

Supplementary Information for

Disentangling the Chemistry of Australian Plant Exudates from a Unique Historical Collection

Rafaella Georgiou, Rachel S. Popelka-Filcoff, Dimosthenis Sokaras, Victoria Beltran, Ilaria Bonaduce, Jordan Spangler, Serge Cohen, Roy Lehmann, Sylvain Bernard, Jean-Pascal Rueff, Uwe Bergmann, Loïc Bertrand

Rachel S. Popelka-Filcoff Uwe Bergmann Loïc Bertrand E-mail: rachel.popelkafilcoff@unimelb.edu.au ubergmann@wisc.edu loic.bertrand@ens-paris-saclay.fr

This PDF file includes:

Supplementary text Figs. S1 to S14 Tables S1 to S7 SI References

Supporting Information Text

1. Materials and Methods

A. Materials. A sample-set of native Australian plant genera (*Eucalyptus, Callitris, Xanthorrhoea, Acacia*) was analyzed. The samples were selected from South Australian Museum (SAM) collections and a historical biological collection gathered from diverse locations across Australia and held for years in the South Australia state forensic laboratory (Forensic Sciences South Australia–FSSA), now located at Flinders University. Unfortunately the collectors of this exceptional comprehensive collection are unknown and the original data notebooks no longer exist; however, each amber jar is clearly labeled with the genus and species, collection location and date. Many of these samples in the historical biological collection are over a century old and have been stored in amber glass bottles under ambient conditions and offer an unprecedented study set. By keeping the storage conditions consistent within the collection, it minimizes the effects of temperature and humidity as a variable across the collection.

Table S1. Australian native plant exudates. The raw materials from the historical collection are approximately 100+ years old. Samples from the South Australian Museum (SAM) (denoted by SAM number) are raw plant materials also stored for decades with the exact collection dates unknown.

Sample	Species	Genus	Clade	Location and date
1	X. arborea	Xanthorrhoea	Angiosperms	Helensburgh, Australia, 29 May 1905
2	X. semiplana ssp. tateana	Xanthorrhoea	Angiosperms	Munston, Kangaroo Island, South Australia [SAM No. A66781]
3	X. semiplana	Xanthorrhoea	Angiosperms	West of Clarendon, South Australia [SAM No A66782]
4	C. calcarata	Callitris	Gymnosperms	Quiedong, near Bombala, New South Wales Australia, 03 March 1887
5	C. glauca	Callitris	Gymnosperms	Deniliquin, New South Wales, Australia
6	C. preissii ssp. verrucosa	Callitris	Gymnosperms	Shuttleton, New South Wales, Australia
7	White Mallee	Eucalyptus	Angiosperms	Not recorded, 06 May 1911
8	E. largiflorens	Eucalyptus	Angiosperms	Gunbar, New South Wales, Australia, Decem ber 1900
9	Not recorded	Acacia	Angiosperms	Coopers Creek [SAM No. A2182]
10	A. bakeri	Acacia	Angiosperms	Tubulgum, Australia, February 1899

B. XRS spectroscopy. X-ray Raman measurements (carbon K-edge and oxygen K-edge) were performed using the high-resolution X-ray Raman spectrometer operated at beamline 6–2 at the Stanford Synchrotron Radiation Laboratory (1). The instrument consists of 40 crystal Si(440) analyzers arranged on overlapping Rowland circles of 1000 mm and at a fixed Bragg angle of 88°, resulting in a fixed detection energy at 6462 eV and an overall energy resolution of 0.29 eV. Spectra are recorded using an inverse scanning approach, wherein the detection energy is held constant at 6462 eV while the incident energy is scanned through a Si(311) double-crystal monochromator. The average momentum transfer is $q \approx 1.23$ a.u., which ensures that dipole transitions dominate the spectra and therefore the X-ray Raman spectra are formally equivalent to conventional soft X-ray absorption spectra.

For each sample, the mean carbon K-edge spectrum shown in Fig. S1 A was obtained from a collection of multiple consecutive spectra on numerous sample regions to ensure sufficient signal-to-noise ratio and high data quality. No visual damage was observed on the samples under study. The experimental conditions were optimized to minimize radiation damage. We measured the samples at a grazing geometry ($\approx 6^{\circ}-7^{\circ}$) and therefore the projected beam footprint on the sample was quite long (>5 mm) in the horizontal direction. Vertically the beam FWHM was $\approx 150 \,\mu$ m. The incident flux was 2–3 × 10¹² ph/sec. A dwell time of 1.5 s per energy step and an energy step size of 2.5 eV over the 254–279 eV range, 0.5 eV over the 279–284 eV range, 0.1 eV over the 284–298 eV range, 0.25 eV over the 298–328 eV range, 1 eV over the 328–333 eV range, 2 eV over the 333-343 eV range and 2.5 eV up to 363 eV were used for each scan. Mean spectra were energy-calibrated using the average center of the elastic scattering peaks collected before and between XRS measurements, in order to take into account the energy drift of the monochromator with time. Using the empirical formula y = a/(x - b) (2), we performed a non-linear fit to the pre-edge region and subtracted the background. The background-corrected spectra were then normalized to the post-edge region with no significant oscillation (Fig. S1).

XRS oxygen K-edge experiments were also performed at GALAXIES beamline (Synchrotron SOLEIL, France) with an energy resolution of 1.2 eV (3). The spectrometer is equipped with 4 spherically bent Si(444) analyzer crystals (4). For the XRS spectrum acquisition, we scanned the beamline energy along the absorption edge of interest, at a fixed analyzer energy. Experiments were performed at a Si(444) analyzer energy of $\approx 8.0 \text{ keV}$ operated at a Bragg angle of 86°. The data was collected at backscattered geometry (scattering angle of $2\theta \approx 130^{\circ}$) to maximize the XRS signal. The momentum transfer is $q \approx 7.5 \text{ Å}^{-1}$. The beamline is equipped with a cryogenically cooled Si(111) double-crystal monochromator. The monochromatized X-ray beam is focused by a toroidal mirror to a spot size (V × H) of approximately $30 \times 80 \text{ µm}^2$.

For each sample, the mean oxygen K-edge spectrum shown in Fig. S1 B was obtained from a collection of 8–10 consecutive scans. The incident flux at 8 keV was 5×10^{13} ph/sec. A dwell time of 0.5 s per energy step and an energy step size of 0.25 eV

2 of 24 Rafaella Georgiou, Rachel S. Popelka-Filcoff, Dimosthenis Sokaras, Victoria Beltran, Ilaria Bonaduce, Jordan Spangler, Serge Cohen, Roy Lehmann, Sylvain Bernard, Jean-Pascal Rueff, Uwe Bergmann, Loïc Bertrand

A. Carbon K-edge

B. Oxygen K-edge

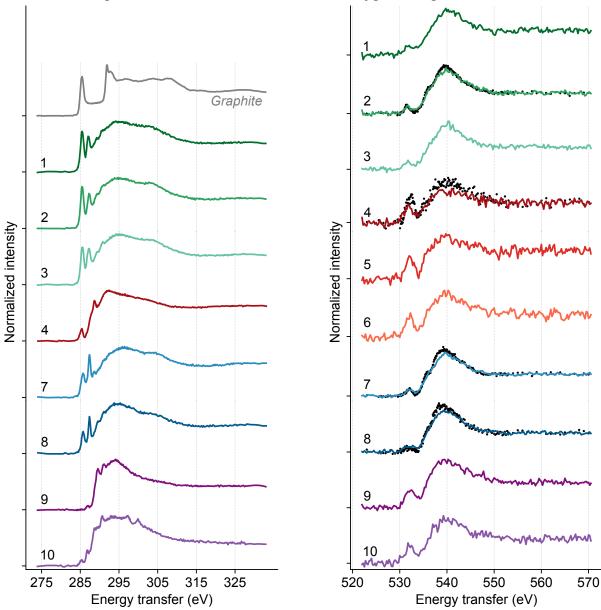


Fig. S1. XRS data of native Australian plant exudates. A. Carbon K-edge (energy resolution: 0.3 eV, SSRL beamline 6-2b). B. Oxygen K-edge (energy resolution: 1.2 eV; SOLEIL beamline GALAXIES). In black color (spectra 2,4,7,8) are higher resolution data for comparison (0.3 eV, SSRL). 1, X. arborea; 2, X. semiplana ssp. tateana; 3, X. semiplana; 4, C. calcarata; 5, C. glauca; 6, C. preissii ssp. verrucosa; 7, White Mallee; 8, E. largiflorens; 9, Acacia sp.; 10, A. bakeri.

over the 500–600 eV range were used for each scan. For the sake of comparison, oxygen K-edge data was collected at beamline 6–2 (SSRL) at low momentum transfer ($q \approx 0.65 \text{ Å}^{-1}$) for four samples (X. semiplana ssp. tateana, C. calcarata, White Mallee, E. largiflorens). Comparison of the oxygen K-edge XRS spectra collected at Galaxies (high-q) and SSRL (low-q) exhibits weak spectral differences, which are possibly related to differences in background subtraction, spectrometer resolution and effects from radiation damage induced during the experiments at Galaxies (Fig. S1 B).

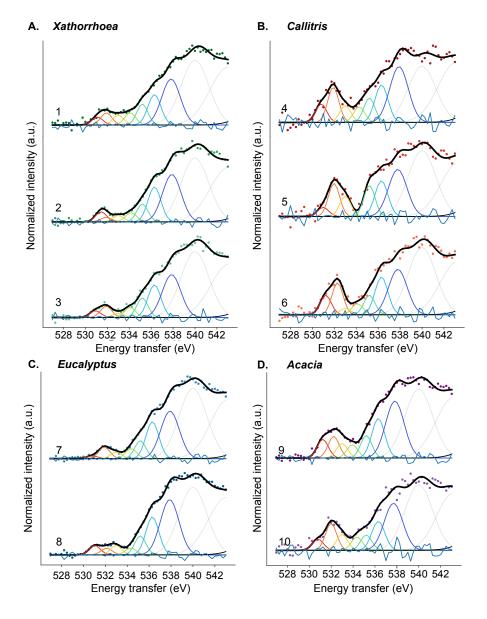


Fig. S2. XRS Oxygen K-edge spectra decomposition. A. 1, X. arborea; 2, X. semiplana ssp. tateana; 3, X. semiplana; B. 4, C. calcarata; 5, C. glauca; 6, C. preissii ssp. verrucosa; C. 7, White Mallee; 8, E. largiflorens; D. 9, Acacia sp., 10, A. bakeri. The centers of Gaussians correspond to core electron transitions at specific functional groups as reported in the literature of soft X-rays absorption and/or EELS studies (see Table S3).

B.1. XRS Data Reduction. The spectral decomposition of the XRS signal in the region of interest 283.0–292.5 eV and 530.8–538.5 eV was performed using a linear combination of Gaussians. The parameters of the model to be fitted are the intensity, the full width half maximum (FWHM) and the center of each Gaussian. The FWHM has a lower (FWHM of the elastic line) and an upper bound (selected in such a way to minimize the post-fit residual). The center of each Gaussian is allowed to vary slightly during the fit in a constrained energy transfer region that corresponds to electronic transitions in specific oxygen-containing functional groups expected for these families of compounds. All the parameters are summarized in Tables S2 and S3. The shift of the center and the width of the fitted Gaussian provides information on the overlap of functional groups of which their transitions are expected to be in the same energy transfer region. Post-fit residuals were calculated to check for possible missing features as in Gueriau et al. (5). The intensity, FWHM and center of each Gaussian are fitted using a nonlinear least squares regression implemented in Python (Tables S4 and S5). The followed approach is semiquantitative, and the area under each Gaussian should only be used to compare relative intensities between samples and not absolute concentrations of each species.

Table S2. XRS carbon	K-edge spectral	decomposition
----------------------	-----------------	---------------

	Center (eV)	FWHM (eV)	Assignment	Transitions
-Gaussian 1	283.0-284.5	0.3–0.9	Aromatic carbonyl (C=O) and/or quinone (6)	1s–π*
– Gaussian 2	285.3–285.6	0.3–0.9	Aromatic and/or olefinic carbon (7)	1s–π*
– Gaussian 3	286.0-287.0	0.3–0.9	Ketones (R ₂ C=O) / nitriles (C=N) (7) / ethers (R-O-R') (8)	1s–π*
– Gaussian 4	287.0–288.0	0.3–0.9	Phenols (Ar-OH) (287.0-287.3 eV) (9) / aliphatics (C _x H _y) (287.3-288.0 eV) (7)	1s–π* 1s–3p/σ*
– Gaussian 5	288.3–288.8	0.3–0.9	Amides (R-C(=O)-N(R',R")) / carboxyl(-C(=O)OH) / acetal carbon (R ₂ C(OR') ₂ , R' \neq H) (7)	1s–π*
– Gaussian 6	289.0–289.6	0.3–0.9	Alcohols (R-OH) (10)	1s–3p/σ*
– Gaussian 7	290.0–290.9	0.3–1.2	Additional 1s– π^* transitions of carboxylic or peroxide groups / alkyl carbon (10, 11)	1s–π* 1s–4p

Table S3. XRS oxygen K-edge spectral decomposition

	Center (eV)	FWHM (eV)	Assignment	Transitions
–Gaussian 1	530.8–531.6	1.2	$C{=}O\frac{(\text{Ketones }(\text{R}_{2}\text{C}{=}\text{O})~(531.3~\text{eV})~(12)~/}{\text{Amides }(\text{R}{-}\text{C}({=}\text{O}){-}\text{N}(\text{R}',\text{R}''))~(531.5~\text{eV}))~(13)}$	1s–π*
– Gaussian 2	531.9–532.3	1.2	C=O (Carboxyls (-C(=O)OH)) (14)	1s-π*
– Gaussian 3	533.0–533.5	1.2	O-O (Peroxides (ROOR)) (15)	1s–σ*
– Gaussian 4	533.9–534.7	1.2	O-H (Alcohols (R-OH)) (16)	1s–σ*
– Gaussian 5	535.2–535.7	1.2	O-H (Phenols (Ar-OH) (17) / carboxyls (-(C=O)OH)) (14)	1s-π*
– Gaussian 6	536.3–537.0	1.2–1.4	Excitations to Rydberg MO's (Peroxides (ROOR) (15))	1s–3s
– Gaussian 7	537.0–538.5	1.2–2	C-O ^{(Peroxides (ROOR) (15) /} Alcohols(R-OH) (16)	1s–σ*

	-	– Gaussian 1	'n	- 0	–Gaussian 2	2	<mark>0 -</mark>	–Gaussian 3	ω	- 0	– Gaussian 4	14	- 0	– Gaussian 5	сл	<mark>0 -</mark>	–Gaussian 6	6	- G	– Gaussian 7	
	μ _{Fit}	σ _{Fit}	l _{Fit}	HFit	σ _{Fit}	l _{Fit}	HFit	σ _{Fit}	l _{Fit}	μ _{Fit}	σ _{Fit}	I _{Fit}	HFit	σ _{Fit}	l _{Fit}	HFit	σ _{Fit}	l _{Fit}	HFit	σ _{Fit}	
X. arborea	284.5	0.38	0.18	285.5	0.38	1.46	286.7	0.38	0.62	287.3	0.38	0.96	288.3	0.38	0.85	289.2	0.38	0.97	290.2	0.51	1.20
X. semiplana ssp. tateana	284.5	0.38	0.14	285.5	0.38	1.44	286.6	0.38	0.43	287.4	0.38	1.12	288.4	0.38	0.76	289.2	0.38	0.95	290.2	0.51	
X. semiplana	284.5	0.38	0.13	285.5	0.38	1.32	286.7	0.38	0.68	287.4	0.38	0.99	288.4	0.38	0.72	289.2	0.38	0.89	290.3	0.51	
C. calcarata	284.5	0.36	0.06	285.4	0.38	0.34	286.8	0.38	0.25	287.7	0.38	0.68	288.6	0.38	1.05	289.5	0.38	0.70	290.4	0.51	
White Mallee	284.5	0.38	0.08	285.6	0.38	0.65	286.3	0.38	0.39	287.4	0.37	1.20	288.5	0.38	0.72	289.4	0.38	0.58	290.3	0.51	
E. largiflorens	284.5	0.31	0.07	285.6	0.38	0.66	286.3	0.38	0.35	287.4	0.35	1.24	288.5	0.38	0.65	289.3	0.38	0.87	290.2	0.51	
Acacia sp.	284.4	0.14	0.00	285.4	0.38	0.01	286.9	0.38	0.16	287.9	0.35	0.13	288.8	0.38	0.79	289.5	0.38	1.36	290.5	0.51	
A. bakeri	284.5	0.31	0.04	285.4	0.38	0.28	286.8	0.38	0.69	287.8	0.38	0.73	288.6	0.38	1.52	289.5	0.38	1.42	290.5	0.51	

Та
PIe
ő
4
Ä
õ
Ca
5
ĭ
Ž
ed
ge
spec
ē
tra
ē
eco
Ĕ
ğ
sit
itio
2
S
ň
ę
Ē
ŧ
, v
an
da
đ
đ
Š.
ati
ñ
â
Ð
a
Ы
<u>n</u>
en
IS:
ž
Ē
Č
Ť
he
Ħ
of the fitted
6
à
SSI
iar
2
list
fi
ŭ
ğ
ns.
-

6 of 24 Rafaella Georgiou, Rachel S. Popelka-Filcoff, Dimosthenis Sokaras, Victoria Beltran, Ilaria Bonaduce, Jordan Spangler, Serge Cohen, Roy Lehmann, Sylvain Bernard, Jean-Pascal Rueff, Uwe Bergmann, Loïc Bertrand

	5 -	–Gaussian 1	1	9 -	– Gaussian 2	2	ອ -	Gaussian 3	3	– Gé	– Gaussian 4	4	-G£	–Gaussian 5	5	- Gê	– Gaussian 6	9	- 0	–Gaussian 7	7
	μ _{Fit}	σ _{Fit}	lFit	μFit	σ _{Fit}	l _{Fit}	μFit	σ _{Fit}	l _{Fit}	μFit	σ _{Fit}	<u>Et</u>	μ _{Fit}	σ _{Fit}	leit	μFit	σ _{Fit}	LEI.	μFit	σ _{Fit}	Let.
X. arborea	531.1	0.51	0.18	531.9	0.51	0.28	533	0.51	0.23	534.1	0.51	0.29	535.2	0.51	0.49	536.3	0.59	0.71	537.8	0.85	1.07
X. semiplana ssp. tateana	531.4	0.51	0.28	531.9	0.51	0.12	533	0.51	0.18	534.2	0.51	0.26	535.2	0.51	0.50	536.3	0.59	0.97	537.9	0.85	1.32
X. semiplana	530.9	0.51	0.18	531.9	0.51	0:30	533	0.51	0.11	534.1	0.51	0.27	535.2	0.51	0.53	536.3	0.59	0.89	537.9	0.85	1.16
C. calcarata	530.8	0.51	0.40	531.9	0.51	0.78	533	0.51	0.37	534.3	0.51	0.34	535.2	0.51	0.54	536.3	0.59	0.83	537.9	0.85	1.23
C. glauca	531	0.51	0.19	532	0.51	0.70	533	0.51	0.43	533.9	0.51	0.02	535.2	0.51	0.65	536.3	0.59	0.75	537.8	0.85	0.99
C. preissii ssp. verrucosa	531.2	0.51	0.52	532.3	0.51	0.85	533	0.51	0.17	534.1	0.51	0.30	535.2	0.51	0.52	536.3	0.59	1.01	537.8	0.85	1.21
White Mallee	530.8	0.51	0.06	531.9	0.51	0.28	533	0.51	0.16	534.4	0.51	0.24	535.2	0.51	0.43	536.3	0.59	0.89	537.9	0.85	1.15
E. largiflorens	531	0.51	0.25	532.1	0.51	0.17	533	0.51	0.25	534.3	0.51	0.18	535.2	0.51	0.50	536.3	0.59	1.00	537.9	0.85	1.47
<i>Acacia</i> sp.	531.2	0.51	0.44	532.2	0.51	0.51	533	0.51	0.34	533.9	0.51	0.31	535.2	0.51	0.54	536.3	0.59	0.93	537.8	0.85	1.35
A. bakeri	530.8	0.51	0.27	532	0.51	0.66	533	0.51	0.39	534.4	0.51	0.33	535.2	0.51	0.37	536.3	0.59	0.71	537.7	0.85	1.16

Table S5. XRS oxygen K-edge spectra decomposition. Center (µ_{Fit}), standard deviation (σ_{Fit}) and Intensity (I_{Fit}) of the fitted Gaussian distributions)

	– Gaussian 1	– Gaussian 2	– Gaussian 3	-Gaussian 4	– Gaussian 5	– Gaussian 6	– Gaussian 7
	Aromatic C =O			^a OH ^b C _x H _y		R—OH	$\begin{array}{c} C \\ C \\ C \\ C \\ C \\ C_n H_{2n+1} \end{array}$
Energy transfer (eV)	284.4–284.5	285.4–285.6	286.3–286.9	287.3–287.9	288.3–288.8	289.2–289.5	290.2–290.5
Transitions	1s-π*	1s–π*	1s-π*	^a 1s–π* ^b 1s–3p/σ*	1s–π*	1s–3p/σ*	^c 1s-π* ^d 1s-4p*
			A_{f}	_{fit} (Normalized intensi	ity, eV)		
X. arborea	0.17	1.4	0.59	0.92	0.81	0.93	1.53
X. semiplana ssp. tateana	0.13	1.38	0.41	1.07	0.73	0.91	1.48
X. semiplana	0.13	1.26	0.65	0.96	0.69	0.85	1.43
C. calcarata	0.05	0.33	0.24	0.65	1.01	0.67	1.28
White Mallee	0.07	0.62	0.37	1.11	0.69	0.55	0.9
E. largiflorens	0.06	0.64	0.33	1.08	0.62	0.82	1.27
Acacia sp.	0.00	0.01	0.16	0.12	0.76	1.31	1.55
A. bakeri	0.03	0.27	0.66	0.69	1.46	1.36	2.62

Table S6. Contribution of functional groups to the XRS carbon K-edge spectra. Integrated area of the fitted Gaussian distributions (calculated as $A_{\text{fit}} = I_{\text{fit}} \cdot \sigma \cdot \sqrt{2\pi}$)

C. Statistical Testing of XRS Data Classification. Fisher linear discrimination is finding the projections which maximizes the ratio J of the *between-group scatter* divided by the *within-group scatter*. Formally, for a given set of projections defined by their matrix \mathbf{W} , noting \mathbf{S}_B and \mathbf{S}_W respectively the *between-group* and *within-group* scatter matrices and for any square matrix \mathbf{M} noting $|\mathbf{M}|$ its determinant, J is defined by:

$$J(\mathbf{W}) = \frac{|\mathbf{W}^t \mathbf{S}_B \mathbf{W}|}{|\mathbf{W}^t \mathbf{S}_W \mathbf{W}|}$$

Once the LDA is performed, the maximum J can also be computed as the product of the singular values attached to each of the proposed projections. Whilst J is a good metric of the ability of the set of projection to separate the dataset into predefined groups, its scale is too dependent upon the number of samples, the number of groups and the number of variables. Hence the value of J can not be used alone to judge if a given LDA is producing a meaningful discrimination. To assess the quality of the discrimination obtained based on the genera of the samples, we compared the value of J with the one obtained for all other possible (random) grouping of the samples with the same group sizes, that is with respectively 3, 2, 2 and 1 members. From the 8! = 40,320 possible permutations of the samples, we kept a single representative from all equivalent groupings: two groupings are equivalent if they differ only by the order of the samples within each groups (leading to $6 \times 2 \times 2 = 24$ equivalents) or by swapping the two groups made of exactly 2 samples (leading to 2 equivalents). This way, we obtained all the 840 non-equivalent groupings of the samples. For each of these (random) grouping, the LDA is performed and the corresponding value of J is recorded and reported on a histogram of possible $\log(J)$ values, Fig. 3B. Using the grouping based on the genera of the samples, we obtained J = 24.91 from the LDA, this being the eleventh highest value of the J computed for all (840) possible groupings. In regard of this rank, the null hypothesis "this value of J is obtained by pure chance" can be rejected with a *p*-value of 11/840 = 1.31%. Note that in this test the lowest possible *p*-value would have been 1/840 = 0.12% and that the standard threshold (p = 5%) corresponds to a value of J = 11.62. All statistical computations were performed using the R Language and Environment for Statistical Computing (18). LDA was performed using the Ida function from the MASS package (19).

D. FT-IR. FT-IR spectra were collected with a spectrometer (Alpha II, Bruker) equipped with a DTGS detector. Measurements were performed in transmission mode using KBr pellets (Sigma Aldrich, FT-IR grade, ref. 221864). A total of 128 scans were accumulated for each sample at a resolution of 4 cm⁻¹ in the range of 4000 to 400 cm⁻¹. The spectra presented have not been corrected to avoid any kind of distortion.

D.1. Xanthorrhoea resins. The FT-IR absorbance spectra of Xanthorrhoea species (X. arborea, X. semiplana ssp. tateana, X. semiplana) (Fig. 3) are dominated by the following features: (a) a broad signal in the $3550-3250 \text{ cm}^{-1}$ region attributed to -OH stretching, (b) a weak intensity peak observed at 3023 cm^{-1} attributed to the CH stretching in alkene groups, (c) weak intensity peaks centered at 2962 cm^{-1} , 2936 cm^{-1} , 2840 cm^{-1} , attributed to the CH₃ and CH₂ antisymmetric stretch and CH₂ symmetric stretch respectively, (d) a peak centered at 1636 cm^{-1} (mainly observed in X. arborea spectrum) attributed to C=C stretching in conjugated C=C bonds, (e) high intensity sharp peaks centered at 1602 cm^{-1} , 1513 cm^{-1} and 1444 cm^{-1} attributed to C=C stretching indicating the presence of aromatic groups, (f) several sharp bands in the range $1275-1000 \text{ cm}^{-1}$ attributed to in-plane C-H deformation, (g) the bands at 1337 cm^{-1} and 1205 cm^{-1} linked to the O-H deformation and the C-OH stretching, and (h) a strong band centered at 832 cm^{-1} attributed to out-of-plane C-H deformation.

D.2. Callitris resins. The FT-IR absorbance spectra of Callitris species (C. glauca, C. calcarata, C. preissii ssp. verrucosa) (Fig. 3) are dominated by the following features: (a) a broad band centered at 3450 cm^{-1} corresponding to the stretching of polymeric O-H, (b) a band at 3080 cm^{-1} corresponding to the stretching vibration of C=C bonds, (c) strong intensity peaks centered at 2957 cm^{-1} , 2937 cm^{-1} , 2873 cm^{-1} , 2845 cm^{-1} attributed to the CH₂ and CH₃ antisymmetric stretch and CH₂ and CH₃ symmetric stretch respectively, the presence of these groups is confirmed by the bands around 1467 cm^{-1} (-CH₂-) and around 1384 cm^{-1} (deformation of CH₃ and CH₂ groups), (d) a strong broad signal centered at 1695 cm^{-1} attributed to C=O stretching from the carboxylic acid; (e) a band centered at 1644 cm^{-1} attributed to the presence of conjugated C=C bonds.

D.3. Eucalyptus kino. The FT-IR spectra of *Eucalyptus* species (White Mallee, *E. largiflorens*) (Fig. 3) exhibit the following features: (a) an intense broad band centered at 3326 cm^{-1} corresponding to the stretching of polymeric O–H, (b) a peak at 1700 cm^{-1} attributed to C=O stretching probably from COOH groups, (c) high intensity bands at 1612 cm^{-1} , 1534 cm^{-1} and 1447 cm^{-1} attributed to the C=C stretching from aromatic compounds, (d) two bands centered at $762 \text{ and } 734 \text{ cm}^{-1}$ attributed to the =CH out-of-plane bending vibrations.

D.4. Acacia. The FT-IR spectra of Acacia species (A. bakeri, Acacia sp.) exhibit the following features (Fig. 3): (a) an intense broad band in the 3600–3000 cm⁻¹ range attributed to O–H stretching, (b) a peak centered at 2940 cm⁻¹ attributed to the CH₂ antisymmetric stretching, and a band at 2890 cm⁻¹ linked to the CH stretching, (c) a broad asymmetrical band centered at 1609 cm^{-1} possibly linked with the presence of absorbed water, (d) two intense bands at 1420 cm^{-1} and at 1380 cm^{-1} possibly attributed to the deformation of CH₃ and CH₂ groups, (e) a broad band centered around 1070 cm^{-1} and two bands centered at 1160 cm^{-1} and 1130 cm^{-1} linked to the skeletal vibrations of polysaccharides. The differences in intensity among the two species is possibly related to the proportion of monosaccharides or their organization in the polysaccharide chain.

E. Thermally assisted hydrolysis and methylation (THM) using tetramethylammonium hydroxide (TMAH) coupled on line to GC-MS (THM/GC-MS). Pyrolysis gas chromatography mass spectrometry experiments were performed at Flinders University, using a Frontier Laboratories single-Shot pyrolyzer PY-3030S-220 (Shimadzu North America, USA) coupled to an Agilent

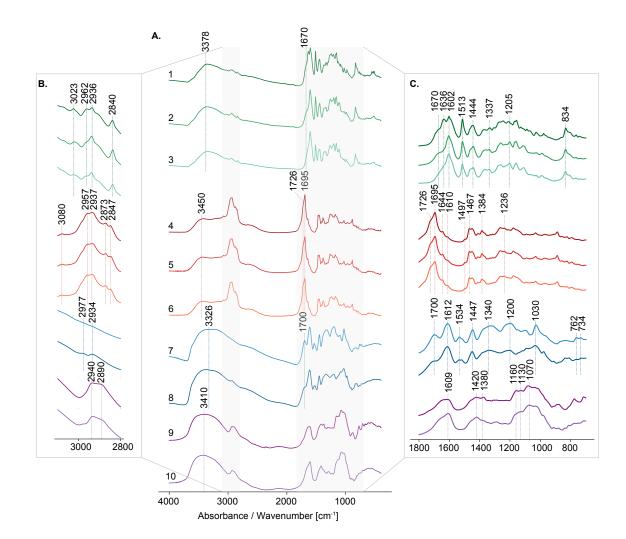


Fig. S3. FT-IR spectra of native Australian plant exudates. 1, X. arborea; 2, X. semiplana ssp. tateana; 3, X. semiplana; 4, C. calcarata; 5, C. glauca; 6, C. preissii ssp. verrucosa; 7, White Mallee; 8, E. largiflorens; 9, Acacia sp., 10, A. bakeri

Technologies 7890A GC System (Agilent Technologies, Santa Clara, USA) with an Agilent Technologies 5975C inert XL EI/CI MSD with triple axis detector (Agilent Technologies, Santa Clara, USA). The sample preparation procedures followed Reeves et al. (20), with preparation of the samples using in-situ derivatisation with TMAH. A HP-5MS Ultra Inert capillary GC column ($30 \text{ m} \times 250 \text{ µm} \times 0.25 \text{ µm}$) with a 5% phenyl 95% methyl polysiloxane stationary phase was used. The carrier gas was helium with a flow rate of $1 \text{ mL} \cdot \text{min}^{-1}$, a split ratio of 50:1, and a solvent delay of 4 min. The sample was subject to pyrolysis for 3 min at a continuous temperature of 600°C, and interface temperature was maintained at 300°C for the duration of pyrolysis. The GC oven temperature program was: 50°C for 5 min, then ramped to 240°C at 5°C · min⁻¹ before being held for 5 min, then ramped to 300°C at 5°C · min⁻¹ and held for a further 8 min. Electron Ionization (EI) at 70 eV was used for mass spectra collection over the m/z range of 30–350. The MS ion source and quadrupole were held at 230°C and 150°C, respectively. Identification of fragments was completed using Mass spectral software MSD ChemStation E.02.02.1431 software (Agilent Technologies, Inc.), which utilized the National Institute of Standards and Technology (NIST) Standard Reference Database Number 69 – NIST Chemistry WebBook (21).

E.1. Xanthorrhoea resins. The literature on the chemical composition of the resins produced by members of the genus Xanthorrhoea is scarce and dates back several years. It is thus not possible to find a well-defined description of the chemical species that can be found in these resins, and data are strongly related to the analytical procedures and methodologies used in the various studies. Nonetheless, it is known that Xanthorrhoea resins are phenolic resins, and comprise a polymer fraction, which is reported to be a polyflavanoid material, and smaller molecules, which include chalcones and flavanones mostly accompanied by varying amounts of free acids, essentially p-coumaric and cinnamic acids (see Duewell et al. (22) and references therein; and Birch et al. (23) and references therein.)

The three *Xanthorrhoea* resins characterized in this study present very similar compositions from a qualitative point of view. Chromatographic profiles and identified compounds are shown in Fig. S4.

The analysis of the pyrolitic profiles obtained allows to assert that the three resins contain:

- (i) Cinnamic acid (shown in red in Fig. S4).
- (ii) Chalcones: given the identification at high retention times of 2',4,4',6'-tetramethoxychalcone (56.42 min) and of its pyrolysis products, 3',4',5'-trimethoxyacetophenone and 3,4,5-trimethoxybenzaldehyde (shown in blue in Fig. S4). At higher retention times (59.20 min) there is a peak whose mass spectrum presents a similar fragmentation pattern of 2',4,4',6'-tetramethoxychalcone, and can be ascribed to a molecule with an extra methoxyl group. As mass spectra acquisition was up to m/z 350, the molecular ion, which should be at m/z 358, is not visible. Given the identification of methyl (2E)-3-(3,4-dimethoxyphenyl)prop-2-enoate in the pyrogram, the mass spectrum is tentatively assigned to 2',3,4,4',6'-pentamethoxychalcone. The mass spectra and molecular structures are shown in Fig. S5.
- (iii) Flavanones and polyflavanoids: Their identification is possible thanks to the detection, among several possible pyrolysis products, of 1-(2,4-dimethoxyphenyl)propan-1-one and 1-(2,4-dimethoxyphenyl)ethanone (in green in Fig. S4), whose structure can very clearly be related to the structure of flavanones. Moreover, at high retention times, some relatively abundant peaks were present, whose mass spectra were not found in the libraries available (Fig. S5, peaks A and B). The molecular ion of the intense peak at 54.5 min is at m/z 298 (Fig. S5, peak A). This value is compatible with the molecular mass of a trimethoxyflavan (with an extra unsaturation in its structure) or that of a hydoxy-dimethoxy-flavone, compounds whose identification has been reported in previous studies (23). At 57.8 min another intense peak is visible whose mass spectrum presents a molecular ion at m/z 328, which is 30 Da higher than the previous one, possibly due to the same molecule with an extra methoxyl group (Fig. S5, peak B). Again this mass is compatible with that of flavanes reported in the literature (22). The mass spectra of the peaks at 54.5 min and 57.8 min are shown in Fig. S5.
- (iv) In addition, at 33.69 min an intense peak (labeled with '*' in Fig. S4) is present, especially in X. semiplana and X. tateana, whose mass spectrum was not present in the libraries of mass spectra available. The mass spectrum is shown in Fig. S6 and the molecular ion is compatible with that of xanthorrhoeol methyl ether, which has been identified among Xanthorrhoea resins' constituents (24). More research is necessary to confirm the attribution.

E.2. Callitris resins. The pyrogram of *Callitris calcarata* is dominated by a peak ascribable to methyl sandaracopimarate. Methyl sandaracopimarate and methyl callitrisate are the main peaks in the pyrogram of *Callitris glauca*. Methyl callitrisate dominates the pyrogram of *Callitris verrucosa*. Compounds are identified by comparison of their mass spectra with the literature, and in particular with Scalarone et al. (25) and Simoneit et al. (26). Trace amounts of methyl callitrisate are detected also in the pyrogram of *Callitris calcarata*. Methyl isopimarate is also detected with relatively high intensity in the pyrograms of both *Callitris glauca* and *Callitris verrucosa* (Fig. S7).

These are products of methylation produced upon the thermally assisted reaction with TMAH. Sandaracopimaric and isopimaric acids are pimarane diterpenoid acids and callitrisic acid is an abietane diterpenoid acid. Small amounts of *iso*-communic acid (methyl ester), a labdane diterpenoid, are also detected, together with common pyrolysis products of polycommunic acid (25) (Fig. S8).

At high retention times, shown in Fig. S7, the peaks whose mass spectra present molecular ions at m/z 346 and 344 are compatible with the methoxy derivatives of methyl sandaracopimarate, methyl isopimarate and methyl callitrisate. Hydroxycallitrisic acid is commonly found in resins from Callitris species (26). 7-oxocallitrisic acid methyl ester is also detected

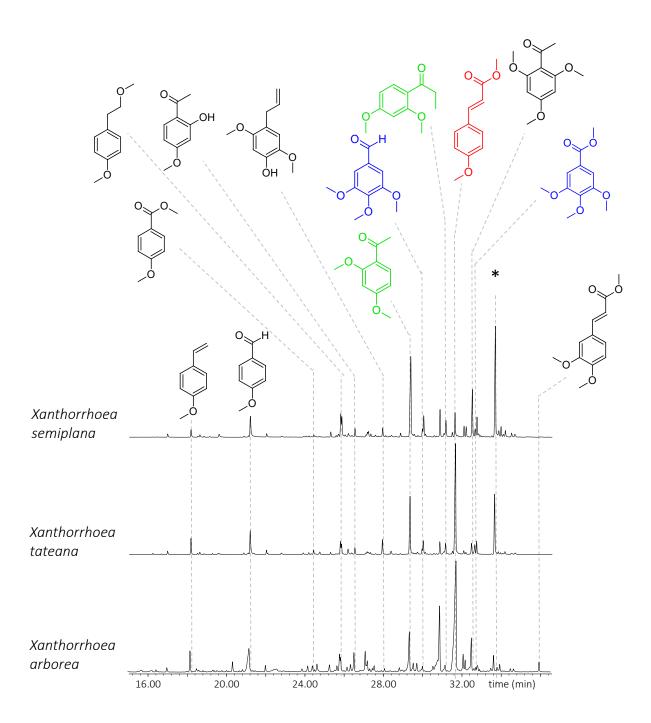


Fig. S4. THM/GC-MS chromatograms relative to X. semiplana, X. tateana and X. arborea with main compounds identified. In red cinnamic acid; in green pyrolysis products of flavanones; in blue pyrolysis products of chalcones/flavanones/polyflavanoids. The peaks labeled with "*" are characterized by the mass spectrum shown in Fig. S5.

at molecular ion m/z 328. The chromatographic profiles of these products of oxidation of abietane and pimarane terpenoid acids are shown in Fig. S9.

Finally monoterpenoids are also identified at low retention times. Pyrolytic profiles are quite similar among the three resins. The main compounds identified are shown in Fig. S10.

E.3. Eucalyptus kinos. The THM/GC-MS chromatograms of both *Eucalyptus* kinos are dominated by the peak ascribable to the fully methylated derivative of gallic acid (Fig. S11). In addition, in the pyrogram of *E. largiflorens* other aromatic compounds are found, including 1,3,5-trimethoxybenzene and 1,3,5-trimethoxy-2-methyl-benzene. Sugars are present in both

12 of 24 Rafaella Georgiou, Rachel S. Popelka-Filcoff, Dimosthenis Sokaras, Victoria Beltran, Ilaria Bonaduce, Jordan Spangler, Serge Cohen, Roy Lehmann, Sylvain Bernard, Jean-Pascal Rueff, Uwe Bergmann, Loïc Bertrand

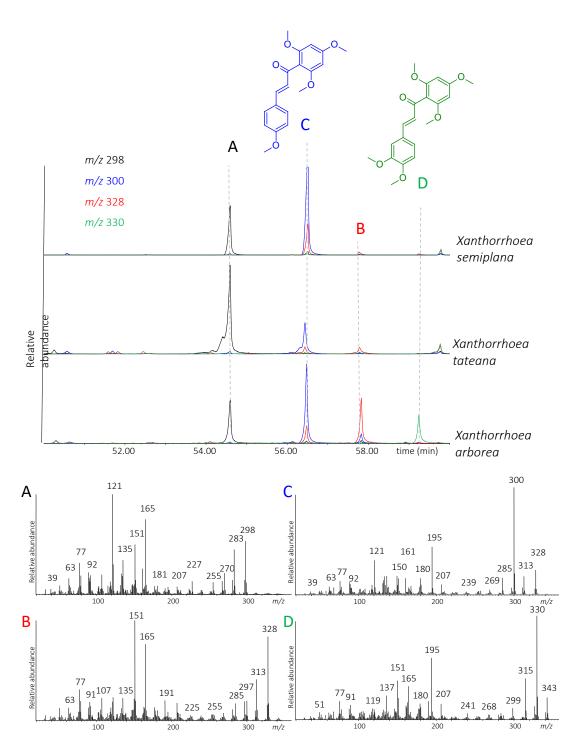


Fig. S5. THM/GC-MS extracted ion chromatograms of fragment ions m/z 298 (black) 300 (blue) 328 (red) and 330 (green) and mass spectra of peaks A, B, C and D.

kinos, identified thanks to the presence of variably methylated 3,6-deoxy-hexonic acid methyl esters in the pyrograms (27).

Methylated derivatives of gallic acid and deoxy-sugar acids derive from the hydrolysis and methylation of hydrolysable tannins. These comprise structures in which gallic acid is linked via ester bonds to sugars. One of the simplest hydrolysable tannin is pentagalloyl glucose, whose likely products of the TMAH thermally assisted reaction are shown in Fig. S12.

Glucose is the most common sugar in hydrolysable tannins (28). Variably methylated 3,6-deoxy-hexonic acid methyl esters have been reported as products of pyrolysis of rhamnose and fucose (27, 29). Their detection in the pyrograms reported in this study is most likely the result of thermally assisted hydrolysis and methylation of a glucose unit in which every hydoxyl moiety is linked via an ester bond.

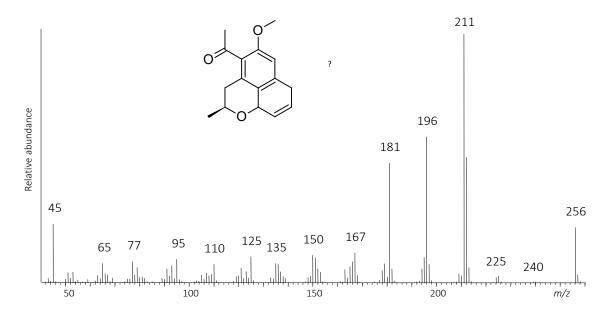


Fig. S6. Mass spectrum of the peak at 33.69 min, and chemical structure of xanthorrhoeol methyl ether, whose molecular mass is 256 Da.

The other aromatic compounds detected in both pyrograms, especially in the pyrogram of *E. largiflorens*, cannot be ascribed to the pyrolysis of hyrolysable tannins, but to that of condensed tannins. Condensed tannins are oligomers and polymers derived from poly hydroxyflavan-3-ols linked via carbon-carbon bonds between the flavanol subunits (28). Two common flavan-3-ol structures are catechin and gallocatechin, whose dimer and possible pyrolysis products are shown in Fig. S13.

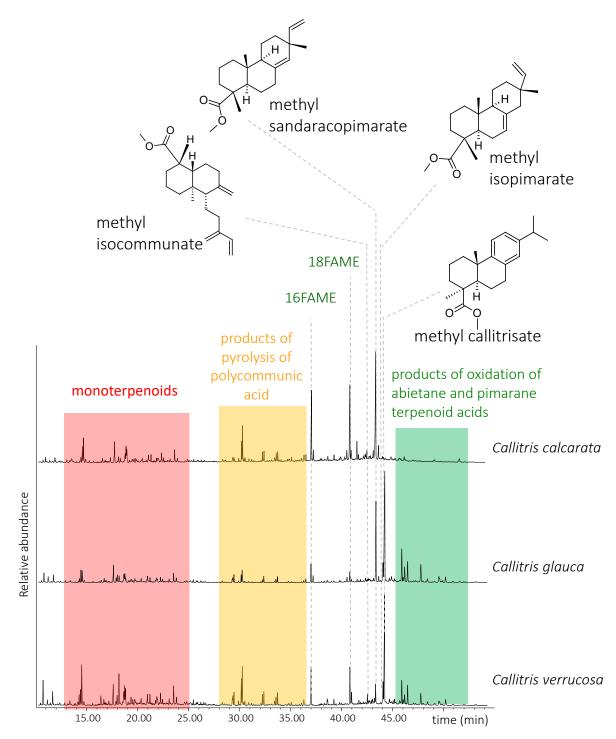


Fig. S7. THM/GC-MS chromatograms relative to Callitris calcarata, Callitris glauca, Callitris verrucosa. 16 FAME: palmitic acid methyl ester; 18 FAME: stearic acid methyl ester.

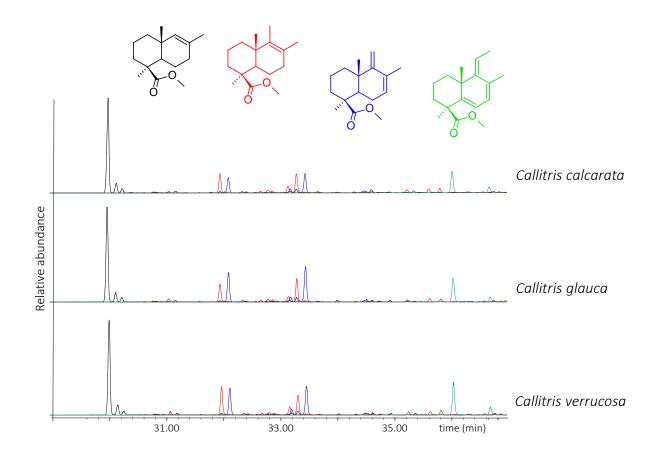


Fig. S8. THM/GC-MS extracted ion chromatograms relative to Callitris calcarata, Callitris glauca, Callitris verrucosa of ions m/z 236 (black), 248 (red), 250 (blue), 270 (green).

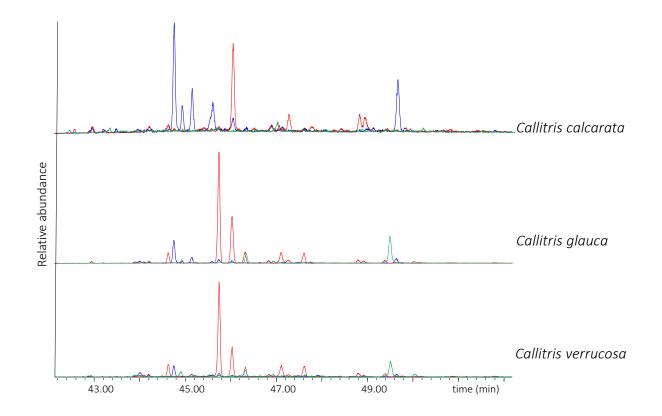


Fig. S9. THM/GC-MS extracted ion chromatograms relative to Callitris calcarata, Callitris glauca, Callitris verucosa of ions m/z 346 (blue; compatible with the methoxy derivatives of methyl sandaracopimarate), 344 (red; compatible with the methoxy derivatives of methyl callitrisate), 328 (green; 7-oxocallitrisic acid methyl ester).

17 of 24

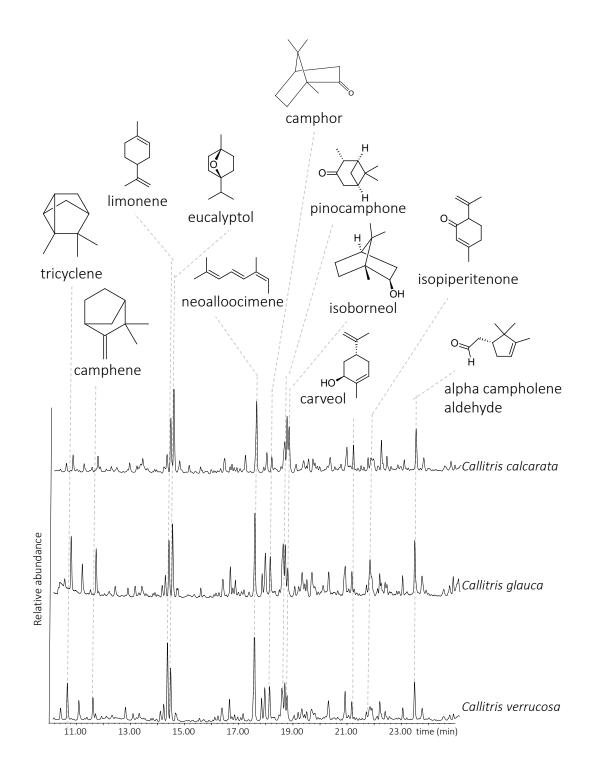


Fig. S10. THM/GC-MS extracted ion chromatograms relative to Callitris calcarata, Callitris glauca, Callitris verrucosa in the area of elution of monoterpenenoids, with main compounds identified.

18 of 24 Rafaella Georgiou, Rachel S. Popelka-Filcoff, Dimosthenis Sokaras, Victoria Beltran, Ilaria Bonaduce, Jordan Spangler, Serge Cohen, Roy Lehmann, Sylvain Bernard, Jean-Pascal Rueff, Uwe Bergmann, Loïc Bertrand

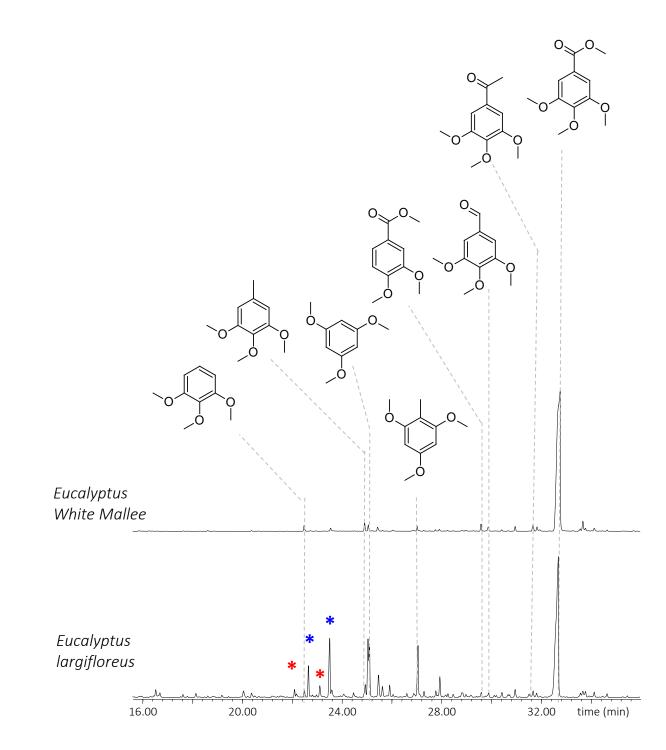


Fig. S11. THM/GC-MS extracted ion chromatograms relative to *Eucalyptus White Mallee* and *E. largiflorens* with main compounds identified. Peaks labeled with red asterisks are relative to partially methylated 3,6-deoxy-pentonic acid methyl esters, and those labeled with blue asterisks are relative to permethylated 3,6-deoxy-pentonic acid methyl esters.

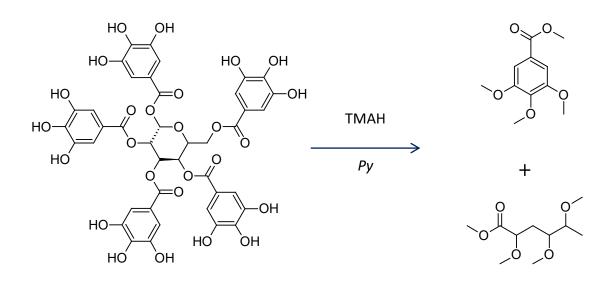


Fig. S12. Pentagalloyl glucose and main products of TMAH thermally assisted reaction.

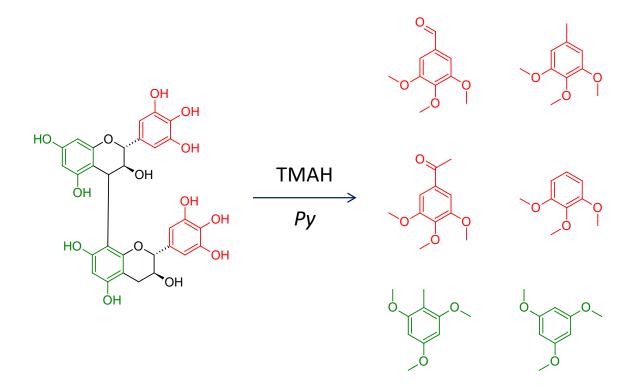


Fig. S13. Gallocatechin dimer and possible main products of TMAH thermally assisted reaction.

F. Gas chromatography mass spectrometry. The analytical procedure used was adapted from a previously published work (30). Samples are subject to acidic hydrolysis assisted by microwaves with trifluoroacetic acid 2 M (power 500 W, temperature 120°C, duration 20 min). After hydrolysis, an aliquot is transferred into conic vials, added with a water solution of mannose as internal standard (126 ppm) and dried under nitrogen flow. Then, 25 µl of ethanethiol/trifluoroacetic acid (2/1, v/v) are added and the resulting solution is kept at room temperature for 10 min. The mercaptalation mixture is then subjected to silylation prior to GC/MS analysis. The silylation is performed in two steps. In the first step, 100 µl of BSTFA is added to the mercaptalation mixture and the mixture is kept for 15 min at 60°C. The solution is then dried under a nitrogen flow, and subsequently added with 50 µl of BSTFA (1% TMCS) as a derivatising agent and 100 µl pyridine as a solvent, and kept at 60°C for 45 min. The reaction mixture is then dried under a nitrogen flow and reconstituted in 50 µl of hexane. 2 µl of this solution, containing diethyl-dithioacetal trimethylsilyl derivatives of the parental sugars, is then injected into the gas chromatograph.

Apparatus A microwave oven MLS MEGA Milestone 1200W (Milestone Microwave Laboratory System, Monroe, CT, USA) was used. Hydrolysis conditions were: power 500 W, temperature 120°C, duration 20 min. 6890N GC System Gas Chromatograph (Agilent Technologies, Palo Alto, CA, USA), coupled with a 5975 Mass Selective Detector (Agilent Technologies, Palo Alto, CA, USA) single quadrupole mass spectrometer, equipped with a PTV injector were used at DCCI. Instrumental parameters are reported elsewhere (30).

F.1. Acacia gums. Acacia gums are polysaccharide materials. Polysaccharides are hydrolysable polymers, they are thus more suited to GC-MS analysis after hydrolysis and derivatisation, rather than analytical pyrolysis, where establishing a relationship among the saccharide composition and the pryolytic profile is not straightforward (27, 31).

The chromatographic profiles of the two gums are shown in Fig. S14, and the sugar composition is reported in Table S7 together with those of *Acacia* gums from some African species, determined with the same analytical procedure (30).

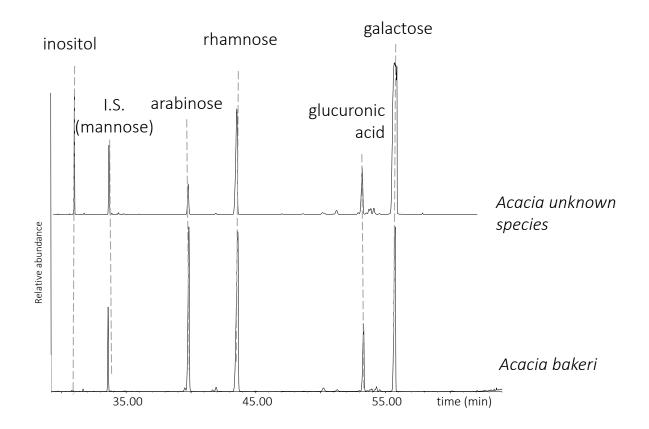


Fig. S14. Chromatographic profiles of the two Acacia gums after hydrolysis, derivatisation and GC-MS analysis. I.S.: internal standard. Inositol is detected with all OH moieties as TMS ethers. Sugars are detected as diethylmercaptal derivatives with all OH moieties as TMS ethers.

Table S7. Sugar composition of *Acacia* sp. and *Acacia bakeri*, and comparison with *Acacia* gums from other African species, determined with the same analytical procedure (30).

<i>Acacia</i> gum	General origins of plant family	xylose	arabinose	rhamnose	fucose	galacturonic acid	glucuronic acid	glucose	mannose	galactose
Gum obtained from the sap of trees Acacia giraffe	Africa (India has several spp.)	-	30	4	-	-	-	-	1	65
Gum obtained from the sap of trees Acacia karoo	South Africa	-	59	2	-	-	-	-	-	39
Gum obtained from the sap of trees Acacia seyal	Senegal to Sudan, Africa	-	63	0	-	-	-	-	-	36
Gum obtained from the sap of trees Acacia senegal	Tropical Africa	-	26	13	-	-	12	-	-	29
Gum obtained from the sap of the trees of Acacia sp.	Australia	-	12	51	-	-	7	7	-	24
Gum obtained from the sap of the trees of Acacia bakeri	Australia	-	2	31	-	-	6	6	-	55

23 of 24

References

- 1. D Sokaras, et al., A high resolution and large solid angle x-ray Raman spectroscopy end-station at the Stanford Synchrotron Radiation Lightsource. *Rev. Sci. Instruments* 83, 043112 (2012).
- U Bergmann, et al., Carbon K-edge X-ray Raman spectroscopy supports simple, yet powerful description of aromatic hydrocarbons and asphaltenes. *Chem. Phys. Lett.* 369, 184–191 (2003).
- 3. JP Rueff, et al., The GALAXIES beamline at the SOLEIL synchrotron: inelastic X-ray scattering and photoelectron spectroscopy in the hard X-ray range. J. Synchrotron Radiat. 22, 175–179 (2015).
- JM Ablett, et al., The GALAXIES inelastic hard X-ray scattering end-station at Synchrotron SOLEIL. J. Synchrotron Radiat. 26, 263–271 (2019).
- 5. P Gueriau, et al., Noninvasive synchrotron-based x-ray Raman scattering discriminates carbonaceous compounds in ancient and historical materials. Anal. Chem. 89, 10819–10826 (2017).
- K Heymann, J Lehmann, D Solomon, MW Schmidt, T Regier, C 1s K-edge near edge X-ray absorption fine structure (NEXAFS) spectroscopy for characterizing functional group chemistry of black carbon. Org. Geochem. 42, 1055–1064 (2011).
- C Le Guillou, S Bernard, F De la Pena, Y Le Brech, XANES-based quantification of carbon functional group concentrations. Anal. Chem. 90, 8379–8386 (2018).
- V Rouchon, S Bernard, Mapping iron gall ink penetration within paper fibres using scanning transmission X-ray microscopy. J. Anal. At. Spectrom. 30, 635–641 (2015).
- 9. J Alleon, et al., Organic molecular heterogeneities can withstand diagenesis. Sci. Reports 7, 1–9 (2017).
- G Cody, H Ade, S Wirick, GD Mitchell, A Davis, Determination of chemical-structural changes in vitrinite accompanying luminescence alteration using c-nexafs analysis. Org. Geochem. 28, 441–455 (1998).
- S Bernard, et al., Geochemical evolution of organic-rich shales with increasing maturity: A STXM and TEM study of the Posidonia Shale (Lower Toarcian, northern Germany). Mar. Petroleum Geol. 31, 70–89 (2012).
- A Hitchcock, C Brion, Inner-shell excitation of formaldehyde, acetaldehyde and acetone studied by electron impact. J. Electron Spectrosc. Relat. Phenom. 19, 231–250 (1980).
- 13. I Ishii, A Hitchcock, A quantitative experimental study of the core excited electronic states of formamide, formic acid, and formyl fluoride. *The J. Chem. Phys.* 87, 830–839 (1987).
- 14. M Robin, I Ishii, R McLaren, A Hitchcock, Fluorination effects on the inner-shell spectra of unsaturated molecules. J. Electron Spectrosc. Relat. Phenom. 47, 53–92 (1988).
- 15. K Harding, et al., Inner-shell excitation spectroscopy of peroxides. Chem. Phys. 461, 117–124 (2015).
- 16. I Ishii, A Hitchcook, The oscillator strengths for C1s and O1s excitation of some saturated and unsaturated organic alcohols, acids and esters. J. Electron Spectrosc. Relat. Phenom. 46, 55–84 (1988).
- 17. J Francis, A Hitchcock, Inner-shell spectroscopy of p-benzoquinone, hydroquinone, and phenol: distinguishing quinoid and benzenoid structures. *The J. Phys. Chem.* **96**, 6598–6610 (1992).
- 18. R Core Team, R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, Austria), (2020).
- 19. WN Venables, BD Ripley, *Modern Applied Statistics with S.* (Springer, New York), Fourth edition, (2002) ISBN 0-387-95457-0.
- 20. T Reeves, RS Popelka-Filcoff, CE Lenehan, Towards identification of traditional European and indigenous Australian paint binders using pyrolysis gas chromatography mass spectrometry. *Anal. Chimica Acta* **803**, 194–203 (2013).
- P Linstrom, W Mallard, NIST Chemistry WebBook, NIST standard reference database. (National Institute of Standards and Technology) Vol. 69, (2003).
- 22. H Duewell, Chemotaxonomy of the genus Xanthorrhoea. Biochem. Syst. Ecol. 25, 717-738 (1997).
- A Birch, C Dahl, Some constituents of the resins of Xanthorrhoea preissii, australis and hastile. Aust. J. Chem. 27, 331–334 (1974).
- 24. H Duewell, Studies on Xanthorrhoea resins. III. Xanthorrhoeol. Aust. J. Chem. 18, 575–581 (1965).
- 25. D Scalarone, M Lazzari, O Chiantore, Ageing behaviour and analytical pyrolysis characterisation of diterpenic resins used as art materials: Manila copal and sandarac. J. Anal. Appl. Pyrolysis 68, 115–136 (2003).
- BR Simoneit, RE Cox, DR Oros, A Otto, Terpenoid compositions of resins from *Callitris* species (*Cupressaceae*). Molecules 23, 3384 (2018).
- 27. C Riedo, D Scalarone, O Chiantore, Advances in identification of plant gums in cultural heritage by thermally assisted hydrolysis and methylation. *Anal. Bioanal. Chem.* **396**, 1559–1569 (2010).
- S Von Martius, K Hammer, C Locher, Chemical characteristics and antimicrobial effects of some *Eucalyptus* kinos. J. Ethnopharmacol. 144, 293–299 (2012).
- D Fabbri, R Helleur, Characterization of the tetramethylammonium hydroxide thermochemolysis products of carbohydrates. J. Anal. Appl. Pyrolysis 49, 277–293 (1999).
- 30. A Lluveras-Tenorio, J Mazurek, A Restivo, MP Colombini, I Bonaduce, Analysis of plant gums and saccharide materials in paint samples: comparison of GC-MS analytical procedures and databases. *Chem. Cent. J.* **6**, 1–16 (2012).
- 31. A Andreotti, I Bonaduce, MP Colombini, F Modugno, E Ribechini, A diagnosis of the yellowing of the marble high reliefs and the black decorations in the chapel of the tomb of Saint Anthony (Padua, Italy). Int. J. Mass Spectrom. 284, 123–130 (2009).