South African Journal of Botany

Supplementary Material

Article title: **Integrating chemotaxonomic-based metabolomics data with DNA barcoding for plant identification: A case study on south-east African Erythroxylaceae species.**

Authors: P.S.F. Alberts* and J.J.M. Meyer

Department of Plant and Soil Sciences, University of Pretoria, Pretoria, 0002, South Africa
*Author for correspondence:
P.S.F. Alberts
Tel: +27 12 420 2224
E-mail: sewes.alberts@up.ac.za

The following Supplementary Material is available for this article:

Fig. S. 1. The mass fragmentation patterns of the identified biomarkers and library references.

Fig. S. 2. Flow diagram describing the main experimental and analysis steps used, and the average time required (in an ideal scenario) to complete each section.

Table S. 1. The collection sites, PRU numbers and GenBank accession numbers of the plant species assessed.

Table S. 2. DNA primer sequences used in the study.

Table S. 3. Summary of the financial cost comparison between different analytical techniques used in plant identification.

Additional data for downloading from the Figshare (https://figshare.com/) data repository:

1) Alberts & Meyer. SA Journal of Botany. 2021.
Full financial cost evaluation and calculations:
https://doi.org/10.25403/UPresearchdata.16587338

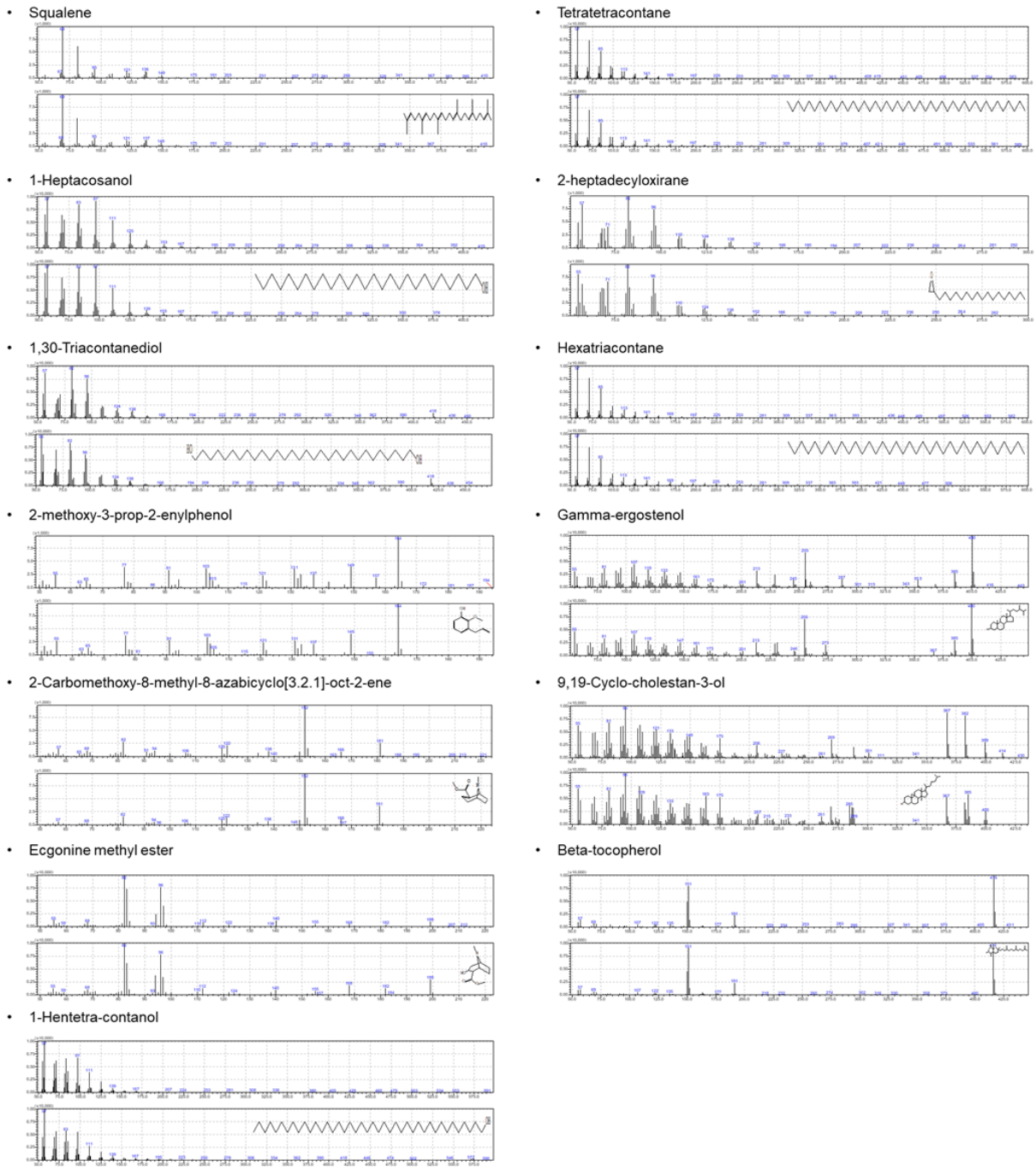

Fig. S. 1. The mass fragmentation patterns of the biomarker (top) compared to the predicted standard compound (bottom) present in the NIST 11 reference database.

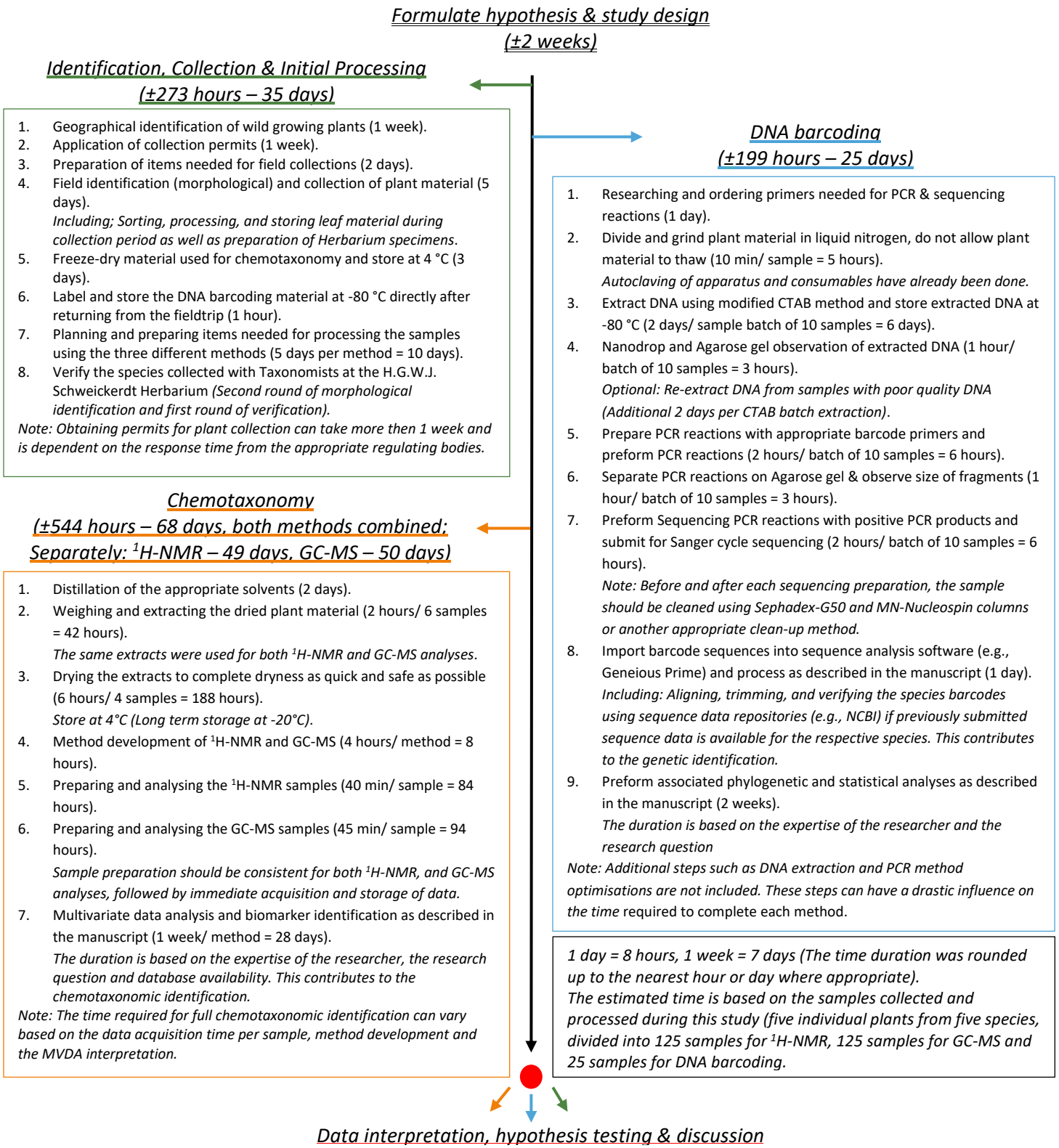

Fig. S. 2. Flow diagram describing the main experimental and analysis steps used, and the average time required (in an ideal scenario) to complete each section.

Table S. 1. The collection site, H.G.W.J. Schweickerdt Herbarium, University of Pretoria (PRU) numbers and the National Centre for Biotechnology Information (NCBI, GenBank) accession numbers of the plant species assessed.

Plant taxon name (Tree #)	**Collection sites**	**PRU numbers**	**GenBank accession numbers***
***Erythroxylum delagoense* Schinz**			
(Tree 1)	Enseleni Nature Reserve	122499	MT476106 (*matK*); MT476131 (*rbcL*); MT476167 (*trnH-psbA*)
(Tree 2)	Enseleni Nature Reserve	122500	MT476107 (*matK*); MT476132 (*rbcL*); MT476168 (*trnH-psbA*);
(Tree 3)	Enseleni Nature Reserve	122501	MT476108 (*matK*); MT476133 (*rbcL*); MT476169 (*trnH-psbA*)
(Tree 4)	Enseleni Nature Reserve	122502	MT476109 (*matK*); MT476134 (*rbcL*); MT476170 (*trnH-psbA*)
(Tree 5)	Enseleni Nature Reserve	122503	MT476110 (*matK*); MT476135 (*rbcL*); MT476171 (*trnH-psbA*)

Table S. 1. (continued)

***Erythroxylum emarginatum* Thonn.**			
(Tree 1)	Mpenjati Nature Reserve	122484	MT476111 (*matK*); MT476136 (*rbcL*); MT476172 (*trnH-psbA*)
(Tree 2)	Mpenjati Nature Reserve	122485	MT476112 (*matK*); MT476137 (*rbcL*); MT476173 (*trnH-psbA*)
(Tree 3)	Mpenjati Nature Reserve	122485	MT476113 (*matK*); MT476138 (*rbcL*); MT476174 (*trnH-psbA*)
(Tree 4)	Mpenjati Nature Reserve	122486	MT476114 (*matK*); MT476139 (*rbcL*); MT476175 (*trnH-psbA*)
(Tree 5)	Mpenjati Nature Reserve	122487	MT476115 (*matK*); MT476140 (*rbcL*); MT476176 (*trnH-psbA*)

Table S. 1. (continued)

***Erythroxylum pictum* E.Mey. ex Harv. & Sond.**			
(Tree 1)	Umtamvuna Nature Reserve	122479	MT476116 (*matK*); MT476141 (*rbcL*); MT476177 (*trnH-psbA*)
(Tree 2)	Umtamvuna Nature Reserve	122480	MT476117 (*matK*); MT476142 (*rbcL*); MT476178 (*trnH-psbA*)
(Tree 3)	Umtamvuna Nature Reserve	122481	MT476118 (*matK*); MT476143 (*rbcL*); MT476179 (*trnH-psbA*)
(Tree 4)	Umtamvuna Nature Reserve	122482	MT476119 (*matK*); MT476144 (*rbcL*); MT476180 (*trnH-psbA*)
(Tree 5)	Umtamvuna Nature Reserve	122483	MT476120 (*matK*); MT476145 (*rbcL*); MT476181 (*trnH-psbA*)

Table S. 1. (continued)

***Nectaropetalum capense* (Bolus) Stapf & Boodle**			
(Tree 1)	Umtamvuna Nature Reserve	122494	MT476121 (*matK*); MT476146 (*rbcL*); MT476182 (*trnH-psbA*)
(Tree 2)	Umtamvuna Nature Reserve	122495	MT476122 (*matK*); MT476147 (*rbcL*); MT476183 (*trnH-psbA*)
(Tree 3)	Umtamvuna Nature Reserve	122495	MT476123 (*matK*); MT476148 (*rbcL*); MT476184 (*trnH-psbA*)
(Tree 4)	Umtamvuna Nature Reserve	122497	MT476124 (*matK*); MT476149 (*rbcL*); MT476185 (*trnH-psbA*)
(Tree 5)	Umtamvuna Nature Reserve	122498	MT476125 (*matK*); MT476150 (*rbcL*); MT476186 (*trnH-psbA*)

Table S. 1. (continued)

***Nectaropetalum zuluense* (Schönland) Corbishley**			
(Tree 1)	Umtamvuna Nature Reserve	122489	MT476126 (*matK*); MT476151 (*rbcL*); MT476187 (*trnH-psbA*)
(Tree 2)	Umtamvuna Nature Reserve	122490	MT476127 (*matK*); MT476152 (*rbcL*); MT476188 (*trnH-psbA*)
(Tree 3)	Umtamvuna Nature Reserve	122491	MT476128 (*matK*); MT476153 (*rbcL*); MT476189 (*trnH-psbA*)
(Tree 4)	Umtamvuna Nature Reserve	122492	MT476129 (*matK*); MT476154 (*rbcL*); MT476190 (*trnH-psbA*)
(Tree 5)	Umtamvuna Nature Reserve	122493	MT476130 (*matK*); MT476155 (*rbcL*); MT476191 (*trnH-psbA*)

*Each GenBank accession number is associated to a specific DNA barcode region i.e., *matK* – maturase K; *rbcL* – ribulose-1,5-bisphosphate carboxylase/ oxygenase large subunit; *trnH-psbA* – intergenic spacer region.

Table S. 2. The DNA primer sequences used in this study.

DNA barcode	Forward primer (5’ to 3’)	Reverse primer (5’ to 3’)	Reference
rbcL	ATGTCACCACAAACAGAAAC	TCGCATGTACCTGCAGTAGC	(Kress et al., 2009)
matK	TGAATGGGCTAGTTTCCGGT	CGATTCCGGCAATTGCTCAA	(Primer 3, version 4.1.0)
trnH-psbA	CGCGCATGGTGGATTCACAATCC	CGAAGCTCCATCTACAAATGG	(Tate and Simpson, 2003)

Table S. 3. Summary of the financial cost comparison between different analytical techniques used in plant identification.

Instrumentation	Average full-service* price for 10 samples at 30 min/ sample (South African Rand)**	Average half-service* price for 10 samples at 30 min/ sample (South African Rand)**
^{1}H-NMR	I – ZAR 5124.60	I – ZAR 5124.60
	A – ZAR 10638.00	A – ZAR 8684.10
	C – ZAR 21165.40	C – ZAR 10047.30
GC-MS	I – ZAR 2683.00	I – ZAR 2683.00
	A – ZAR 12195.10	A – ZAR 8938.80
	C – ZAR 17772.08	C – ZAR 12734.00
DNA barcoding (Sanger Cycle Sequencing	I – ZAR 4860.70	I – ZAR 4860.70
	A – ZAR 8938.80	A – ZAR 7567.10
	C – ZAR 12447.00	C – ZAR 12447.00
HPLC	A – ZAR 5203.30	A – ZAR 3671.60
	C – ZAR 6158.40	C – ZAR 4081.80
LC-MS	A – ZAR 12951.62	A – ZAR 9084.40
	C – ZAR 16781.42	C – ZAR 11924.70
MALDI-MS	A – ZAR 7216.05	A – ZAR 4990.80
	C – ZAR 10267.05	C – ZAR 7102.10
UV-Vis	A – ZAR 3880.60	A – ZAR 3052.10
	C – ZAR 4617.10	C – ZAR 3788.60
ESI-MS	A – ZAR 8354.90	A – ZAR 5151.10
	C – ZAR 11485.00	C – ZAR 8281.20

[A] I – In-house rates, A – Academic rates, C – Commercial rates

[B] Refer to the link under 'Additional data for downloading' for the full financial cost evaluation.

*Full-service prices include the cost of initial data interpretation or report generation by the analysis facility. Half-service represents the cost of data generation without initial interpretation or report generation by the analysis facility.

**The cost of analysis is represented in South African Rand (ZAR) and calculated based on the exchange rates at the time of reporting (USD/ZAR = 14.73, EUR/ZAR = 17.37, GBP/ZAR = 20.60).

Note: These prices are intended for comparative purposes and are subject to change based on the financial exchange rates, availability of instrumentation, sample number, analysis time and the professional relationship between the researcher and analysis facility. We strongly urge any interested persons to contact the analysis facility directly for a detailed cost evaluation based on the research project.