Supporting Information

Polyphenol and Ellagitannin Constituents of Jabuticaba (*Myrciaria cauliflora*) and Chemical Variability at Different Stages of Fruit Development

Luciane Dias Pereira[1], João Marcos Gonçalves Barbosa[1], Antonio Jorge Ribeiro da Silva[2], Pedro Henrique Ferri[1], Suzana Costa Santos[1*]

[1]*Instituto de Química, Universidade Federal de Goiás, 74690-900, Goiânia, Goiás, Brazil*

[2]*Núcleo de Pesquisas de Produtos Naturais, Centro de Ciências da Saúde, Universidade Federal do Rio de Janeiro, 21941-590, Rio de Janeiro, Rio de Janeiro, Brazil*

Corresponding author

*Phone: + 55 62 3521 1008. E-mail: suzana.quimica.ufg@hotmail.com.

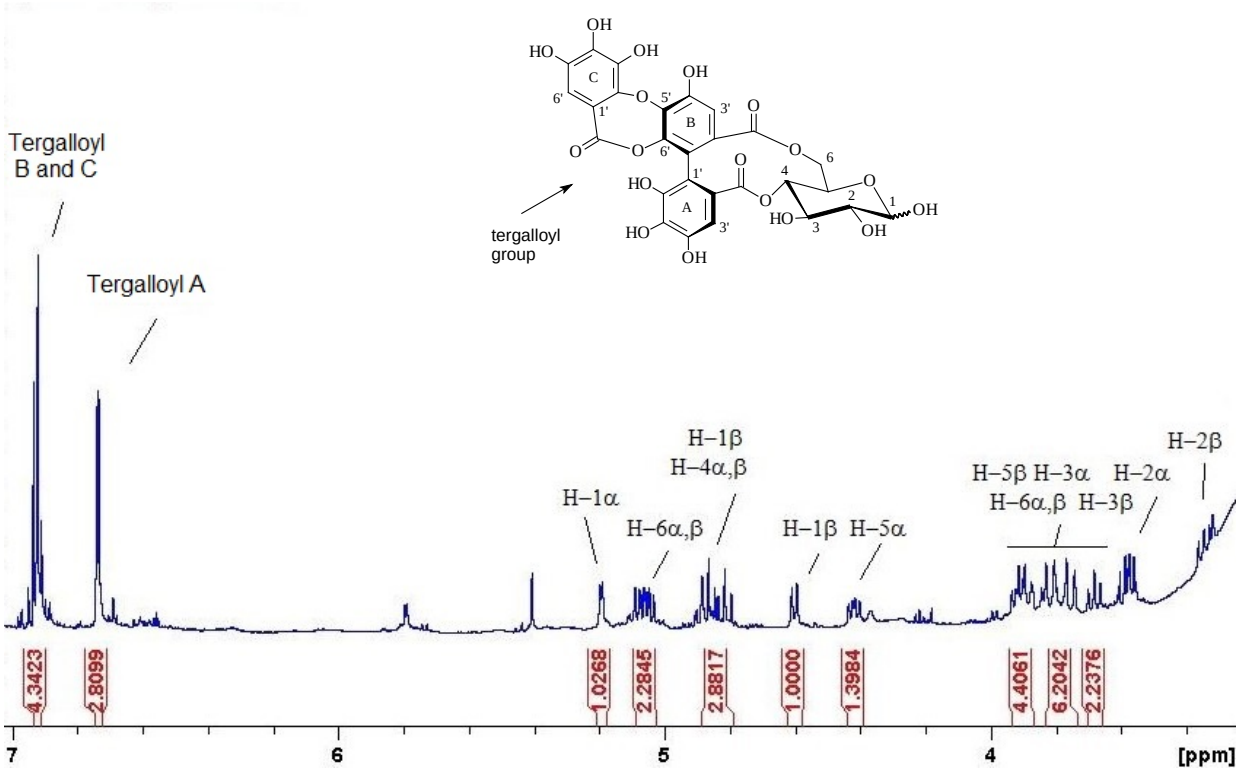


Figure S1. ^{1}H NMR (500 MHz, acetone-d_6) spectrum of cauliflorin (**1**) (From 7.0 to 3.2 ppm).

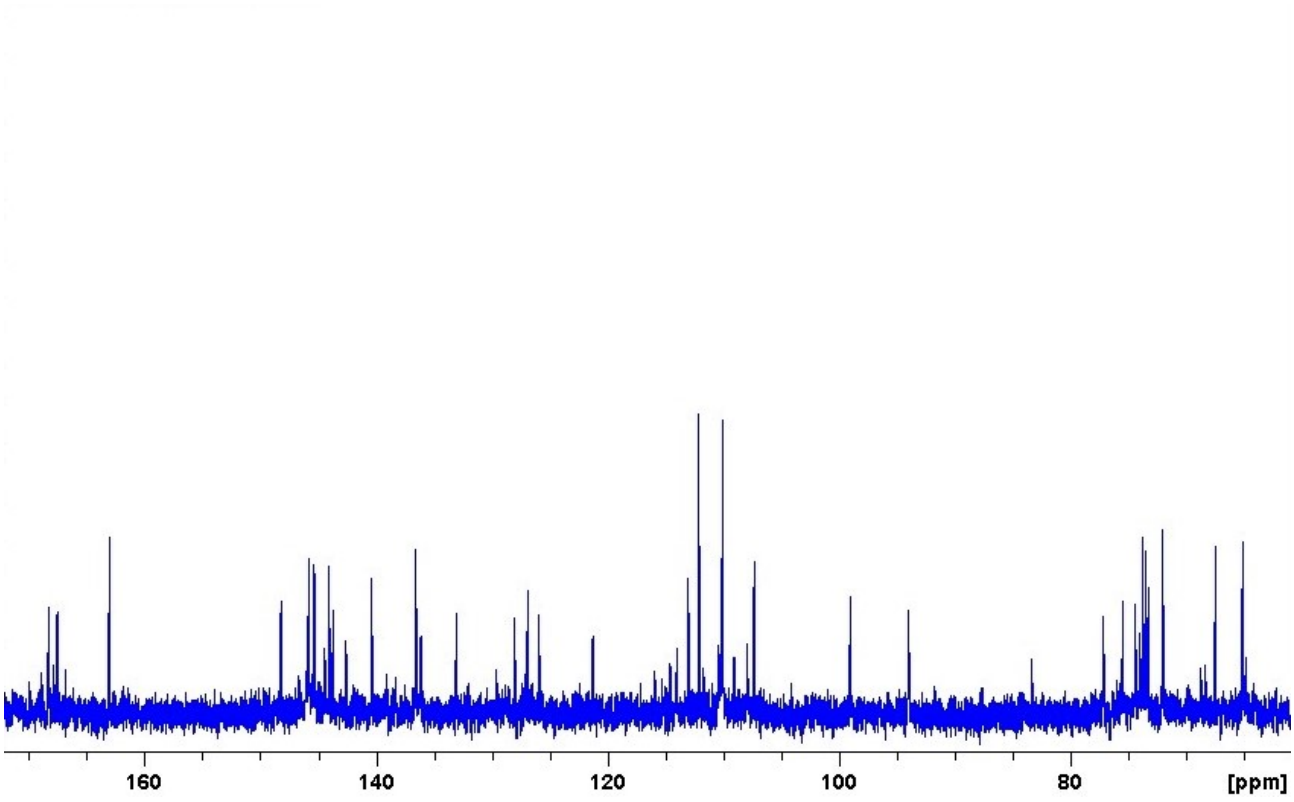


Figure S2. ^{13}C NMR (125 MHz, acetone-d_6) spectrum of cauliflorin (**1**).

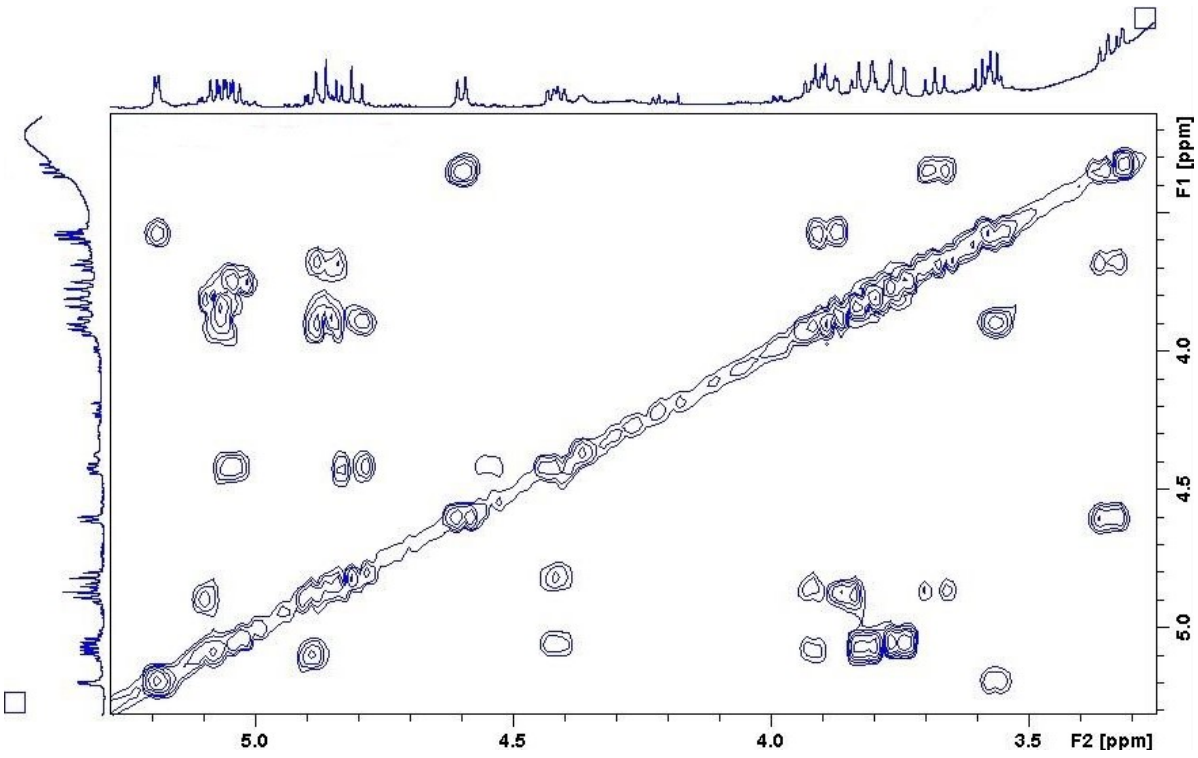


Figure S3. Expansion of the COSY NMR experiment of cauliflorin (**1**).

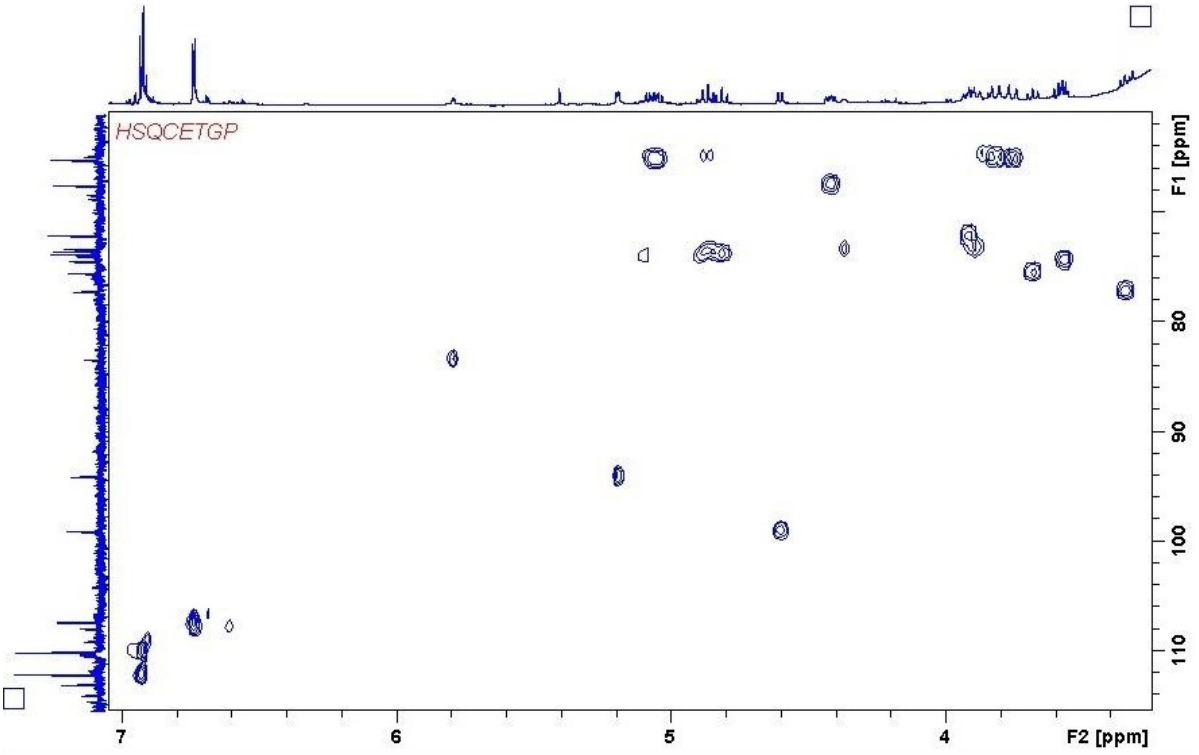


Figure S4. Expansion of the HSQC NMR experiment of cauliflorin (**1**).

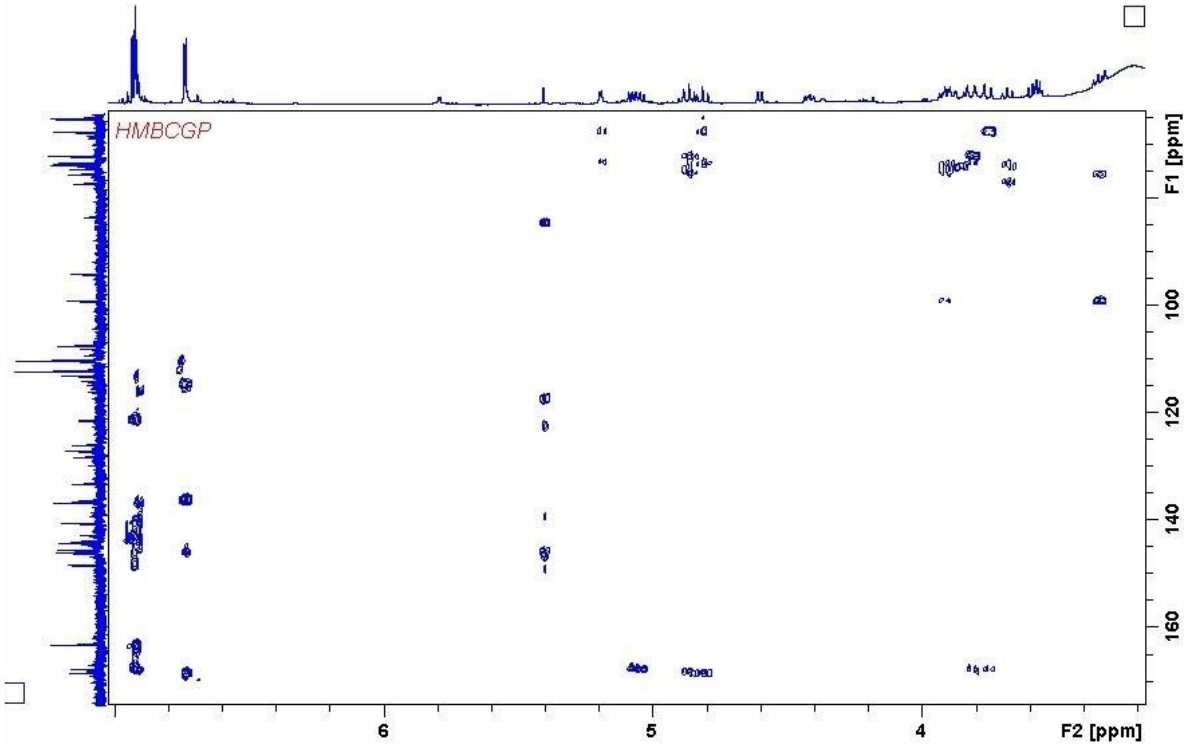


Figure S5. Expansion of the HMBC NMR experiment of cauliflorin (**1**).

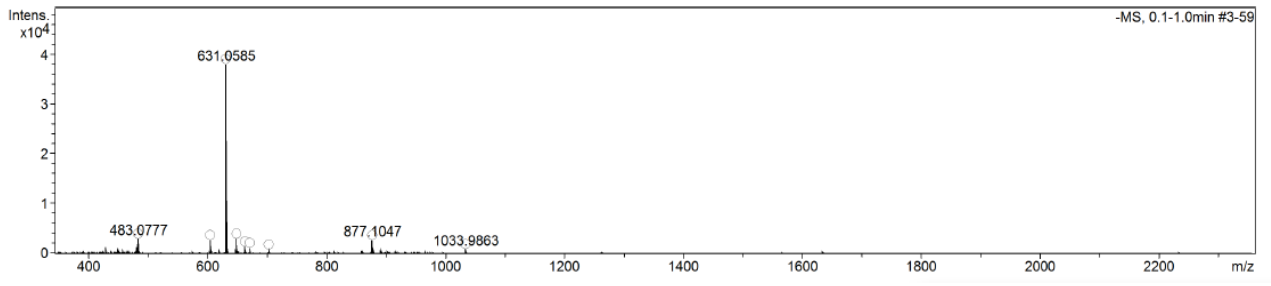
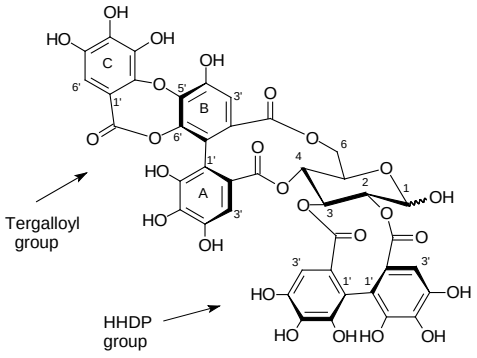


Figure S6. ESI-TOF mass spectrum of cauliflorin (**1**).

Table S1. $^{1}H–^{13}C$ long-range correlation data for cauliflorin (**1**)

position	δ_H (ppm)	$^{1}H–^{13}C$ long-range correlations
α anomer		
1	5.19	C-3α, C-5α
2	3.57	C-3α
3	3.89	C-2α, C-4α
4	4.81	C-3α, C-5α, C-6α, C-7'$_A$
5	4.42	C-1α, C-4α
6	5.05, 3.76	C-4α, C-5α, C-7'$_B$
β anomer		
1	4.60	C-5β
2	3.34	C-1β, C-3β
3	3.68	C-2β, C-4β
4	4.86	C-3β, C-5β, C-6β, C-7'$_A$
5	3.91	C-1β, C-4β
6	5.07, 3.82	C-5β, C-7'$_B$
tergalloyl A ring		
3'	6.73, 6.74	C-2'$_A$, C-4'$_A$, C-5'$_A$, C-7'$_A$
tergalloyl B ring		
3'	6.92, 6.93	C-2'$_B$, C-4'$_B$, C-5'$_B$, C-7'$_B$
tergalloyl C ring		
6'	6.91	C-1'$_C$, C-2'$_C$, C-4'$_C$, C-5'$_C$, C-7'$_C$

Spectral data of the known isolated ellagitannins



alnusiin (**2**)

Alnusiin (**2**) is a mixture of α-anomer and β-anomer. It is an off-white amorphous powder, ESI-TOF MS *m/z* 933.0642 $[M - H]^-$ (calcd for $C_{41}H_{25}O_{26}$, 933.0640). ^{1}H NMR (acetone-d_6, 500 MHz), δ: 3.87 (1H, dd, *J* = 13, 2 Hz, H-6α), 3.94 (1H, dd, *J* = 13, 1 Hz, H-6β), 4.26 (1H, ddd, *J* = 10, 6, 1 Hz, H-5β), 4.63 (1H, ddd, *J* = 9.8, 7, 2 Hz, H-5α), 4.88 (1H, dd, *J* = 9, 8 Hz, H-2β), 5.07 (1H, d, *J* = 8 Hz, H-1β), 5.09 (1H, dd, *J* = 9.8, 3.8 Hz, H-2α), 5.16 (1H, t, *J* = 9.8 Hz, H-4α), 5.16 (1H, t, *J* = 10 Hz, H-4β), 5.21 (1H, dd, *J* = 13, 7 Hz, H-6α), 5.24 (1H, dd, *J* = 13, 6 Hz, H-6β), 5.27 (1H, dd, *J* = 10, 9 Hz, H-3β), 5.47 (1H, d, *J* = 3.8, H-1α), 5.50 (1H, , t, *J* = 9.8 Hz, H-3α), 6.41 and 6.64 (2H, s, HHDP H-3'β), 6.40 and 6.65 (2H, s, HHDP H-3'α), 6.60 and 6.64 (2H, s, tergalloyl A H-3'α/β), 6.95 and 6.96 (2H, s, tergalloyl C H-6'α/β), 6.99 and 7.01 (2H, s, tergalloyl B H-3'α/β). ^{13}C NMR (acetone-d_6, 125 MHz), δ: 64.5 (2C, C-6α/β), 67.4 (C-5α), 69.6 (C-4β), 70.0 (C-4α), 72.4 (C-5β), 75.9 (C-2α), 75.9 (C-3α), 77.6 (C-3β), 78.5 (C-2β), 91.8 (C-1α), 95.5 (C-1β), 107.5 and 107.6 (2C, HHDP C-3'α/β), 114.4 and 114.8 (2C, HHDP C-1'α/β), 136.1 and 136.3 (2C, HHDP C-5'α/β), 145.1 (2C, HHDP C-4'α/β), 168.7 and 169.7 (2C, HHDP C-7'α/β), 107.6 (2C, tergalloyl A C-3'α/β), 113.9 and 114.7 (2C, tergalloyl A C-1'α/β), 136.3 (2C, tergalloyl A C-5'α/β), 145.1 and 145.7 (2C, tergalloyl A C-4'α/β), 168.3 and 168.7 (2C, tergalloyl A C-7'α/β), 113.0 (2C, tergalloyl B C-3'α/β), 120.8 (2C, tergalloyl B C-1'α/β), 143.0 (2C, tergalloyl B C-5'α/β), 148.4 (2C, tergalloyl B C-4'α/β), 167.5 (2C, tergalloyl B C-7'α/β), 110.0 (2C, tergalloyl C C-6'α/β), 112.3 (2C, tergalloyl C C-1'α/β), 140.5 (2C, tergalloyl C C-2'α/β), 143.6 (2C, tergalloyl C C-4'α/β), 144.0 (2C, tergalloyl C C-5'α/β), 163.4 (2C, tergalloyl C C-7'α/β).

Pedunculagin (**3**) is a mixture of α-anomer and β-anomer. It is a light brown amorphous powder, ESI-TOF MS *m/z* 783.0671 $[M - H]^-$ (calcd for $C_{34}H_{23}O_{22}$, 783.0686). ^{1}H NMR (acetone-d_6, 500 MHz), δ: 3.79 (1H, dd, *J* = 13, 2 Hz, H-6α), 3.85 (1H, dd, *J* = 13, 1 Hz, H-6β), 4.22 (1H, dd, *J* = 10, 6 Hz, H-5β), 4.61 (1H, dd, *J* = 10, 7 Hz, H-5α), 4.86 (1H, dd, *J* = 9, 8 Hz, H-2β), 5.06 (1H, d, *J* = 8 Hz, H-1β), 5.07 (1H, dd, *J* = 10, 4 Hz, H-2α), 5.08 (1H, t, *J* = 10 Hz, H-4α), 5.08 (1H, t, *J* = 10

Hz, H-4β), 5.24 (1H, dd, *J* = 10, 9 Hz, H-3β), 5.26 (1H, dd, *J* = 13, 7 Hz, H-6α), 5.30 (1H, dd, *J* = 13, 6 Hz, H-6β), 5.46 (1H, d, *J* = 4, H-1α), 5.47 (1H, , t, *J* = 10 Hz, H-3α), 6.33 and 6.52 (2H, s, HHDP-H-3''β), 6.34 and 6.57 (2H, s, HHDP H-3''α), 6.60 and 6.67 (2H, s, HHDP H-3''β), 6.61 and 6.66 (2H, s, HHDP H-3''α). ^{13}C NMR (acetone-d_6, 125 MHz), δ: 63.6 (2C, C-6α/β), 67.5 (C-5α), 69.7 (C-4β), 69.9 (C-4α), 72.6 (C-5β), 75.6 (C-2α), 75.9 (C-3α), 77.7 (C-3β), 78.5 (C-2β), 91.8 (C-1α), 95.4 (C-1β), 107.3, 107.4, 107.6, 107.7, 107.8, 107.9, 108.4 and 108.5 (8C, HHDP C-3''α/β).

Strictinin (**4**) is an off-white amorphous powder, ESI-TOF MS *m/z* 633.0732 $[M - H]^-$ (calcd for $C_{27}H_{21}O_{18}$, 633.0733). ^{1}H NMR (acetone-d_6, 500 MHz), δ: 3.72 (1H, dd, *J* = 9, 8 Hz, H-2), 3.74 (1H, t, *J* = 10 Hz, H-3), 3.85 (1H, d, *J* = 13 Hz, H-6), 4.18 (1H, m, H-5), 4.94 (1H, t, *J* = 10 Hz, H-4), 5.26 (1H, dd, *J* = 13, 6 Hz, H-6), 5.72 (1H, d, *J* = 8 Hz, H-1), 6.52 and 6.68 (2H, s, HHDP H-3'/3'), 7.21 (2H, s, H-2'/6'). ^{13}C NMR (acetone-d_6, 125 MHz), δ: 63.6 (C-6), 72.0 (C-4), 72.0 (C-5), 73.3 (C-2), 73.3 (C-3), 94.8 (C-1), 106.0 and 107.2 (2C, HHDP C-3'/3'), 109.4 (2C, C-2'/6').

Casuarictin (**5**) is a light brown amorphous powder, ^{1}H NMR (acetone-d_6, 500 MHz), δ: 3.90 (1H, d, *J* = 13 Hz, H-6), 4.50 (1H, dd, *J* = 10, 6.4 Hz, H-5), 5.15 (1H, t, *J* = 10 Hz, H-4), 5.20 (1H, t, *J* = 9 Hz, H-2), 5.34 (1H, dd, *J* = 13, 6.4 Hz, H-6), 5.44 (1H, dd, *J* = 10, 9 Hz, H-3), 6.22 (1H, d, *J* = 9 Hz, H-1), 6.41 and 6.60 (2H, s, HHDP H-3'/3'), 6.49 and 6.69 (2H, s, HHDP H-3'/3'), 7.20 (2H, s, H-2'/6'). ^{13}C NMR (acetone-d_6, 125 MHz), δ: 62.7 (C-6), 68.6 (C-4), 72.5 (C-5), 75.1 (C-2), 76.4 (C-3), 91.4 (C-1), 106.1, 106.2, 106.6 and 106.8 (4C, HHDP C-3'/3'), 109.3 (2C, C-2'/6').

Castalagin (**7**) is a white amorphous powder, ESI-TOF MS *m/z* 933.0631 $[M - H]^-$ (calcd for $C_{41}H_{25}O_{26}$, 933.0640). ^{1}H NMR (acetone-d_6, 500 MHz), δ: 4.01 (1H, d, *J* = 13 Hz, H-6), 5.03 (1H, dd, *J* = 7.0, 1.4 Hz, H-3), 5.04 (1H, dd, *J* = 4.7, 1.4 Hz, H-2), 5.10 (1H, dd, *J* = 13, 2.6 Hz, H-6), 5.24 (1H, dd, *J* = 7.4, 7.0 Hz, H-4), 5.62 (1H, ddd, 7.4, 2.6, 1.0 Hz, H-5), 5.74 (1H, d, *J* = 4.7 Hz, H-1), 6.64 (1H, s, H-3', HHDP B), 6.79 and 6.91 (2H, s, H-3' HHDP A and flavogalloyl C). ^{13}C NMR (acetone-d_6, 125 MHz), δ: 64.5 (C-6), 65.6 (C-3), 66.7 (C-1), 68.6 (C-4), 70.4 (C-5), 73.3 (C-2).

Vescalagin (**8**) is a white amorphous powder, ESI-TOF MS *m/z* 933.0562 $[M - H]^-$ (calcd for $C_{41}H_{25}O_{26}$, 933.0642). ^{1}H NMR (acetone-d_6, 500 MHz), δ: 4.01 (1H, d, *J* = 13 Hz, H-6), 4.58 (1H, dd, *J* = 7.0, 1.5 Hz, H-3), 4.92 (1H, d, *J* = 2.3 Hz, H-1), 5.10 (1H, dd, *J* = 13, 2.6 Hz, H-6), 5.22 (1H, dd, *J* = 7.4, 7.0 Hz, H-4), 5.24 (1H, dd, *J* = 2.3, 1.5 Hz, H-2), 5.65 (1H, ddd, 7.4, 2.6, 1.0 Hz,

H-5), 6.63 (1H, s, H-3', HHDP B), 6.79 and 6.79 (2H, s, H-3' HHDP A and flavogalloyl C). ^{13}C NMR (acetone-d_6, 125 MHz), δ: 64.5 (C-6), 64.6 (C-1), 67.7 (C-3), 68.6 (C-4), 70.4 (C-5), 77.0 (C-2).

Table S2. ^{1}H and ^{13}C NMR spectroscopic data for compounds in jabuticaba's fruit extracts[a]

Compounds	δ_H	Multiplicity	*J*	Assignments	δ_C
sucrose	5.42	d	3.9	C1H α-glucose	92.1
α-glucose	5.24	d	3.8	C1H	95.2
β-glucose	3.25	dd	9.2, 8.1	C2H	77.2
α-fructose	4.12	m	-	C3H	84.7
β-fructose	4.12	m	-	C3H, C4H	77.6
citric acid	3.05	d	15.8	*α′, γ*'CH	46.5
malic acid	2.94	dd	16.6, 4.6	*β*'CH	41.5

[a] Recorded in HCl/H_2O/D_2O solution (pH 1.0) at 500 MHz for ^{1}H and at 125 MHz for ^{13}C NMR spectra. Chemical shifts and coupling constants are given in ppm and Hz, respectively. d, doublet; dd, double doublet; m, multiplet.

Table S3. Compounds, retention times and UV data of HPLC analysis

Compounds	Time (min)	UV max (λ/nm)[a]	Analyze (λ/nm)
Gallic acid	4.8	213, 271	216
Vescalagin	7.0	200, 221	216
Pedunculagin	8.2 and 13.7	196, 219	216
Castalagin	10.2	200, 221	216
Cauliflorin	18.2 and 21.5	214, 278	216
Delphinidin-3-*O*-glucose	27.4	274, 519	370
Cyanidin-3-*O*-glucose	31.4	279, 516	370
Ellagic acid	48.9	254, 368	370

[a]Mobile phase: acetonitrile and 0.01 M H_3PO_4: 0.01 M KH_2PO_4

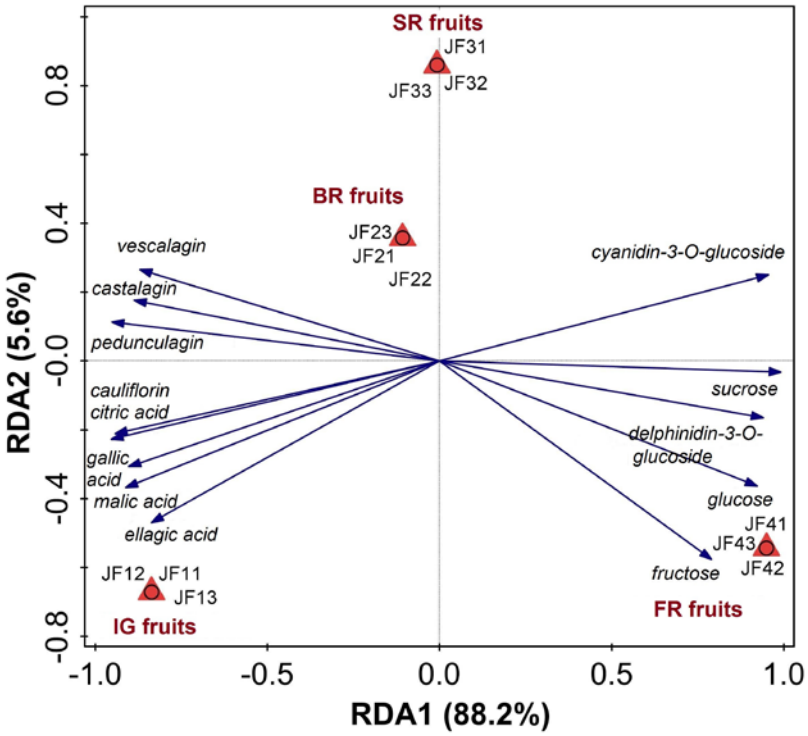


Figure S7. RDA ordination of the first two axes showing the distribution of jabuticaba sampling maturity stages (IG, BR, SR, FR). Chemical constituents of fruits are represented by long arrows from the origin. Red triangles represent cluster centroids. Values in brackets refer to the explained variance on each canonical axis.

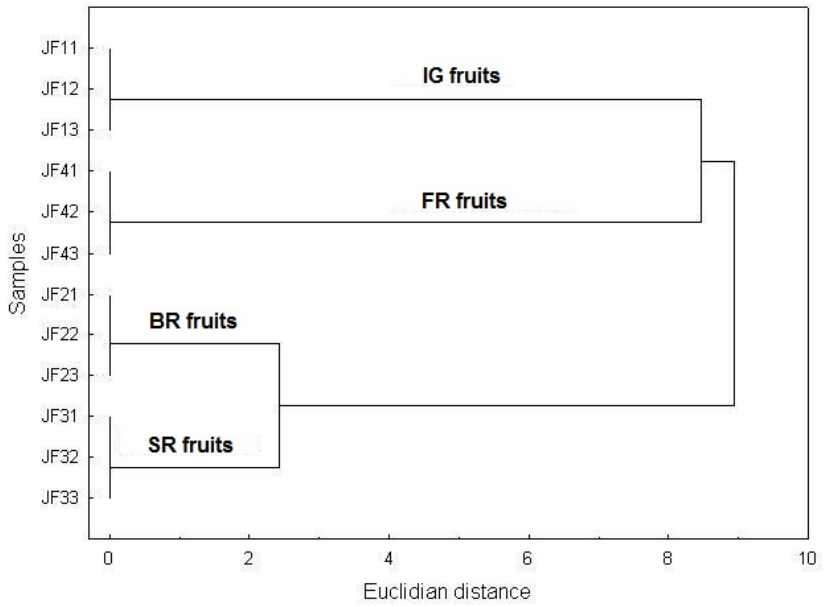


Figure S8. Dendrogram representing chemical composition similarity relationships among jabuticaba's fruit samples and to whose cluster it belongs: immature green (IG), breaker turning purple (BR), semiripe (SR), full-ripe (FR).

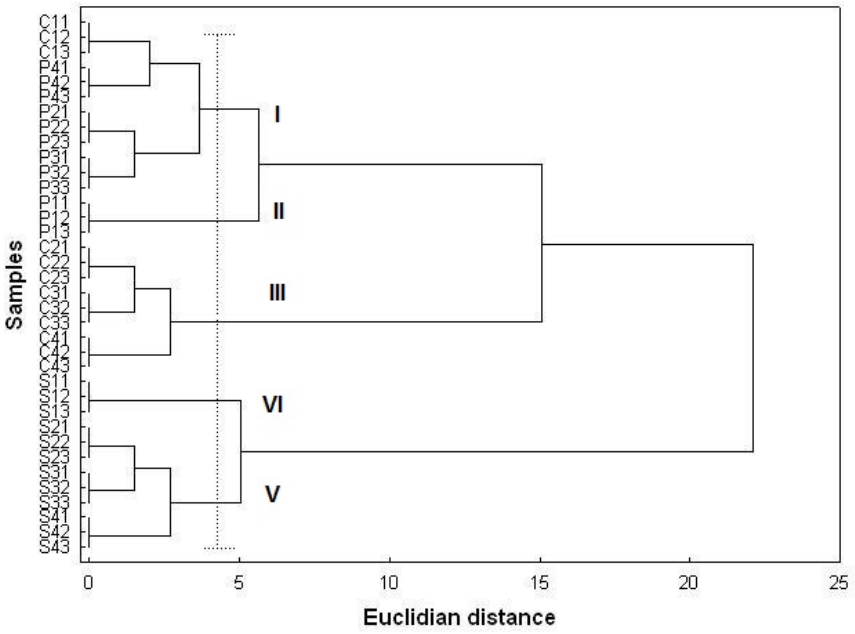


Figure S9. Dendrogram representing chemical composition similarity relationships among jabuticaba's fruit parts samples (C, peel; P, pulp; S, seed) to whose cluster it belongs: I with IG peel, BR, SR and FR pulp; II with IG pulp; III with BR, SR and FR pulp; VI with IG seed; V with BR, SR and FR seed.

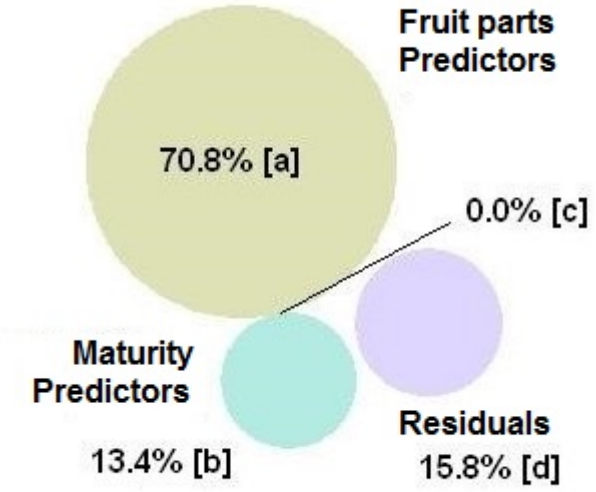


Figure S10. Venn diagram illustrating percentages of variation in jabuticaba's fruit parts chemical composition that are attributable to two sets of explanatory variables: fruit parts and degree of maturity predictors. Letters in brackets correspond to variation explained by each component of variation according to Table 4.