powder; UV (MeOH) λ_{max} nm: 213, 277; ^{1}H (700 MHz) and ^{13}C NMR (175 MHz), see Tables 2 and 3; ESIMS m/z 355 $[M + Na]^+$, 331 $[M - H]^-$.

Picraquassioside D (**5**). White amorphous powder; UV (MeOH) λ_{max} nm: 212, 270; ^{1}H (700 MHz) and ^{13}C NMR (175 MHz), see Tables 2 and 3; ESIMS m/z 325 $[M + Na]^+$, 301 $[M - H]^-$.

Table 2: ^{1}H (700 MHz) data of compounds **4**–**7** (δ in ppm)[a].

Position	**4**[b]	**5**[b]	**6**[b]	**7**[c] (s)
		δ (J in Hz)		
2		6.70 t (2.1)	7.20 d (2.1)	7.06
3	6.58 s			
4		6.62 t (2.1)		
5	6.58 s			
6		6.91 t (2.1)	7.14 d (2.1)	7.06
7			2.98 t (7.0)	
8			4.07 dt (7.0, 1.4)	
1′	5.62 d (7.7)	5.67 d (7.7)	5.51 d (7.7)	
2′	4.35[d]	4.34 t (9.1)	4.27 t (9.1)	
3′	4.35[d]	4.37 m	4.30 t (9.1)	
4′	4.35[d]	4.38 t (9.1)	4.35 t (9.1)	
5′	3.94 m	4.04 m	3.98 m	
6′	4.43 dd (11.9, 2.1) 4.36 dd (11.9, 4.9)	4.47 dd (11.9, 2.1) 4.38 dd (11.9, 4.9)	4.51 dd (11.9, 2.1) 4.41 dd (11.9, 4.9)	
OCH_3	3.71 s	3.67 s		

[a]Assignments confirmed by ^{1}H-^{1}H COSY, HSQC, and HMBC experiments. [b]Measured for pyridine-d_5 solution. [c]Measured for CD_3OD solution. [d]Overlapped signals are reported without designating multiplicity.

4,5-Dihydroxyphenylethanol-3-O-β-D-glucopyranoside (Careyarboside B, ***6****).* White amorphous powder; UV (MeOH) λ_{max} nm: 200, 272; ^{1}H (700 MHz) and ^{13}C NMR (175 MHz), see Tables 2 and 3; ESIMS m/z 355 $[M + Na]^+$, 331 $[M - H]^-$.

*Gallic acid (***7***).* White amorphous powder; UV (MeOH) λ_{max} nm: 214, 271; ^{1}H (700 MHz) and ^{13}C NMR (175 MHz), see Tables 2 and 3; ESIMS m/z 169 $[M - H]^-$.

3.3. Identification of Components. MS and various 1D and 2D NMR experiments, such as DEPT, HMBC, ^{1}H-^{1}H COSY, and HSQC, were used to determine the structures of compounds **1**–**7**.

Table 3: ^{13}C (175 MHz) data of compounds **1**–**7** (δ in ppm)[a].

Position	**1**[b]	**2**[b]	**3**[b]	Position	**4**[b]	**5**[b]	**6**[b]	**7**[c]
	δ mult.				δ mult.			
2	83.2 d[d]	83.7 d	80.5 d	1	129.7 s	161.2 s	131.3 s	122.1 s
3	67.9 d	68.5 d	67.3 d	2	154.9 s	94.9 d	111.8 d	110.5 d
4	29.1 t	29.5 t	29.8 t	3	95.4 d	162.7 s	147.8 s	146.5 s
5	158.3 s	158.7 s	158.9 s	4	156.2 s	97.1 d	136.7 s	139.7 s
6	97.1 d	96.9 d	96.9 d	5	95.4 d	161.3 s	148.5 s	146.5 s
7	159.1 s	159.0 s	158.9 s	6	154.9 s	98.0 d	113.6 d	110.5 d
8	98.2 d	95.8 d	96.1 d	7			40.5 t	170.5 s
9	156.8 s	157.6 s	157.9 s	8			64.2 t	
10	103.0 s	101.2 s	100.6 s	1′	106.2 d	102.6 d	105.6 d	
1′	137.0 s	132.0 s	131.7 s	2′	76.4 d	75.3 d	75.5 d	
2′	107.9 d	107.9 d	108.0 d	3′	78.7 d	78.8 d	78.7 d	
3′	152.7 s	148.4 s	148.1 s	4′	72.0 d	71.5 d	71.4 d	
4′	137.1 s	135.4 s	135.1 s	5′	78.9 d	79.2 d	79.3 d	
5′	152.7 s	148.4 s	148.1 s	6′	63.0 t	62.6 t	62.5 t	
6′	107.9 d	107.9 d	108.0 d	OCH_3	56.7 q	55.5 q		
1″	102.7 d							
2″	75.3 d							
3″	79.0 d							
4″	71.3 d							
5″	79.1 d							
6″	62.5 t							
OCH_3	60.5 q							

[a]Assignments confirmed by ^{1}H-^{1}H COSY, HSQC, and HMBC experiments. [b]Measured for pyridine-d_5 solution. [c]Measured for CD_3OD solution. [d]Carbon multiplicity deduced by HSQC and DEPT.

Compound **1** was isolated in the form of a white amorphous powder, ($[\alpha]_D^{25}$ −24.4°) (c = 0.2, MeOH), exhibiting a UV absorption maximum at 205 and 272 nm. The molecular formula, $C_{22}H_{26}O_{12}$, was deduced from its HRESIMS data (*m/z* 483.1504 $[M+H]^+$, calcd for $C_{22}H_{27}O_{12}$, 483.1503), indicating 10 degrees of unsaturation. The NMR spectra data (Table 1) combined with the HSQC spectrum revealed a signal characteristic of a flavan-3-ol skeleton with a phloroglucinol pattern for ring A (δ_H 7.10 (1H, d, J = 2.1 Hz, H-6; δ_C 97.1), δ_H 6.72 (1H, d, J = 2.1 Hz, H-8; δ_C 98.2)), a pyrogallol pattern for ring B (δ_H 7.18 (2H, s, H-2′, 6′; δ_C 107.9)), and a catechin moiety for ring C (δ_H 5.18 (1H, d, J = 7.0 Hz, H-2; δ_C 83.2), 4.55 (1H, m, H-3; δ_C 67.9), 3.50 (1H, dd, J = 16.1, 4.9 Hz, H_a-4) and 3.32 (1H, dd, J = 16.1, 7.7 Hz, H_b-4); δ_C 29.1)) [21]. In addition, the remaining ^{1}H NMR signals of **1** revealed the presence a methoxy group (δ_H 4.00 (3H, s; δ_C 60.5)) and one anomeric proton at δ_H 5.64 (1H, d, J = 7.7 Hz) correlated with the HSQC spectrum with the anomeric carbon at δ_C 102.7 in the ^{13}C NMR spectrum. The ^{13}C NMR spectrum revealed the presence of 22 carbon signals, of which six were ascribed to the sugar moieties (δ_C 102.7, 79.1, 79.0, 75.3, 71.3, and 62.5) and the remaining 16 were ascribed to the aglycone part, including two oxygenated methine (δ_C 83.2 and 67.9), a methylene at δ_C 29.1, four aromatic carbons (δ_C 107.9 × 2, 98.2, and 97.1), one methoxy group at δ_C 60.5, and eight aromatic quaternary carbons (δ_C 159.1, 158.3, 156.9, 152.7 × 2, 137.1, 137.0, and 103.0). The fragment ion peak at *m/z* 321 $[M+H-162]^+$ suggested a hexose unit. The ^{13}C NMR spectrum revealed not only signals corresponding to the dihydroflavonol skeleton and one methoxy group but also signals indicative of a hexose monosaccharide moiety. Based on these data, **1** was identified as a flavan-3-ol glycoside. In the HMBC spectrum, the one methoxy group at δ_H 4.00 showed a long-range correlation with C-4′ (δ_C 137.1), indicating its attachment to C-4′. The anomeric proton of sugar at δ_H 5.64 showed a correlation with C-5 (δ_C 158.3), which suggests that the hexose moiety is located at C-5 (Figure 2). The connections through–CH(2)–CH(3)–CH_2(4)– can also be derived from the correlations identified by ^{1}H–^{1}H COSY. The doublet at δ_H 5.18 (H-2, J = 7.0 Hz; C-2 at δ_C 83.2) and the multiplet at δ_H 4.55 (H-3; C-3 at δ_C 67.9) are characteristic of the 2,3-*trans* relationship of the flavan-3-ol unit [21–25]. The D-glucose identity of the free sugar was established based on acid hydrolysis and TLC analysis of an actual sample. The ^{1}H-NMR coupling constant of the anomeric proton ($J_{1'',2''}$ > 7.0 Hz) indicated the β-configuration of the sugar moiety [26]. Furthermore, the aglycone was found to have a negative optical rotation value ($[\alpha]_D^{25}$ −47.1°) [27, 28]. The analysis of 1D and 2D NMR spectroscopic data revealed the same planar structure as (−)-4′-methylepigallocatechin 5-*O*-β-glucopyranoside [29] and (+)-gallocatechin 4′-*O*-methyl ether 5-*O*-β-D-glucopyranoside [30], suggesting that **1** was a stereoisomer of these compounds (Figure 3). The singlet signal at H-2 and the multiplet signal at H-3 were characteristic of the 2,3-*cis* configuration of the (−)-4′-methylepigallocatechin 5-*O*-β-glucopyranoside [29]. Moreover, the stereochemistry of (+)-gallocatechin 4′-*O*-methyl ether 5-*O*-β-D-glucopyranoside was determined based on the positive optical rotation value; however, its detailed spectroscopic data, including the 1D-NMR spectra, were not reported. Therefore, **1** was

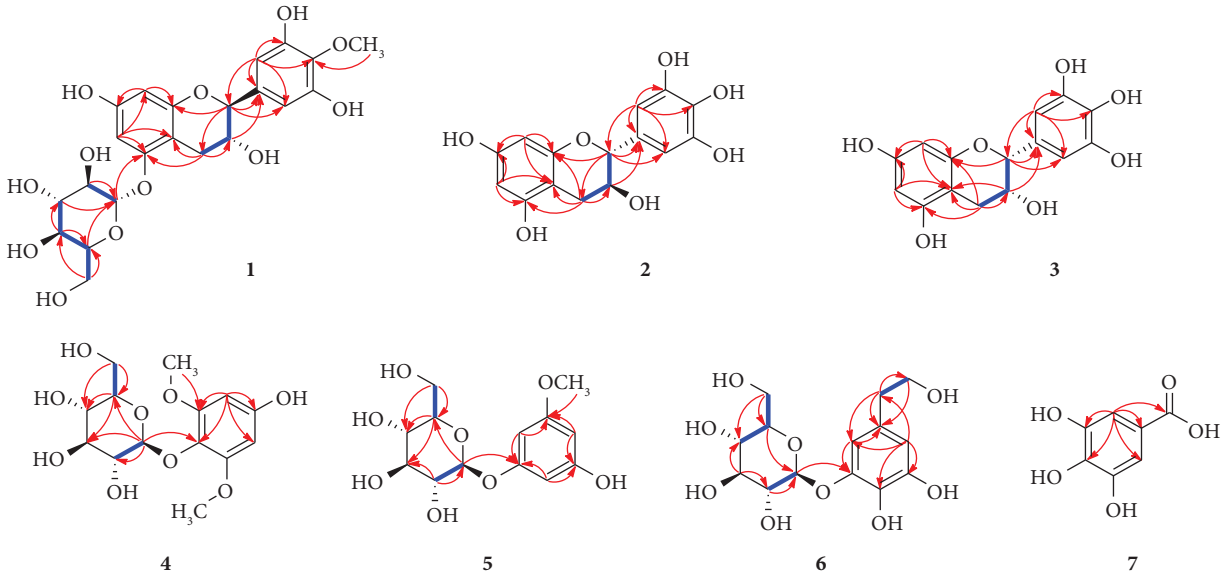

Figure 2: Key ^{1}H-^{1}H COSY (blue bold lines) and HMBC (red arrows) correlations of **1**–**7**.

identified as (−)-gallocatechin 4′-*O*-methyl ether 5-*O*-*β*-D-glucopyranoside and was given the trivial name careyarboside A.

Compounds **2** and **3** showed a UV spectrum similar to that of compound **1**. The ^{1}H-NMR spectra of compounds **2** and **3** were similar to those of compound **1**, showing phloroglucinol signals for ring A at δ_H 6.71–6.70 (1H, d, J = 2.1 Hz, H-6) and 6.66 (1H, d, J = 2.1 Hz, H-8) and pyrogallol signals for ring B at (**2**: δ_H 7.29 (2H, s, H-2′,6′) and **3**: δ_H 7.48 (2H, s, H-2′,6′)). These findings indicate compounds **2** and **3** are flavan-3-ol derivatives. The doublet signal at (**2**: δ_H 5.18 (1H, d, J = 7.7 Hz, H-2)) and singlet at (**3**: δ_H 5.38 (1H, s, H-2)) are characteristic of the 2,3-*trans* (**2**) and 2,3-*cis* (**3**) configurations for the C-2/C-3 position [31]. Additionally, the specific optical rotation of compounds **2** and **3** were $[\alpha]_D^{25}$ + 10.8° and −36.0°, respectively. On the basis of the abovementioned evidence, the structures of **2** and **3** were identified as (+)-gallocatechin (**2**) [32] and (−)-epigallocatechin (**3**) [33]. Compound **4** was isolated in the form of a white amorphous powder. The ^{1}H NMR spectrum revealed signals of one singlet at δ_H 6.58 (2H, s, H-3,5) assigned to two aromatic protons; one singlet at δ_H 3.71 (6H, s) assigned to two methoxyl groups; and a sugar moiety (δ_H 5.62 (1H, d, J = 7.7 Hz, H-1′) and 4.36–3.91 (6H)). In HMBC, a correlation between the two aromatic protons (δ_H 6.58, H-3,5) and four oxygenated aromatic carbons at δ_C 156.2 (C-4), 154.9 (C-2, 6), and 129.7 (C-1) was found. Signals characteristic of the *β*-anomeric form (δ_H 5.62, d, J = 7.7 Hz) of glucopyranosyl were also observed. The attachment of the sugar moiety at C-1 was proven by the HMBC correlation between the anomeric protons at δ_H 5.62 and the carbon signal at δ_C 129.7. A methoxyl group was positioned each at C-2 and C-6 (δ_C 154.9) based on the HMBC correlations of this carbon with the methoxyl group and the two aromatic protons H-2 and H-6 (δ_H 6.58) (Figure 2). The NMR data of compound **4** are consistent with the spectral data in the literature [34]; thus, compound **4** was identified as leonuriside A. In the ESI mass spectrum, compound **5** showed a quasi-molecular ion peak $[M-H]^-$ at *m/z* 301 corresponding to $C_{13}H_{18}O_8$; the *m/z* value was 30 amu less than that of **4**, which indicates the absence of methylene (CH_2) and hydroxy groups in **4**. The ^{1}H-NMR spectrum of **5** revealed three *meta*-coupled aromatic proton signals δ_H 6.91 (1H, d, J = 2.1 Hz, H-6), 6.70 (1H, d, J = 2.1 Hz, H-2), and 6.62 (1H, d, J = 2.1 Hz, H-4), suggesting the presence of a 1,3,5-trisubstituted benzene ring as well as a methoxy group at δ_H 3.67 (3H, s). The ^{1}H and ^{13}C NMR spectra also indicated sets of signals belonging to a *β*-glucopyranosyl group (δ_H 5.67 (1H, d, J = 7.7 Hz, H-1); δ_C 102.6 (C-1′), 79.2 (C-5′), 78.8 (C-3′), 75.3 (C-2′), 71.5 (C-4′), 62.6 (C-6′)). The HMBCs between δ_H 3.67 (OCH_3) and δ_C 162.7 (C-3), as well as that between δ_H 5.67 (H-1′) and δ_C 161.2 (C-1), confirmed the positions of the methoxy group and glucopyranosyl unit (Figure 2). The MS and NMR data of compound **5** agreed well with those reported in the literature [35]; thus, compound **5** was identified as picraquassioside D. Compound **6** was purified as a white amorphous powder with a molecular weight of a quasi-molecular ion peak at *m/z* 331 $[M-H]^-$, based on the ESI-MS data. The ^{1}H (Table 1) and ^{13}C NMR data for **6** suggest the presence of a 3,4,5-trihydroxy-1-substituted benzyl group at (δ_H 7.20 (1H, d, J = 2.1 Hz, H-2) and 7.14 (1H, d, J = 2.1 Hz, H-6); δ_C 148.5 (C-5), 147.8 (C-3), 136.7 (C-4), and 131.3 (C-1)), one oxygenated methylene signal (δ_H 4.07 (2H, dt, J = 7.0, 1.4 Hz, H-8); δ_C 64.2 (C-8)), a methylene signal (δ_H 2.98 (2H, t, J = 7.0 Hz, H-7); δ_C 40.5 (C-7)), and a doublet at δ_H 5.51 (1H, d, J = 7.7 Hz, H-1′) together with six others proton signals between δ_H 4.51 and 3.98, suggesting the presence of a sugar moiety. The low-field carbon chemical shift of the methylene at δ_H 4.07 (δ_C 64.2) and the key ^{1}H-^{1}H COSY (Figure 2) correlation with H-7 (δ_H 2.98)/H-8 (δ_H 4.07) indicated the presence of a terminal hydroxyl group in the ethyl chain. The HMBC correlations observed between the oxygenated methylene (δ_H 4.07, H_2-8) to C-1 (δ_C 131.8) and methylene (δ_H 2.98, H_2-7) to C-1 (δ_C

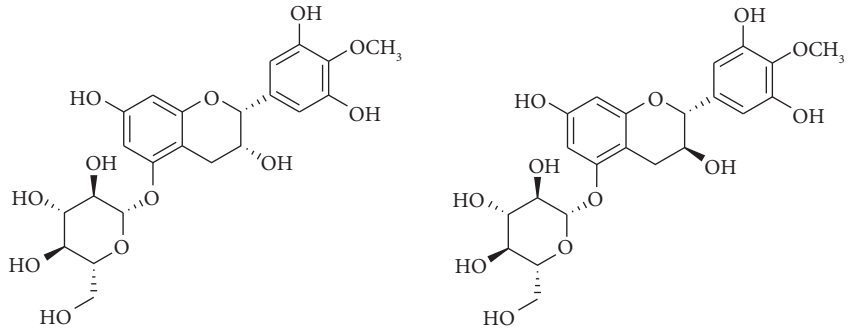

Figure 3: Previously isolated stereoisomer of compound **1**.

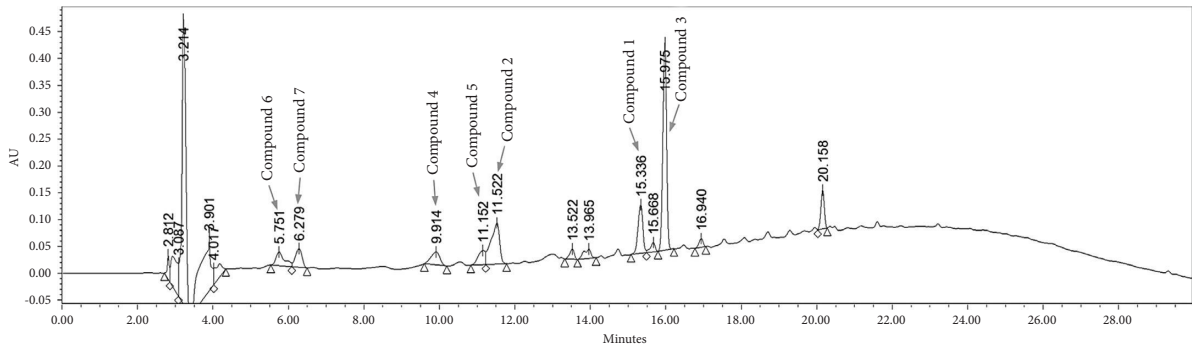

Figure 4: HPLC–PDA chromatographic profile of the EtOH extract from *Careya arborea* stems was recorded at 210 nm.

131.8), C-2 (δ_C 111.8), and C-6 (δ_C 113.6) indicate that the ethyl chain is attached to the C-1 of the aromatic ring. Three oxygenated aromatic carbons were established in the HMBC couplings at δ_C 148.5 (C-5), 147.8 (C-3), and 136.7 (C-4) via correlations with aromatic protons at δ_H 7.20 (H-2) and 7.14 (H-6) (Figure 2). Finally, the sugar moiety was linked to C-3 on the basis of the HMBC correlation from the anomeric proton (δ_H 5.51, H-1′) to C-3 (δ_C 147.9). Based on these data, **6** was identified as 4,5-dihydroxyphenylethanol-3-*O*-*β*-D-glucopyranoside and was given the trivial name careyarboside B [36]. The ^{1}H-NMR spectrum of compound **7** showed one singlet at δ_H 7.06 (2H, s, H-2, 6) attributable to galloyl protons. Based on the proton spectral data of compound **7** and a comparison of the ^{13}C-NMR spectrum of compound **7**, using with literature data, we identified it as gallic acid [37]. The HPLC profiles of the isolated compounds are shown in Figure 4.

4. Conclusion

Seven compounds were isolated from the stems of *C. arborea*, including a new flavan-3-ol glycoside (**1**), with two known flavan-3-ols (**2** and **3**) and four known phenolics (**4–7**). The compounds were identified as careyarboside A (**1**), (+)-gallocatechin (**2**), (−)-epigallocatechin (**3**), leonuriside A (**4**), picraquassioside D (**5**), 4,5-dihydroxyphenylethanol-3-*O*-*β*-D-glucopyranoside (**6**), and gallic acid (**7**). Of these, compounds **4** (leonuriside A), **5** (picraquassioside D), and **6** (4,5-dihydroxyphenylethanol-3-*O*-*β*-D-glucopyranoside) have not previously been reported to have been isolated from any species of the Lecythidaceae family, to the best of our knowledge. Compound **4** was first isolated from *Coix lachryma-jobi* var. *ma-yuen* [29], whereas compound **5** has been isolated from *Sedum sediforme* [38]. Furthermore, compound **6** has been isolated from *Spondias tuberosa* [36]. This study provides valuable insights into the diverse chemical composition of *C. arborea*. Some of the chemical constituents of *C. arborea* identified in this study are consistent with those reported previously on this plant species. Among these, phenolic glycosides (**4–6**) can be considered as chemotaxonomic markers of *C. arborea* as they have not been isolated from any other *Careya* species. Further phytochemical research should be conducted on this genus to understand the functions of the different compounds that occur in the various member species.

Abbreviations

COSY: Correlation spectroscopy
DEPT: Distortionless enhancement by polarization transfer
ESIMS: Electrospray ionization mass spectrometry
HMBC: Heteronuclear multiple bond correlation
HPLC: High-performance liquid chromatography
HRESIMS: High-resolution electrospray ionization mass spectrometry
HSQC: Heteronuclear single quantum coherence
MPLC: Medium pressure liquid chromatography

NMR: Nuclear magnetic resonance
TLC: Thin layer chromatography.

Data Availability

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

The present study was supported by a grant from the National Institute of Biological Resources under the Ministry of Environment of the Korea (grant no. 2019-06-203). The authors would like to thank enago (http://www.enago.co.kr) for English language editing.

Supplementary Materials

Scheme S1: The isolation scheme of compounds **1**–**7**. Figure S1: ^{1}H-NMR spectrum of the compound **1**. Figure S2: ^{13}C-NMR spectrum of the compound **1**. Figure S3: DEPT135 spectrum of the compound **1**. Figure S4: ^{1}H-^{1}H COSY spectrum of the compound **1**. Figure S5: HSQC spectrum of the compound **1**. Figure S6: HMBC spectrum of the compound **1**. Figure S7: HRESIMS spectrum of the compound **1**. Figure S8: Specific optical rotation of compound **1**. Figure S9: Specific optical rotation of compound **1** (aglycone). Figure S10: Specific optical rotation of compound **1** (sugar moiety). Figure S11: ^{1}H-NMR spectrum of compound **2**. Figure S12: ^{13}C-NMR spectrum of the compound **2**. Figure S13: ESI-MS spectrum of the compound **2**. Figure S14: ^{1}H-NMR spectrum of the compound **3**. Figure S15: ^{13}C-NMR spectrum of the compound **3**. Figure S16: ESI-MS spectrum of the compound **3**. Figure S17: ^{1}H-NMR spectrum of the compound **4**. Figure S18: ^{13}C-NMR spectrum of the compound **4**. Figure S19: ESI-MS spectrum of the compound **4**. Figure S20: ^{1}H-NMR spectrum of the compound **5**. Figure S21: ^{13}C-NMR spectrum of the compound **5**. Figure S22: ESI-MS spectrum of the compound **5**. Figure S23: ^{1}H-NMR spectrum of the compound **6**. Figure S24: ^{13}C-NMR spectrum of the compound **6**. Figure S25: ESI-MS spectrum of the compound **6**. Figure S26: ^{1}H-NMR spectrum of the compound **7**. Figure S27: ^{13}C-NMR spectrum of the compound **7**. Figure S28: ESI-MS spectrum of the compound **7**. (*Supplementary Materials*)

References

[1] H. A. Khaliq, "Pharmacognostic, physicochemical, phytochemical and pharmacological studies on *Careya arborea* Roxb.; A review," *The Journal of Phytopharmacology*, vol. 5, no. 1, pp. 27–34, 2016.

[2] M. Kaur and R. Yadav, "Pharmacognostic, ethnopharmacological, phytochemical and pharmacological profile of wild guava I.E. *Careya arborea* Roxb," *International Journal of Current Pharmaceutical Research*, vol. 9, no. 3, p. 1, 2017.

[3] L. B. Nadekar, P. M. Deshmukhe, A. G. Sonawane, and M. A. Gujarkar, "*Careya arborea*: a review," *International Journal of Pharmacy and Pharmaceutical Research*, vol. 21, no. 3, p. 574, 2021.

[4] A. S. Navya and S. Anitha, "Review on pharmacognostic and pharmacological activities of *Careya arborea* plant," *Journal of Pharmacognosy and Phytochemistry*, vol. 8, no. 3, p. 4165, 2019.

[5] S. Natesan, S. Badami, S. H. Dongre, and A. Godavarthi, "Antitumor activity and antioxidant status of the methanol extract of *Careya arborea* bark against Dalton's lymphoma ascites-induced ascitic and solid tumor in mice," *Journal of Pharmacological Sciences*, vol. 103, no. 1, pp. 12–23, 2007.

[6] B. Ramadurga, R. K. Jat, and S. Badami, "Pharmacognostic evaluation and antimicrobial activity of root of *Careya arborea*," *Pharmacognosy Journal*, vol. 11, no. 3, pp. 608–612, 2019.

[7] P. Gupta, D. Patil, and A. Patil, "Antibacterial activity of bark and leaf extracts of *Careya arborea* Roxb," *International Journal of Green and Herbal Chemistry*, vol. 8, no. 3, p. 312, 2019.

[8] R. Y. Nanotkar, C. R. Jangde, D. G. Jagtap, and T. K. Sar, "Effect of *Careya arborea* Roxb. on CCl_4 induced liver damage in rats," *Annals of Phytomedicine: International Journal*, vol. 5, no. 2, pp. 113–117, 2016.

[9] R. Begum, M. Sharma, K. K. Pillai, V. Aeri, and M. A. Sheliya, "Inhibitory effect of *Careya arborea* on inflammatory biomarkers in carrageenan-induced inflammation," *Pharmaceutical Biology*, vol. 53, no. 3, pp. 437–445, 2015.

[10] P. C. Gupta and C. V. Rao, "Gastroprotective effect of standardized leaf extract from *Careya arborea* on experimental gastric ulcers in rats," *Pharmaceutical Biology*, vol. 52, no. 8, pp. 1003–1008, 2014.

[11] J. C. Kalita, A. Haque, E. Kalita, R. Sharma, and M. M. Rahman, "Anti-fertility effects of the methanolic root extract of *Careya arborea* Roxb in albino mice," *The Bioscan*, vol. 6, no. 4, p. 701, 2011.

[12] N. Senthilkumar, S. Badami, S. H. Dongre, and S. Bhojraj, "Antioxidant and hepatoprotective activity of the methanol extract of *Careya arborea* bark in Ehrlich ascites carcinoma-bearing mice," *Journal of Natural Medicines*, vol. 62, no. 3, pp. 336–339, 2008.

[13] M. T. Rahman, O. F. Khan, S. Saha, and M. Alimuzzaman, "Antidiarrhoeal activity of the bark extract of *Careya arborea* Roxb," *Fitoterapia*, vol. 74, no. 1-2, pp. 116–118, 2003.

[14] M. Ahmed, M. W. Rahman, M. T. Rahman, and C. F. Hossain, "Analgesic principle from the bark of *Careya arborea*," *Pharmazie*, vol. 57, no. 10, pp. 698–701, 2002.

[15] R. P. Kamble, P. Ghosh, and A. A. Kulkarni, "Identification of α-amylase inhibitory compounds from leaves of *Careya arborea* Roxb. and in silico docking studies," *South African Journal of Botany*, vol. 151, pp. 493–503, 2022.

[16] D. Mandal, N. Panda, and S. Kumar, "A triterpenoid saponin possessing antileishmanial activity from the leaves of *Careya arborea*," *Phytochemistry*, vol. 67, no. 2, pp. 183–190, 2006.

[17] M. C. Das and M. Shashi B, "Triterpenoid sapogenols from the leaves of *Careya arborea*: structure of careyagenolide," *Phytochemistry*, vol. 21, no. 8, pp. 2069–2073, 1982.

[18] B. Talapatra, A. Basak, and S. K. Talapatra, "Terpenoids and related compounds. Part XX. Careaborin, a new triterpene ester from the leaves of *Careya arborea*," *Journal of the Indian Chemical Society*, vol. 58, no. 8, p. 814, 1981.

[19] S. B. Mahato and N. L. Dutta, "Sterols from *Careya arborea*," *Phytochemistry*, vol. 11, no. 6, pp. 2116-2117, 1972.

[20] X. L. Jiang, L. Wang, and E. J. Wang, "Flavonoid glycosides and alkaloids from the embryos of *Nelumbo nucifera* seeds and their antioxidant activity," *Fitoterapia*, vol. 125, pp. 184–190, 2018.

[21] O. S. Ayoolu, O. O. Ogbole, and E. O. Ajaiyeoba, "Flavanols from *Tetrapleura tetraptera* with cytotoxic activities," *Fitoterapia*, vol. 160, Article ID 105206, 2022.

[22] C. L. Zhang, Y. Wang, and Y. F. Liu, "Two new flavonoid glycosides from *Iris tectorum*," *Phytochemistry Letters*, vol. 15, pp. 63–65, 2016.

[23] Y. Liu, Z. Lu, and M. Wang, "Catechins and lignan from the flower buds of Rosa chinensis Jacq," *Phytochemistry Letters*, vol. 38, pp. 46–48, 2020.

[24] N. Li, M. Chen, and H. T. Zhu, "Theaflavoids A-C, new flavan-3-ols with potent α-glucosidase inhibitory activity from Yunnan black tea 'Jin-Ya'," *LWT-Food Science & Technology*, vol. 168, Article ID 113918, 2022.

[25] M. K. Tao, M. Xu, and H. T. Zhu, "New phenylpropanoid-substituted flavan-3-ols from Pu-er ripe tea," *Natural Product Communications*, vol. 9, no. 8, Article ID 1934578X1400900, 1170 pages, 2014.

[26] S. S. Hong, C. W. Choi, Y. H. Choi, J. S. Oh, and A. Coixlachryside, "A new lignan glycoside from the roots of *Coix lachryma-jobi* L. var. *ma-yuen* Stapf," *Phytochemistry Letters*, vol. 17, pp. 152–157, 2016.

[27] K. T. Cheng, F. L. Hsu, and S. H. Chen, "New constituent from *Podocarpus macrophyllus* var. *macrophyllus* shows anti-tyrosinase effect and regulates tyrosinase-related proteins and mRNA in human epidermal melanocytes," *Chemical & Pharmaceutical Bulletin*, vol. 55, no. 5, pp. 757–761, 2007.

[28] Y. Zhang, B. Jayaprakasam, and N. P. Seeram, "Insulin secretion and cyclooxygenase enzyme inhibition by cabernet sauvignon grape skin compounds," *Journal of Agricultural and Food Chemistry*, vol. 52, no. 2, pp. 228–233, 2004.

[29] G. Hussein, N. Nakamura, and M. R. Meselhy, "Phenolics from *Maytenus senegalensis*," *Phytochemistry*, vol. 50, no. 4, pp. 689–694, 1999.

[30] V. D. Nguyen, T. L. Nguyen, and H. T. Tran, "Flavan-3-ols from the barks of *Barringtonia acutangula*," *Biochemical Systematics and Ecology*, vol. 55, pp. 219–221, 2014.

[31] B. Y. Hwang, H. S. Kim, and J. H. Lee, "Antioxidant benzoylated flavan-3-ol glycoside from *Celastrus orbiculatus*," *Journal of Natural Products*, vol. 64, no. 1, pp. 82–84, 2001.

[32] A. L. Davis, Y. Cai, A. P. Davies, and J. R. Lewis, "^{1}H and ^{13}C NMR assignments of some green tea polyphenols," *Magnetic Resonance in Chemistry*, vol. 34, no. 11, pp. 887–890, 1996.

[33] S. Valcic, J. A. Burr, B. N. Timmermann, and D. C. Liebler, "Antioxidant chemistry of green tea catechins. New oxidation products of (−)-epigallocatechin gallate and (−)-epigallocatechin from their reactions with peroxyl radicals," *Chemical Research in Toxicology*, vol. 13, no. 9, pp. 801–810, 2000.

[34] H. Otsuka, M. Takeuchi, and S. Inoshiri, "Phenolic compounds from *Coix lachryma-jobi var.*Ma-yuen," *Phytochemistry*, vol. 28, no. 3, pp. 883–886, 1989.

[35] K. Yoshikawa, S. Sugawara, and S. Arihara, "Phenylpropanoids and other secondary metabolites from fresh fruits of *Picrasma quassioides*," *Phytochemistry*, vol. 40, no. 1, pp. 253–256, 1995.

[36] M. L. Zeraik, E. F. Queiroz, and L. Marcourt, "Antioxidants, quinone reductase inducers and acetylcholinesterase inhibitors from *Spondias tuberosa* fruits," *Journal of Functional Foods*, vol. 21, pp. 396–405, 2016.

[37] F. M. Abdel Bar, D. S. Ibrahim, and S. R. Gedara, "Nematicidal compounds from the leaves of *Schinus terebinthifolius* against root-knot nematode, *Meloidogyne incognita* infecting tomato," *Natural Product Sciences*, vol. 24, no. 4, pp. 272–283, 2018.

[38] M. K. Sakar, F. Petereit, and A. Nahrstedt, "Two phloroglucinol glucosides, flavan gallates and flavonol glycosides from *Sedum sediforme* flowers," *Phytochemistry*, vol. 33, no. 1, pp. 171–174, 1993.