diversity

MDPI

Article

A Study of Phenolic Compounds and Their Chemophenetic Value in the Genus *Thesium* (Santalaceae)

Natasha Lombard [1,2,*], Maria A. Stander [3], Herman Redelinghuys [4], Margaretha Marianne Le Roux [1,2] and Ben-Erik Van Wyk [2]

1 Foundational Biodiversity Science Division, National Herbarium, South African National Biodiversity Institute, Private Bag X101, Pretoria 0001, South Africa; m.leroux@sanbi.org.za
2 Department of Botany and Plant Biotechnology, University of Johannesburg, Johannesburg 2006, South Africa; bevanwyk@uj.ac.za
3 Department of Biochemistry, University of Stellenbosch, Private Bag X1, Stellenbosch 7600, South Africa; lcms@sun.ac.za
4 Centre for Research on Evaluation, Science and Technology, University of Stellenbosch, Private Bag X1, Stellenbosch 7600, South Africa; hredelinghuys@sun.ac.za
* Correspondence: n.lombard@sanbi.org.za

Abstract: Despite the common use of *Thesium* species as food and medicine and the prevailing systematic (taxonomic) difficulties of the genus, the diversity of phenolic compounds and their chemophenetic value remain largely unknown. As part of ongoing systematic research on *Thesium*, phenolic compounds and their composition were investigated at four taxonomic ranks: generic, infrageneric, interspecific and infraspecific. Liquid chromatography–mass spectrometry, multivariate analyses, comparisons to DNA data and manual investigations of total ion chromatograms were conducted for 156 samples of 50 *Thesium* species, as well the two monotypic genera sister to *Thesium* (*Lacomucinea* and *Osyridicarpos*). A wide diversity of phenolic compounds, primarily flavonols, carboxylic acids, phenolic acids and associated derivatives, were observed. Rutin was the most common compound, followed by citric acid, isorhamnetin O-glucoside O-rhamnoside, kaempferol O-rutinoside, quinic acid and cryptochlorogenic acid. Chemophenetic data proved valuable at the generic and interspecific level and, to a lesser degree, at the infraspecific level. On the other hand, no distinct patterns were observed at the infrageneric level, nor did chemophenetic data correlate with infrageneric classifications based on genetic, geographical or morphological data. At the generic level, the patterns of phenolic compounds in *Lacomucinea* and *Osyridicarpos* overlapped with those of *Thesium* species, but no compound or group of compounds were unique to *Thesium*. At the interspecific level, total ion chromatograms of the species were largely distinct from one another and, with the exception of four species, remarkably consistent. Two related species (*T. triflorum* and *T. scandens*) with doubtful species boundaries had similar but distinct chromatograms, providing evidence of their retention as separate species. At the infraspecific level, the data were mostly inconclusive, but variation in samples of *T. asterias* revealed two morphologically distinct populations, one of which is possibly a species new to science. This study provides valuable preliminary insights into the phenolic diversity and chemophenetic relationships in *Thesium*, *Lacomucinaea* and *Osyridicarpos* and highlights the potential of chemophenetics in taxonomically problematic plant groups.

Keywords: flavonoids; *Lacomucinaea*; LCMS analysis; *Osyridicarpos*; systematics; Thesiaceae

Citation: Lombard, N.; Stander, M.A.; Redelinghuys, H.; Le Roux, M.M.; Van Wyk, B.-E. A Study of Phenolic Compounds and Their Chemophenetic Value in the Genus *Thesium* (Santalaceae). *Diversity* **2022**, *14*, 590. https://doi.org/10.3390/d14080590

Academic Editor: Jesús Fernando Ayala-Zavala

Received: 21 June 2022
Accepted: 20 July 2022
Published: 23 July 2022

1. Introduction

Thesium L. is the largest genus [1] in the parasitic plant family Santalaceae [2], with ±350 species occurring naturally in Africa, Madagascar, Australia, Asia, Europe and South America, and one species being introduced to North America [3]. The centre of diversity for the genus is in southern Africa where ±175 species occur [3]. *Thesium* is monophyletic

and sister to *Lacomucinaea* Nickrent and M.A.García plus *Osyridicarpos* A.DC., two monotypic genera found in South Africa, and southern and tropical Africa, respectively [1,4]. *Thesium* plants are root hemiparasites that usually grow as herbs or subshrubs. They have linear or scale-like leaves that lack petioles and dry, nut-like fruits that often form elaiosomes [1,5,6]. Several species of *Thesium* have traditional and contemporary uses, particularly as medicines and functional foods in Africa and Asia [3].

Thesium has a long history of different circumscriptions and infrageneric classifications (see Ref [7]), since it was described by Linnaeus in 1753 [8]. Prior to a recent molecular study by Zhigila et al. (2020) [7], infrageneric groupings were based predominantly on morphological characters, and the proposed infrageneric classification systems (subgenera, sections, series, etc.) were geographically biased (e.g., Refs [9–12]). Zhigila et al. (2020) [7] presented a revised infrageneric classification for *Thesium* based on molecular data from one nuclear and three plastid gene regions, for 160 samples representing 116 *Thesium* species. They recognised five subgenera within *Thesium* based on five well-supported phylogenetic clades: Subgenus *Hagnothesium* (A.DC.) Zhigila, Verboom and Muasya, Subgenus *Thesium*, Subgenus *Discothesium* (A.DC.) Zhigila, Verboom and Muasya, Subgenus *Psilothesium* (A.DC.) Zhigila, Verboom and Muasya and Subgenus *Frisea* (Rchb.) Hendrych [7]. These five subgenera are supported by a strong geographical component, with Subgenus *Thesium* occurring mainly in Eurasia, Subgenus *Psilothesium* in tropical Africa and South America, and Subgenera *Hagnothesium*, *Discothesium* and *Frisea* in South Africa. The molecular phylogeny also confirmed that previous groupings based on morphology, particularly those proposed by Hill (1915) [9] for South African species, were artificial. However, although some clades, such as Subgenus *Hagnothesium*, and, to a degree, Subgenus *Thesium* and Subgenus *Discothesium*, have clear morphological apomorphies, the prevalence of homoplasious characters within the genus complicates the morphological diagnoses of some clades and subclades proposed by Zhigila et al. (2020) [7]. Furthermore, while their study presented strong molecular support for the five main clades, some internal nodes were unresolved and showed incongruent placements between nuclear and plastid phylogenetic trees [7]. This incongruence was especially prominent in the largest clade, Subgenus *Frisea* (clade 5), which consists of ±103 southern African *Thesium* species [7]. It is clear that although broad groupings within *Thesium* have largely been resolved, uncertainty remains regarding the placement of numerous species, as well as several natural groups within *Thesium*. Furthermore, some molecular clades lack clear diagnostic morphological characters (synapomorphies) to inform taxonomic revisions, and the placement of species not included in any molecular studies to date remains uncertain. Additional non-molecular and non-morphological data, such as information on patterns of variation in phytochemical compounds, as reported by Stander et al. (2019) [13] in a few *Thesium* species used as Cape herbal teas, might shed light on several remaining questions on the systematics of *Thesium* and provide additional evidence for the current groupings.

Chemophenetic information on *Thesium* is limited. The only chemophenetic report for *Thesium* was completed by De Kock and Rapson in 1938 [14] on 17 *Thesium* species endemic to the Greater Cape Floristic Region (GCFR) in South Africa. They showed that the presence or absence of an unidentified phlobatannin and/or an unidentified volatile oil glycoside in *Thesium* species corresponded with infrageneric sections of the time that were proposed by Hill (1915) [9] based on floral morphology. Species from Section *Imberbia* A.W.Hill contained both the phlobatannin and the volatile oil glycoside; species from Section *Annulata* A.W.Hill contained neither of the compounds; and species from Section *Barbata* A.W.Hill contained only the phlobatannin, only the volatile oil glycoside, both compounds or neither compounds [9,14]. While their results show some phytochemical support for the infrageneric sections of the time, several recent molecular studies have not only shown that Hill's sections are paraphyletic [4,7,15] but also that the species grouped together by De Kock and Rapson (1938) [14] are paraphyletic.

As with chemophenetics, knowledge on the phytochemical compounds found in *Thesium* is limited. In a recent literature review, Lombard et al. (2020) [3] found that the

phytochemical compounds of only eight *Thesium* species have been studied. A total of 70 phytochemical compounds (mainly phenolics, fatty acids and alkaloids) were isolated from these eight species. It should be noted that 60 of the 70 phytochemical compounds were isolated from only two species (*T. chinense* Turcz. from Asia and *T. humile* Vahl from the Mediterranean area), and the phytochemical composition of ±340 species remains unknown [3]. This limited knowledge is surprising, considering that *Thesium* species are used in ±40 functional foods, as medicine to treat ±130 ailments and its reported toxicity in some species [3].

Given (1) the remaining knowledge gaps in species relationships and natural groups within *Thesium*, (2) the need for apomorphic characters that support molecular clades, (3) the paucity of information on phytochemical composition patterns in *Thesium* and (4) the outdated nature of current chemophenetic research on the genus, an exploratory chemophenetic study of *Thesium* is warranted. An investigation of phenolic compounds was a logical starting point for such an exploratory study as phenolics; in particular, flavonoids are the most common compounds found thus far in *Thesium* [3], especially those traditionally used as herbal teas. The aims of this study, as part of ongoing systematic studies of *Thesium*, were to (1) conduct an exploratory investigation on the phenolic compounds present in *Thesium* and to (2) investigate the chemophenetic value of patterns in phenolic compound variation in *Thesium* at the generic level, infrageneric level, interspecific level and infraspecific level. This study provides the first account of phenolic variation in *Thesium*, *Lacomucinaea* and *Osyridicarpos*, including the first phytochemical accounts of 49 *Thesium* species and *O. schimperianus*.

2. Materials and Methods

2.1. Plant Material Used

A total of 156 samples of 50 *Thesium* species from Africa, Madagascar, Europe and Asia were included in this study. All five subgenera within *Thesium* are represented: 1 species (of 8) from Subgenus *Hagnothesium*, 3 species (of ±65) from Subgenus *Thesium*, 2 species (of 13) from Subgenus *Discothesium*, 21 species (of ±98) from Subgenus *Psilothesium* and 23 species (of ±103) from Subgenus *Frisea*. Of the 156 *Thesium* samples (50 species) included here, 142 samples (44 species) were collected by the authors from live plants at several sites across South Africa during the summer seasons of 2016 to 2019 and air dried before extraction. Where possible, multiple populations (one to three) of a species were sampled and multiple samples (one to five) collected from each population. Each species is therefore represented by between one and nine samples. A further six air-dried samples of two Eurasian species, *T. ebracteatum* Hayne and *T. ramosum* Hayne, were obtained from collections made by Zigmantas Gudzinskas in Lithuania. Eight samples of four species (*T. chinense* from Asia, *T. cymosum* A.W.Hill and *T.* cf. *ussanguense* Engl. from southern and eastern Africa, and *T.* cf. *leandrianum* Cavaco and Keraudren from Madagascar) were also taken from herbarium specimens. Every attempt was made to include samples from the full geographical range and morphological variation of the genus.

The two monotypic genera that are sister to *Thesium*, *Lacomucinaea* and *Osyridicarpos* [1] were also investigated using three and one samples, respectively. A list of all samples with corresponding voucher information is provided in Table 1.

2.2. Extraction

Samples consisted of above-ground parts (including leaves, stems, flowers and fruits) of plants and were finely ground using a mortar and pestle. Approximately 0.2 g of each sample was extracted using 50% methanol in water containing 1% formic acid (1.5 mL) by soaking it overnight. This was followed by extraction in an ultrasonic bath (0.5 Hz, Integral Systems, South Africa) for 60 min at room temperature. The extracts were lastly centrifuged (Hermle Z160 m, 3000× *g* for 5 min).

Table 1. List of samples used in this study, including their species names, subgenera, sample codes, countries of origin, collection localities and dates, and voucher details. Vouchers are housed in the Pretoria National Herbarium, South Africa (PRE). Samples included in multivariate statistics are indicated with an "*".

Sample No.	Species	Subgenus	Sample Code	Collection Details			
				Country	Locality	Date	Voucher
1 2 3	*Thesium acuminatum* A.W.Hill	*Frisea* (Rchb.) Hendrych	1_T.acumi_A1 * 2_T.acumi_A2 * 3_T.acumi_A3 *	South Africa	Western Cape, Silvermine	2018/09/13	*Visser and van Wyk 305*
4	*Thesium acutissimum* A.DC.	*Psilothesium* (A.DC.) Zhigila, Verboom and Muasya	4_T.acuti_A1 *	South Africa	Eastern Cape, Grahamstown area	2018/11/04	*Visser et al. 369*
5			5_T.acuti_B1 *	South Africa	Eastern Cape, Grahamstown	2018/11/04	*Visser et al. 370*
6 7 8 9 10	*Thesium asterias* A.W.Hill	*Psilothesium*	6_T.aster_A1 7_T.aster_A2 8_T.aster_A3 9_T.aster_A4 10_T.aster_A5	South Africa	KwaZulu-Natal, Cape Vidal	2018/02/02	*van Wyk s.n.*
11 12 13			11_T.aster_B1 * 12_T.aster_B2 * 13_T.aster_B3 *	South Africa	Limpopo, Lekgalameetse Nature Reserve	2017/12/05	*Visser and le Roux 291*
14 15 16	*Thesium carinatum* A.DC.	*Frisea*	14_T.carin_A1 15_T.carin_A2 16_T.carin_A3	South Africa	Western Cape	2017	*van Wyk s.n.*
17 18	*Thesium chinense* Turcz. var. *longipedunculatum* Chu	*Thesium*	31_T.chine_A1 32_T.chine_A2	China	-	2004/05/04	*Guo 1506034*
19 20	*Thesium chinense* Turcz. var. *chinense*	*Thesium*	34_T.chine_B1 * 35_T.chine_B2 *	Japan	Honshu, Tsukuba City	1997/05/07	*Konta and Momose 18144*
21 22 23	*Thesium commutatum* Sond.	*Frisea*	36_T.commu_A1 * 37_T.commu_A2 * 38_T.commu_A3 *	South Africa	Western Cape, Silvermine	2018/09/13	*Visser and van Wyk 304*
24 25 26	*Thesium confine* Sond.	*Psilothesium*	39_T.confi_A1 * 40_T.confi_A2 * 41_T.confi_A3 *	South Africa	Eastern Cape, Middelburg area	2016/10/26	*Visser and le Roux 208*

Table 1. *Cont.*

Sample No.	Species	Subgenus	Sample Code	Collection Details			
				Country	Locality	Date	Voucher
27 28 29 30 31 32	*Thesium costatum* A.W.Hill	*Psilothesium*	42_T.costa_A1 43_T.costa_A2 44_T.costa_A3 48_T.costa_C1 * 49_T.costa_C2 * 50_T.costa_C3 *	South Africa South Africa	Gauteng, Pretoria National Botanical Garden Limpopo, Haenertsburg	2016/10/18 2017/12/05	*le Roux 184* *Visser and le Roux 293*
33 34 35	*Thesium cupressoides* A.W.Hill	*Psilothesium*	51_T.cupre_A1 52_T.cupre_B1 53_T.cupre_C1	South Africa South Africa South Africa	KwaZulu-Natal, Sani Pass KwaZulu-Natal, Umtamvuna Nature Reserve area KwaZulu-Natal, Umtamvuna Nature Reserve	2017/01/31 2017/02/01 2017/02/02	*Visser et al. 252* *Visser et al. 253* *Visser et al. 262*
36	*Thesium cymosum* A.W.Hill	*Psilothesium*	30_T.cymos_A1	Mozambique	Manica, Mavita	1944/10/25	*Mendonça 2565*
37 38	*Thesium davidsoniae* Brenan	*Psilothesium*	54_T.david_A1 * 55_T.david_A2 *	South Africa	Limpopo, Abel Erasmus Pass	2017/12/04	*Visser and le Roux 289*
39 40 41 42 43 44	*Thesium durum* Hilliard and B.L.Burtt	*Frisea*	56_T.durum_A1 * 57_T.durum_A2 * 58_T.durum_A3 * 59_T.durum_B1 * 60_T.durum_B2 * 61_T.durum_B3 *	South Africa South Africa	Eastern Cape, Lady Grey Eastern Cape, Rossouw-Indwe area	2018/10/30 2018/11/02	*Visser et al. 352* *Visser et al. 362*
45 46 47	*Thesium ebracteatum* Hayne	*Thesium*	62_T.ebrac_A1 * 63_T.ebrac_A2 * 64_T.ebrac_A3 *	Lithuania	-	2018/06	*Gudzinskas s.n.*
48	*Thesium* cf. *elatius* Sond.	*Frisea*	17_T.elati_A1 *	South Africa	Western Cape, Betty's Bay	2018/09/14	*Visser et al. 312*
49 50 51	*Thesium euphorbioides* P.J.Bergius	*Frisea*	65_T.eupho_A1 * 66_T.eupho_A2 * 67_T.eupho_A3 *	South Africa	Western Cape, Fernkloof Nature Reserve	2018/09/14	*Visser et al. 316*

Table 1. *Cont.*

Sample No.	Species	Subgenus	Sample Code	Collection Details			
				Country	Locality	Date	Voucher
52 53 54	*Thesium euphrasioides* A.DC.	*Frisea*	68_T.euphr_A1 * 69_T.euphr_A2 * 70_T.euphr_A3 *	South Africa	Western Cape, Franschoek area	2018/09/12	*Visser and van Wyk 300*
55 56 57 58 59 60	*Thesium flexuosum* A.DC.	*Frisea*	71_T.flexu_A1 * 72_T.flexu_A2 * 73_T.flexu_A3 * 74_T.flexu_B1 * 75_T.flexu_B2 * 76_T.flexu_B3 *	South Africa	Eastern Cape, Hankey area	2018/12/30	*Visser 379*
				South Africa	Eastern Cape, Grahamstown area	2018/11/04	*Visser et al. 368*
61 62 63	*Thesium foliosum* A.DC.	*Frisea*	77_T.folio_A1 * 78_T.folio_A2 * 79_T.folio_A3 *	South Africa	Eastern Cape, Van Stadens Wild Flower Local Nature Reserve	2018/12/30	*Visser 380*
64	*Thesium* cf. *frisea* L.	*Frisea*	18_T.frise_A1 *	South Africa	Western Cape, Betty's Bay	2018/09/14	*Visser et al. 310*
65	*Thesium fruticosum* A.W.Hill	*Frisea*	80_T.fruti_A1 *	South Africa	Eastern Cape, Grahamstown	2018/11/04	*Visser et al. 371*
66 67	*Thesium* cf. *gnidiaceum* A.DC.	*Psilothesium*	19_T.gnidi_A1 * 20_T.gnidi_A2 * 21_T.gnidi_A3 *	South Africa	Eastern Cape, Elliot area	2018/11/01	*Visser et al. 360*
69	*Thesium gypsophiloides* A.W.Hill	*Psilothesium*	82_T.gypso_A1 *	South Africa	KwaZulu-Natal, Vernon Crookes Nature Reserve	2017/02/03	*Visser et al. 269*
70 71 72	*Thesium hispidulum* Lam.	*Frisea*	83_T.hispi_A1 * 84_T.hispi_A2 * 85_T.hispi_A3 *	South Africa	Western Cape, Central Cederberg	2018/09/17	*Visser and van Wyk 337*
73 74 75			86_T.hispi_B1 * 87_T.hispi_B2 * 88_T.hispi_B3 *	South Africa	Western Cape, Piekerniekloof Pass	2018/09/16	*Visser and van Wyk 329*
76 77 78			89_T.hispi_C1 * 90_T.hispi_C2 * 91_T.hispi_C3 *	South Africa	Western Cape, Bain's Kloof Pass	2018/09/16	*Visser and van Wyk 327*

Table 1. *Cont.*

Sample No.	Species	Subgenus	Sample Code	Collection Details			
				Country	Locality	Date	Voucher
79 80 81 82 83 84	*Thesium imbricatum* Thunb.	*Frisea*	92_T.imbri_A1 * 93_T.imbri_A2 * 94_T.imbri_A3 * 95_T.imbri_B1 * 96_T.imbri_B2 * 97_T.imbri_B3 *	South Africa South Africa	Eastern Cape, Rossouw-Indwe area Eastern Cape, Maclear area	2018/11/02 2018/10/31	*Visser et al. 363* *Visser et al. 357*
85 86	*Thesium impeditum* A.W.Hill	*Psilothesium*	98_T.imped_A1 * 115_T.imped_B1 *	South Africa South Africa	Eastern Cape, Hogsback area Free State, Frankfort area	2018/11/03 2016/12/12	*Visser et al. 366* *Visser and le Roux 233*
87	*Thesium junceum* Bernh.	*Frisea*	99_T.junce_A1 *	South Africa	Eastern Cape, Maclear area	2018/10/31	*Visser et al. 356*
88 89 90	*Thesium karooicum* Compton	*Frisea*	100_T.karoo_A1 * 101_T.karoo_A2 * 102_T.karoo_A3 *	South Africa	Western Cape, Montagu	2018/09/15	*Visser et al. 318*
91	*Thesium* cf. *leandrianum* Cavaco and Keraudren	*Psilothesium*	22_T.leand_A1 *	Madagascar	Marojejy National Park	1989/02/15	*Miller and Lowry 4171*
92 93	*Thesium magalismontanum* Sond.	*Psilothesium*	108_T.magal_A1 * 109_T.magal_A2 *	South Africa	Gauteng, Dinokeng Nature Reserve	2017/12/12	*Nel 482*
94 95 96	*Thesium microcarpum* A.DC.	*Hagnothesium* (A.DC.) Zhigila, Verboom and Muasya	110_T.micro_A1 * 111_T.micro_A2 * 112_T.micro_A3 *	South Africa	Western Cape, Op-de-Tradouw	2018/09/15	*Visser et al. 323*
97 98	*Thesium multiramulosum* Pilg.	*Psilothesium*	113_T.multi_A1 * 114_T.multi_A2 *	South Africa	Limpopo, Uitsig	2017/12/04	*Visser and le Roux 282*
99 100 101	*Thesium* cf. *nigromontanum* Sond.	*Frisea*	23_T.nigro_A1 * 24_T.nigro_A2 * 25_T.nigro_A3 *	South Africa	Western Cape, Franschoek Pass	2018/09/12	*Visser and van Wyk 301*
102 103	*Thesium ovatifolium* N.Lombard and M.M.leRoux	*Psilothesium*	116_T.ovati_A1 * 117_T.ovati_A2 *	South Africa	KwaZulu-Natal, Ngome area	2018/10/14	*Visser and Lombard 346*

Table 1. *Cont.*

Sample No.	Species	Subgenus	Sample Code	Collection Details			
				Country	Locality	Date	Voucher
104 105	*Thesium pallidum* A.DC.	*Psilothesium*	118_T.palli_A1 * 119_T.palli_A2 *	South Africa	Free State, Platberg Nature Reserve	2016/12/14	*Visser and le Roux 239*
106			120_T.palli_B1 *	South Africa	Eastern Cape, Maclear area	2018/11/01	*Visser et al. 359*
107 108 109	*Thesium procerum* N.E.Br.	*Psilothesium*	121_T.proce_A1 122_T.proce_A2 123_T.proce_A3	South Africa	Gauteng, Pretoria National Botanical Garden	2017/10/13	*Visser 276*
110 111	*Thesium pubescens* A.DC.	*Frisea*	124_T.pubes_A1 * 125_T.pubes_A2 *	South Africa	Western Cape, Citrusdal area	2018/09/16	*Visser and van Wyk 331*
112	*Thesium racemosum* Bernh.	*Psilothesium*	126_T.racem_A1 *	South Africa	Limpopo, Woodbush Forest Reserve area	2017/12/04	*Visser and le Roux 288*
113 114 115	*Thesium ramosum* Hayne	*Thesium*	127_T.ramos_A1 * 128_T.ramos_A2 * 129_T.ramos_A3 *	Lithuania	-	2018/06	*Gudzinskas s.n.*
116	*Thesium resedoides* A.W.Hill	*Psilothesium*	130_T.resed_A1 *	South Africa	Gauteng, Dinokeng Nature Reserve	2017/12/12	*Nel 481*
117 118 119	*Thesium scandens* Sond.	*Discothesium* (A.DC.) Zhigila, Verboom and Muasya	131_T.scand_A1 * 132_T.scand_A2 * 133_T.scand_A3 *	South Africa	Eastern Cape, Paterson area	2018/11/04	*Visser et al. 373*
120			134_T.scand_B1 *	South Africa	Eastern Cape, Baviaan River Conservancy	2018/11/05	*Visser et al. 376*
121	*Thesium scirpioides* A.W.Hill	*Frisea*	135_T.scirp_A1 *	South Africa	Gauteng, Suikerbosrand Nature Reserve	2016/12/12	*Visser and le Roux 232*
122 123 124	*Thesium sonderianum* Schltr.	*Frisea*	136_T.sonde_A1 * 137_T.sonde_A2 * 138_T.sonde_A3 *	South Africa	Eastern Cape, Nieu-Bethesda area	2018/11/06	*Visser et al. 378*

Table 1. *Cont.*

Sample No.	Species	Subgenus	Sample Code	Collection Details			
				Country	Locality	Date	Voucher
125	*Thesium spicatum* L.	*Frisea*	103_T.spica_A1	South Africa	Western Cape	2017	*van Wyk s.n.*
126			104_T.spica_A2				
127			105_T.spica_A3				
128			106_T.spica_A4				
129			107_T.spica_A5				
130	*Thesium strictum* P.J.Bergius	*Frisea*	139_T.stric_A1	South Africa	Western Cape	2017	*van Wyk s.n.*
131			140_T.stric_A2				
132			141_T.stric_A3				
133			142_T.stric_B1 *	South Africa	Western Cape, Tradouw Pass	2018/09/15	*Visser et al. 325*
134			143_T.stric_B2 *				
135			144_T.stric_B3 *				
136			145_T.stric_C1 *	South Africa	Western Cape, Rooi-Els	2018/09/14	*Visser et al. 309*
137			146_T.stric_C2 *				
138			147_T.stric_C3 *				
139	*Thesium transvaalense* Schltr.	*Psilothesium*	148_T.trans_A1 *	South Africa	Gauteng, Waverley	2016/10/22	*Visser and le Roux 201*
140			149_T.trans_B1 *	South Africa	Gauteng, Waverley	2016/10/22	*Visser and le Roux 200*
141			150_T.trans_B2 *				
142	*Thesium triflorum* L.f.	*Discothesium*	26_T.trifl_A1 *	South Africa	KwaZulu-Natal, St. Lucia	2018/02/02	*van Wyk s.n.*
143	*Thesium* cf. *ussanguense* Engl.	*Frisea*	27_T.ussan_A1 *	Mozambique	Manica, Serra Zuira	1965/11/09	*Torre and Pereira 12.809*
144			29_T.ussan_A3 *				
145	*Thesium vahrmeijeri* Brenan	*Psilothesium*	415_T.vahrm_A1 *	South Africa	KwaZulu-Natal, iSimangaliso Wetland Park	2020/01/14	*Lombard and le Roux 415*
146			418_T.vahrm_A2 *	South Africa	KwaZulu-Natal, iSimangaliso Wetland Park	2020/01/15	*Lombard and le Roux 418*
147	*Thesium virgatum* Lam.	*Frisea*	81_T.virgat_A1 *	South Africa	Western Cape, Franschoek area	2018/09/12	*Visser and van Wyk 299B*

Table 1. *Cont.*

Sample No.	Species	Subgenus	Sample Code	Collection Details			
				Country	Locality	Date	Voucher
148	*Thesium zeyheri* A.DC.	*Psilothesium*	151_T.zeyhr_A1 *	South Africa	Free State, Platberg Nature Reserve	2016/12/14	*Visser and le Roux 240*
149			152_T.zeyhr_A2 *				
150			153_T.zeyhr_A3 *				
151			154_T.zeyhr_B1 *	South Africa	Eastern Cape, Maclear area	2018/10/31	*Visser et al. 354*
152			155_T.zeyhr_B2 *				
153			156_T.zeyhr_B3 *				
154			157_T.zeyhr_C1 *	South Africa	Eastern Cape, Rossouw-Indwe area	2018/11/02	*Visser et al. 364*
155			158_T.zeyhr_C2 *				
156			159_T.zeyhr_C3 *				
157	*Osyridicarpos schimperianus* (Hochst. ex A.Rich.) A.DC.	-	Osyridicarpos_292 *	South Africa	Limpopo, Lekgalameetse Nature Reserve	2017/12/05	*Visser and le Roux 292*
158	*Lacomucinaea lineata* (L.f.) Nickrent and M.A.García	-	Lacomucinaea_394 *	South Africa	Northern Cape, Postmasburg area	2019/11/28	*Lombard and le Roux 394*
159			Lacomucinaea_401 *	South Africa	Northern Cape, Bergenaarspadpas area	2019/11/29	*Lombard and le Roux 401*
160			Lacomucinaea_405 *	South Africa	Northern Cape, Witsand area	2019/11/29	*Lombard and le Roux 405*

The uppercase letters (A, B and C) in the sample codes of samples 1–156 refer to different populations of a species and the numbers to different individuals (e.g., A1 = individual one of population one; C2 = individual two of population three).

2.3. Standards

Standards of nine compounds (flavonoids and organic acids), namely caffeic acid, catechin, citric acid, epicatechin, ferulic acid, gallic acid, *p*-coumaric acid, rutin and quinic acid were prepared at concentrations of 1.95 µg/mL, 3.9 µg/mL, 15.6 µg/mL, 31.25 µg/mL, 62.5 µg/mL, 125 µg/mL and 500 µg/mL with the same solvent used for sample extraction (50% methanol in water containing 1% formic acid).

2.4. Liquid Chromatography—Mass Spectrometry (LCMS)

High-resolution UPLC-MS analysis was conducted using a Waters Synapt G2 Quadrupole time-of-flight (QTOF) mass spectrometer (MS) and a Waters Acquity ultra-performance liquid chromatograph (UPLC) (Waters, Milford, MA, USA). Negative mode was used for electrospray ionisation with a cone voltage of 15 V, desolvation gas at 650 L/h and desolvation temperature of 275 °C. The other mass spectrometry settings were optimised for best resolution and sensitivity. Data were attained by scanning from m/z 150 to 1500 in the resolution mode and MS^E mode. In the MS^E mode, two channels of mass spectrometry data were collected: the first at low collision energy (4 V) and the second with a collision energy ramp (40–100 V) for fragmentation data. Accurate mass determination was performed with leucine enkephalin as a mass reference, and the instrument was calibrated with sodium formate. A Waters HSS T3, 2.1 × 100 mm, 1.7 µm column was used for separation. The injection volume was 2 µL, and the mobile phase consisted of 0.1% formic acid (solvent A), as well as acetonitrile with 0.1% formic acid (solvent B). After 1 min of 100% solvent A, the gradient changed linearly to 28% solvent B over 22 min, 40% solvent B over 50 s, followed by a wash step of 1.5 min at 100% solvent B and, finally, re-equilibration to the initial conditions for 4 min. The column temperature was sustained at 55 °C, and the flow rate was 0.3 mL/min.

2.5. Compound Identification

Provisional identifications of compounds are provided based on a combination of the following information: correlation with standards, data in previous publications on *Thesium*, as well as the fragmentation data, elemental composition (based on accurate mass), relative retention times and UV data of peaks. Online databases, including Chemspider (www.chemspider.com, accessed on 11 November 2020), KNApSAck (www.knapsackfamily.com/knapsack.core, accessed on 10 January 2022) and METLIN (metlin.scripps.edu, accessed on 18 January 2022), were used. Quinic acid derivatives showing an m/z 191 fragment ion, ferulic acids (m/z 193), coumaric acids (m/z 163) and combinations of these occurred in many samples. Flavonol bases (fragment ions of kaempferol m/z 285, quercetin m/z 300 and m/z 315 isorhamnetin) with different mono-and diglycosides (loss of m/z 162 for hexose, −146 for rhamnose, −308 for dihexoside) were also detected. Further identifications of observed compounds are beyond the scope of this study and may be addressed in future studies, especially in those species with interesting and as yet unidentified compounds.

2.6. Multivariate Analysis

To ensure accurate peak alignment, only samples with little retention time drift were included in the multivariate analysis. A total of 128 samples (46 species) fitted this criterion and are indicated with an "*" in Table 1. The remaining 28 samples were manually studied and compared. Raw mass spectrometry data were processed using the MarkerLynx XS function of MassLynx version 4.1 software (Waters Inc., Johannesburg, South Africa, 2012). The peak detection analysis was used to align peaks, as well as to convert raw data to retention time–mass pairs with the signal intensity for each peak. Only peaks between 4 and 25 min retention time were included, and the analysis was run with the following method parameters: low mass = 120, high mass = 1300, XIC window (Da) = 0.01, use relative retention time = no, peak width at 5% height (seconds) = 15, apply smoothing = yes, marker intensity threshold = 500, mass window = 0.05, retention time window = 0.80, noise elimination level = 200 and deisotope data = yes. Peak alignment was checked manually for accuracy. Pareto scaling was applied.

Generic, infrageneric, interspecific and infraspecific relationships were investigated and visualised using Python version 3.8 (Python Software Foundation, 2001–2022; https://www.python.org/downloads/, accessed on 8 December 2020) with the Scikit-learn 0.23.2 and Matplotlib 3.3 Libraries, as well as MetaboAnalyst version 5.0 software (Xia Lab, McGill University, 2020; https://www.metaboanalyst.ca/home.xhtml, accessed on 8 December 2020). In Python, the Markerlynx mass spectra data were firstly normalised to compensate for the variance in concentration and to ensure equal representation in the dataset, thereby facilitating comparative analysis. Normalisation involved scaling each sample vector to a unit norm, independently of other samples, so that all values were represented on a common scale. The dimensionality of the dataset was then reduced using principal component analysis (PCA). This was performed both to gain a better understanding of the important features in the dataset and to reduce dimensional complexity for further analysis steps without significant reduction in accuracy. The minimum number of PCA components to be evaluated were selected, so that the amount of variance that needed to be explained by this number of components was greater than two times the standard deviation (95.45%) of data coverage. In this case, 44 PCA components were sufficient to explain the variation in the data to more than two standard deviations, and by examining the loading factors for each of these PCA components, it is possible to visualise the causative factors, which contributed the most variation within the dataset. Next, unsupervised clustering analysis was preformed using the mean shift clustering algorithm (or the mode-seeking algorithm). This is a non-parametric kernel density estimation-based algorithm that attempts to seek the densest region of samples iteratively. The mean shift algorithm does not require specifying the number of clusters in advance; the number of clusters is conferenced by the data. The algorithm initially assigns each point in the feature space as a possible cluster centroid, and the distribution distances of points in the feature space are used to estimate the kernel bandwidth parameter, also known as the attractive interaction distance between samples. Iteratively, the mean of all additional points within the bandwidth of the initial centroids determines the new set of centroids (kernel density estimation) and assigns each data point to the closest cluster centroid until convergence is achieved. After convergence, the remaining cluster centres define the number of clusters, and the data points associated with the same centroid are members of the same cluster. The kernel density estimate bandwidth parameter significantly impacts the accuracy and precision performance of the mean shift algorithm. The bandwidth parameter was estimated for this feature space to be the average of the distance of each vector to its k-nearest neighbour (kNN), relative to the furthest distance between the feature vectors. Setting the number of k-nearest neighbours to 6–7% of the total sample set yielded optimal results.

In MetaboAnalyst, raw peak intensity data from MarkerLynx were normalised by sum, log transformed and scaled using automatic scaling, after which two- and three-dimensional PCA plots and hierarchical clustering dendrograms (using Euclidean distance and Ward clustering) were computed to show infraspecific relationships.

2.7. Phylogenetic Comparison

Generic, infrageneric and interspecific relationships were also investigated in a phylogenetic framework by plotting the main 33 compounds observed (see Table 2) in *Thesium* and *Lacomucinaea* (see Table 3) on the latest and most complete phylogeny [7]. A total of 27 *Thesium* species spanning all five subgenera overlapped between this study and the phylogeny, as well as *L. lineata*.

Table 2. The most common phenolic compounds and organic acids (peaks) detected in a total of 156 samples from 50 species of *Thesium*, along with their retention time, molecular mass, molecular formula, fragments, standards used and preliminary identifications with associated references. The number of samples and species in which each compound was observed at high levels (peak intensity > 30%) are also provided.

	Retention Time	*m*/*z* [M-H]$^-$	Formula	MS/MS	Standard	Identification	Reference	Number of Samples	Number of Species
F1	3.05	191.0195	$C_6H_7O_7$	170, **155**	Citric acid	Citric acid		38	26
F6	4.30	612.1365	$C_{26}H_{30}NO_{14}S$	559, 377, 325, **279**, 161, 97		Unknown sulphated compound		12	4
F11	5.99	389.1207	$C_{20}H_{21}O_8$	**165**, 183, 209		Unknown		9	6
F12	6.01	339.1283	$C_{14}H_{25}O_{12}$/$C_{13}H_{23}O_{10}$	**339**, 215, 193/193		Dehydrodiferulic acid		20	10
F16	6.32	215.0564	$C_9H_{11}O_6$	156, 155, 137, 111, 93		Trimethyl Aconitate		10	4
F18	6.65	191.0555	$C_7H_{11}O_6$	129, 115, **101**		Quinic acid		35	14
F20	6.86	527.2552	$C_{19}H_{43}O_{16}$	361, 309, 263, 191, 179, 119, 89		Unknown		22	5
F28	9.50	353.0873	$C_{16}H_{17}O_9$	**191**, 179, 135		Neochlorogenic acid (3-caffeoylquinic acid)	[16]	11	7
F33	10.01	173.0459	$C_7H_9O_5$	129, **111**, 85		Shikimic acid derivative		18	7
F41	11.08	337.0923	$C_{16}H_{17}O_8$	191, **163**, 119		Coumaroyl quinic acid		19	9
F49	11.67	353.0871	$C_{16}H_{17}O_9$	**191**		Chlorogenic acid (5-caffeoylquinic acid)		20	7
F50	11.90	353.0865	$C_{16}H_{17}O_9$	191, 179, 173, 135		Cryptochlorogenic acid (4-caffeoylquinic acid)	[16]	25	12
F58	12.39	367.1033	$C_{17}H_{19}O_9$	**193**, 134		Feruloylquinic acid		11	7
F78	13.78	337.0955	$C_{16}H_{17}O_8$	**191**, 173		Coumaroyl quinic acid		12	6
F90	15.05	367.1046	$C_{17}H_{19}O_9$	193, **173**		Feruloylquinic acid		7	6
F94	15.10	367.1024	$C_{17}H_{19}O_9$	191		Feruloylquinic acid		14	7
F100	15.49	755.2222	$C_{33}H_{39}O_{20}$	**300**, 271, 255		Quercetin rhamnosyl-rhamnosyl-glucoside		18	11
F111	16.28	609.1619	$C_{27}H_{29}O_{16}$	**300**, 271, 255		Quercetin robinobioside		18	10
F113	16.42	739.2306	$C_{30}H_{43}O_{21}$	**284**, 255, 227		Kaempferol O-rhamnosyl-rhamnosyl-hexoside		11	5
F122	17.20	609.1416	$C_{27}H_{29}O_{16}$	**300**, 271, 255	Rutin	Rutin		80	27

Table 2. *Cont.*

	Retention Time	*m*/*z* [M-H]$^-$	Formula	MS/MS	Standard	Identification	Reference	Number of Samples	Number of Species
F125	17.27	449.1227	$C_{21}H_{21}O_{11}$	**287**, 151, 135		Eriodictyol -O-glucoside		10	4
F133	17.55	1043.2959	$C_{48}H_{51}O_{26}$	**755**, 609, 447		Quercetin glycoside		11	4
F138	17.99	579.1522	$C_{23}H_{31}O_{17}$	**300**, 271, 255		Quercetin pentoside hexoside		11	5
F145	18.81	593.1492	$C_{27}H_{29}O_{15}$	**285**, 255		Kaempferol O-rutinoside	[17]	36	15
F154	19.28	355.1417	$C_{17}H_{23}O_{8}$	193, **161**		Ferulic acid derivative		12	6
F156	19.36	623.1616	$C_{28}H_{31}O_{16}$	**315**, 300, 271		Isorhamnetin O-glucoside O-rhamnoside		51	20
F173	21.46	933.269	$C_{43}H_{49}O_{23}$	**661**, **353**		Unknown		10	4
F177	22.11	933.2667	$C_{43}H_{49}O_{23}$	771, 661, **353**		Unknown		10	3
F193	23.93	301.0364	$C_{15}H_{9}O_{7}$	299, 271, **151**		Quercetin		11	5

Table 3. Comparison of the main phenolic compounds and organic acids (peak intensity >30%) detected in *Lacomucinaea*, *Osyridicarpos* and *Thesium*.

	Retention Time	*m/z* [M-H]$^-$	Formula	MS/MSc	*Lacomucinaea*	*Osyridicarpos*	*Thesium*	Identification	Reference
F1	3.05	191.0195	$C_6H_7O_7$	170, **155**	+	+	+	Citric acid	
F18	6.65	191.0555	$C_7H_{11}O_6$	129, 115, **101**			+	Quinic acid	
F25.4	8.76	463.1452	$C_{19}H_{27}O_{13}$	293, 191, 169, 149, 125, 89	+			Unknown	
F26	8.90	315.1080	$C_{14}H_{19}O_8$	**153**, 123		+	+	Hydroxytyrosol O-glucoside	
F28	9.50	353.0873	$C_{16}H_{17}O_9$	**191**, 179, 135	+	+	+	Neochlorogenic acid (3-caffeoylquinic acid)	[16]
F50	11.90	353.0865	$C_{16}H_{17}O_9$	191, 179, 173, 135	+	+	+	Cryptochlorogenic acid (4-caffeoylquinic acid)	[16]
F57	12.29	385.1160	$C_{17}H_{21}O_{10}$	**223**, 208, 179, 164	+		+	Unknown, possibly sinapoyl hexose	
F58	12.39	367.1033	$C_{17}H_{19}O_9$	**193**, 134	+	+	+	Feruloylquinic acid	
F78.5	13.81	769.1800	$C_{33}H_{37}O_{21}$	667, 625		+		Unknown	
F90	15.05	367.1046	$C_{17}H_{19}O_9$	193, **173**	+	+	+	Feruloylquinic acid	
F94.1	15.15	753.1843	$C_{40}H_{33}O_{15}$	**591**, 489, 447, 191, 179, 161		+		Unknown	
F94.2	15.15	625.1400	$C_{27}H_{29}O_{17}$	**367**, 300, 191	+			Quercetin-O-dihexoside	
F114.1	16.58	601.2300	$C_{31}H_{37}O_{12}$	**479**, 171	+	+		Unknown	
F121.2	17.15	639.1600	$C_{28}H_{31}O_{17}$	315	+			Unknown	
F122	17.20	609.1416	$C_{27}H_{29}O_{16}$	**300**, 271, 255			+	Rutin	
F136	17.67	463.0978	$C_{25}H_{19}O_9$	**300**, 271, 255	+		+	Quercetin-O-hexoside	
F145	18.81	593.1492	$C_{27}H_{29}O_{15}$	**285**, 255			+	Kaempferol O-rutinoside	[17]
F156	19.36	623.1616	$C_{28}H_{31}O_{16}$	**315**, 300, 271			+	Isorhamnetin O-glucoside O-rhamnoside	
F207	24.89	265.1429	$C_{15}H_{21}O_4$	**97**, 183	+		+	Unknown	

3. Results

3.1. Overview of Phenolic Compounds

More than 200 individual peaks were recorded during the manual examination of 156 total ion chromatograms of 50 *Thesium* species, *L. lineata* and *O. schimperianus* (all chromatograms are provided in Supplementary Figure S1). While the identification of all peaks was not the aim of this study, preliminary identifications and information on the ±30 most common peaks observed in *Thesium* are provided in Table 2 and those of *L. lineata* and *O. schimperianus* in Table 3. In all three genera, flavonols, carboxylic acids, phenolic acids, organic acids and associated derivatives were the main compounds observed. A flavanone (eriodictyol-O-glucoside) was also detected in 10 samples from four species.

In *Thesium*, rutin was the most common compound observed. Rutin was detectable (peak intensity > 0) in all but 2 species (*T. cymosum* and *T.* cf. *leandrianum*) and was the main peak (peak intensity > 30%) in 27 species. Other common compounds, present at peak intensities of more than 30%, included citric acid (26 species), isorhamnetin O-glucoside O-rhamnoside (20 species), kaempferol O-rutinoside (15 species), quinic acid (14 species), cryptochlorogenic acid (4-caffeoylquinic acid) (12 species), quercetin rhamnosyl-rhamnosyl-glucoside (11 species), dehydrodiferulic acid (10 species) and quercetin robinobioside (10 species) (Table 2). Numerous peaks of coumaroylquinic acid, feruloylquinic acid and flavonoid glycosides were also observed (Table 2).

Lacomucinaea had several prominent peaks, including citric acid, cryptochlorogenic acid, neochlorogenic acid, sinapoyl hexose, quercetin-O-dihexoside and an unknown compound (F121.2, Table 3). The main peaks observed in *Osyridicarpos* were citric acid, cryptochlorogenic acid, neochlorogenic acid, hydroxytyrosol O-glucoside, feruloylquinic acid and one unknown compound (F94.1, Table 3).

3.2. Chemophenetic Value at Generic Level

The PCA analysis and dendrogram showed both *L. lineata* and *O. schimperianus* to be deeply embedded within *Thesium* (Figures 1 and 2), and no compound or group of compounds were found to universally distinguish *Thesium* from *Lacomucinaea* or *Osyridicarpos*. Similarly, the main compounds plotted on the Zhigila et al. (2020) [7] phylogeny overlapped between *Thesium* and *Lacomucinaea* (Figure 3).

With the exception of *T. microcarpum* A.DC. (see Section 4), all *Thesium* species had one or more of the following six compounds in detectable amounts (peak intensity > 50): rutin, citric acid, isorhamnetin O-glucoside O-rhamnoside, kaempferol O-rutinoside, quinic acid and cryptochlorogenic acid. Three of the previously mentioned six compounds (isorhamnetin O-glucoside O-rhamnoside, kaempferol O-rutinoside and quinic acid) were not detected in *L. lineata* or *O. schimperianus*, and rutin was only detected at very low levels (peak intensity < 0.2 in *Lacomucinaea* and <35 in *Osyridicarpos*) (Table 3). All three genera had relatively high levels of citric acid, cryptochlorogenic acid, neochlorogenic acid and feruloylquinic acid (Table 3; Supplementary Figure S1).

Lacomucinaea and *Osyridicarpos* seem to be distinguishable from one another based on several compounds observed at high levels (peak intensity > 180) in each genus (see Table 3). While these compounds are observed at high levels in *Lacomucinaea* and *Osyridicarpos*, respectively, low levels (peak intensity < 100) were also occasionally present in the other genera studied. A larger sampling of *Lacomucinaea* and *Osyridicarpos* plants is needed to affirm these results.

3.3. Chemophenetic Value at Infrageneric Level

No discernible groupings were apparent among *Thesium* species using multivariate statistics (Figures 1 and 2), comparison with the Zhigila et al. (2020) phylogeny [7] (Figure 3) or manual examination of species' total ion chromatograms (Supplementary Figure S1). In the multivariate framework, neither the subgenera of the current DNA-based classification (Figure 2) nor the species from the same geographic regions (Supplementary Figure S2) or species with similar morphological characters (Supplementary Figure S3) grouped together.

One exception was a cluster of three morphologically similar species endemic to South Africa and Lesotho: *T. imbricatum* Thunb., *T. karooicum* Compton and *T. sonderianum* Schltr. (Figure 2; Supplementary Figure S3). *Thesium scandens* E.Mey. ex Sond., *T. zeyheri* A.DC., *T. ramosum* and *T. transvaalense* Schltr. were somewhat separated from the other species in the PCA analysis (Figure 1) and *T. zeyheri* and *T. transvaalense* in the dendrogram (Figure 2). *Thesium zeyheri* was the only species in which F7 (unknown sulphated caffeic acid derivative) was detected at high levels (peak intensity > 3), and both *T. zeyheri* and *T. transvaalense* had low levels of rutin (peak intensity < 12). Other compounds responsible for variation in principal components one to three and their details are provided in Supplementary Figure S4 and Supplementary Table S1. Another exception were the slight differences in the phenolic and organic acid composition of Subgenus *Hagnothesium* (see Section 4).

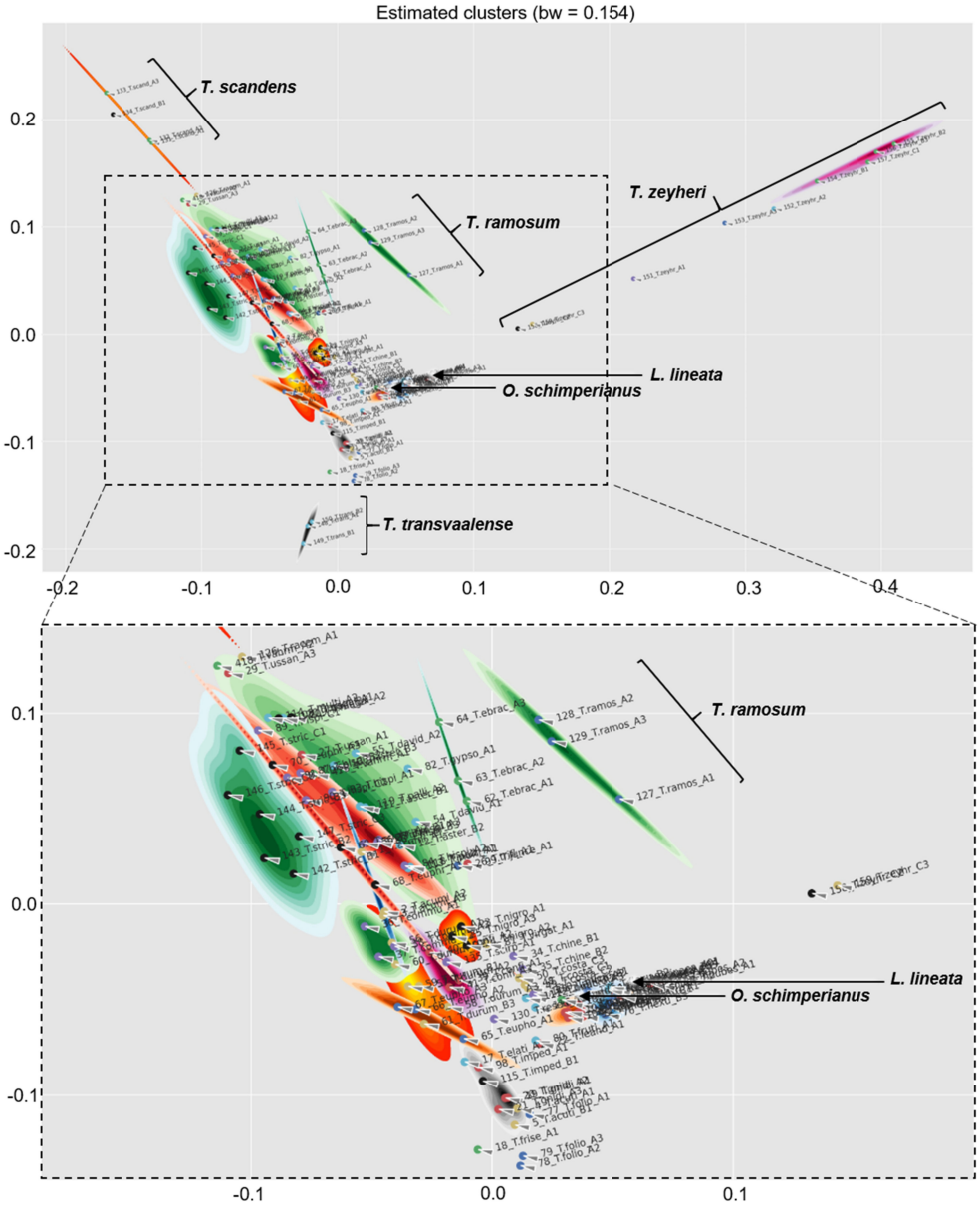

Figure 1. Principal component analysis with cluster map showing the relationships among 46 *Thesium* species, *Lacomucinaea lineata* and *Osyridicarpos schimperianus*, based on phenolic compound composition. Clusters are indicated with different colours. Species names corresponding to sample labels are provided in Table 1.

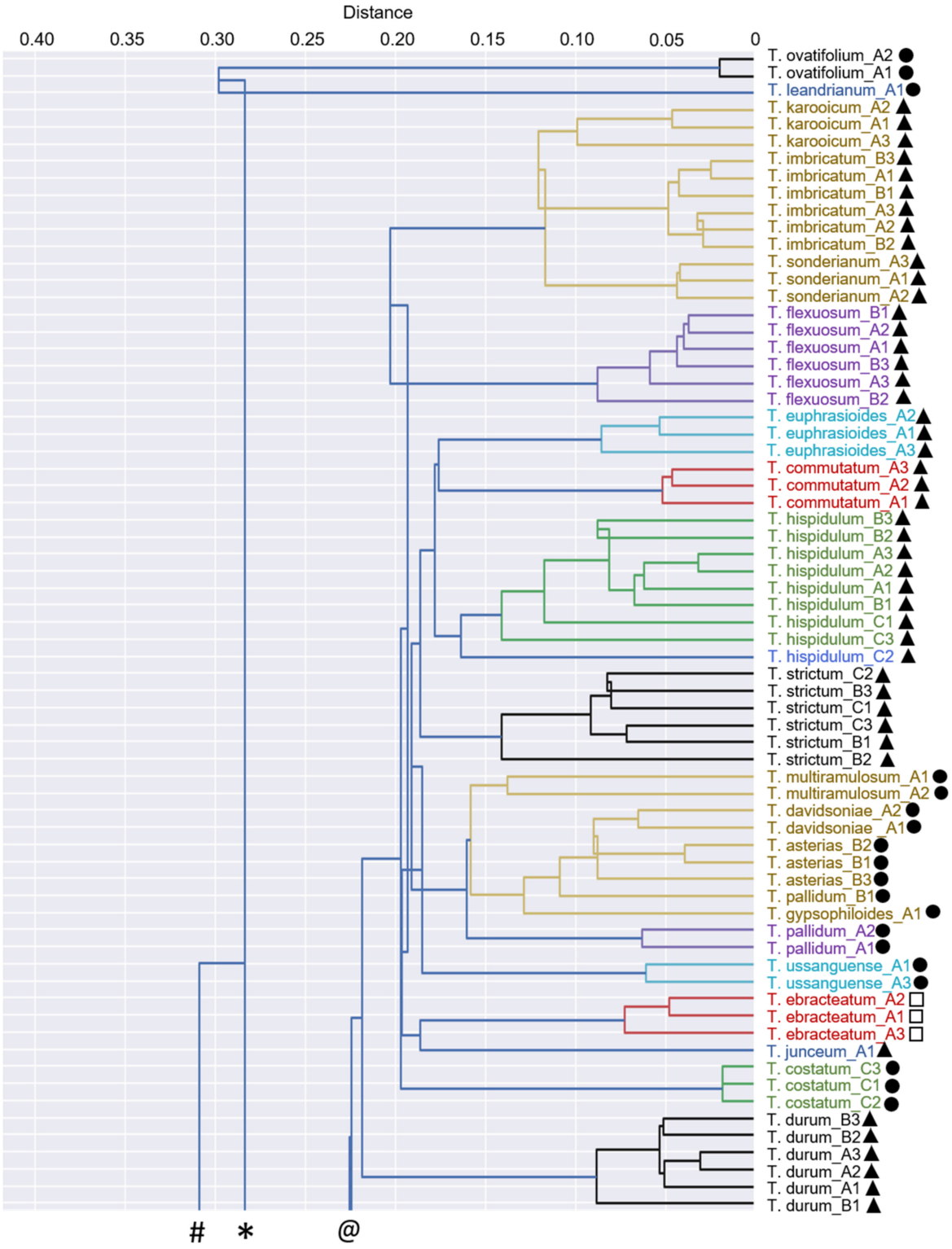

Figure 2. *Cont.*

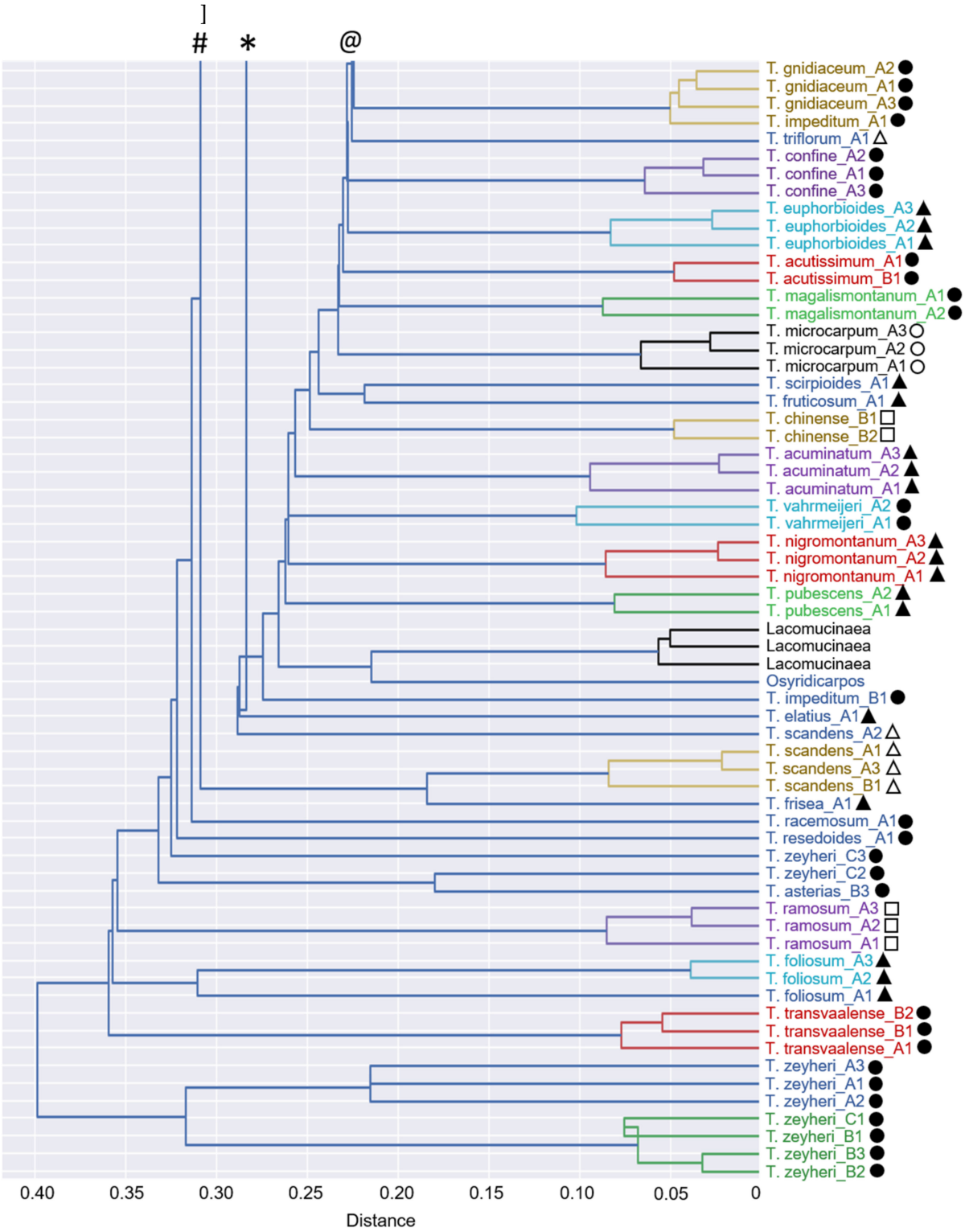

Figure 2. Dendrogram with clusters showing the relationships among 46 *Thesium* species, *Lacomucinaea lineata* and *Osyridicarpos schimperianus*, based on phenolic compound composition. The placements of subgenera are also shown: *Discothesium* (clear triangles), *Frisea* (solid triangles), *Hagnothesium* (clear circles), *Psilothesium* (solid circles) and *Thesium* (clear squares). Clusters are shown in different colours. Species names corresponding to sample labels are provided in Table 1.

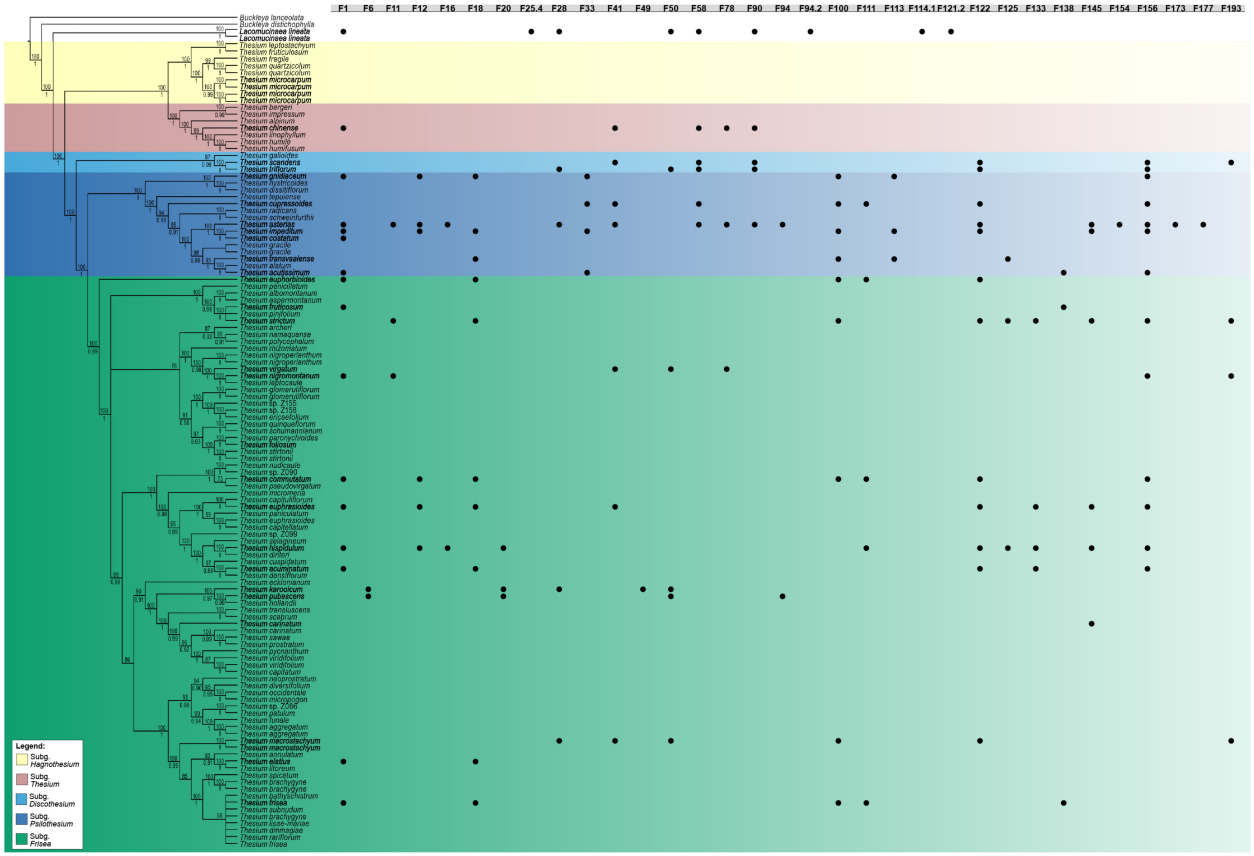

Figure 3. Patterns of phenolic compound variation observed in *Thesium* (27 species) and *Lacomucinaea* (1 species) within a phylogenetic framework. The most common phenolic compounds (peaks) detected in *Thesium* (see Table 2) and *Lacomucinaea* (see Table 3) are plotted on the most recent molecular phylogeny (adapted from [7]). The presence of compounds (peaks) in species is indicated with solid circles, and the five subgenera are shown in different colours. Compound names correspond with the information provided in Tables 2 and 3.

3.4. Chemophenetic Value at Interspecific Level

While no interspecific level signals were found in the phylogenetic framework (Figure 3), the multivariate framework and manual examination of species' total ion chromatograms revealed an interesting result. Almost all *Thesium* species, as well as both *L. lineata* and *O. schimperianus*, were largely distinct from one another in the combination of peaks and, to a lesser extent, relative peak intensities (Supplementary Figures S1 and S5). Indeed, 42 of the 50 species had at least one main peak that was not present in any of the other species. Furthermore, in the 38 species represented by multiple samples, the samples of each species not only grouped together (Figure 2) but also correlated more closely to one another than to those of other species (Figure 4). This indicates that phenolic total ion chromatograms might be useful in species identification when used in combination with morphological characteristics. In the three species (*T. pallidum* A.DC., *T. impeditum* A.W.Hill and *T. zeyheri*; Figure 2) where samples of a species did not group together, manual examination of their total ion chromatograms proved them to be conspecific (Supplementary Figure S1). Conversely, four species (*T. asterias* A.W.Hill, *T. costatum* A.W.Hill, *T. strictum* P.J.Bergius and *T.* cf. *ussanguense*) showed substantial variation (Supplementary Figures S1 and S2) (see Section 4).

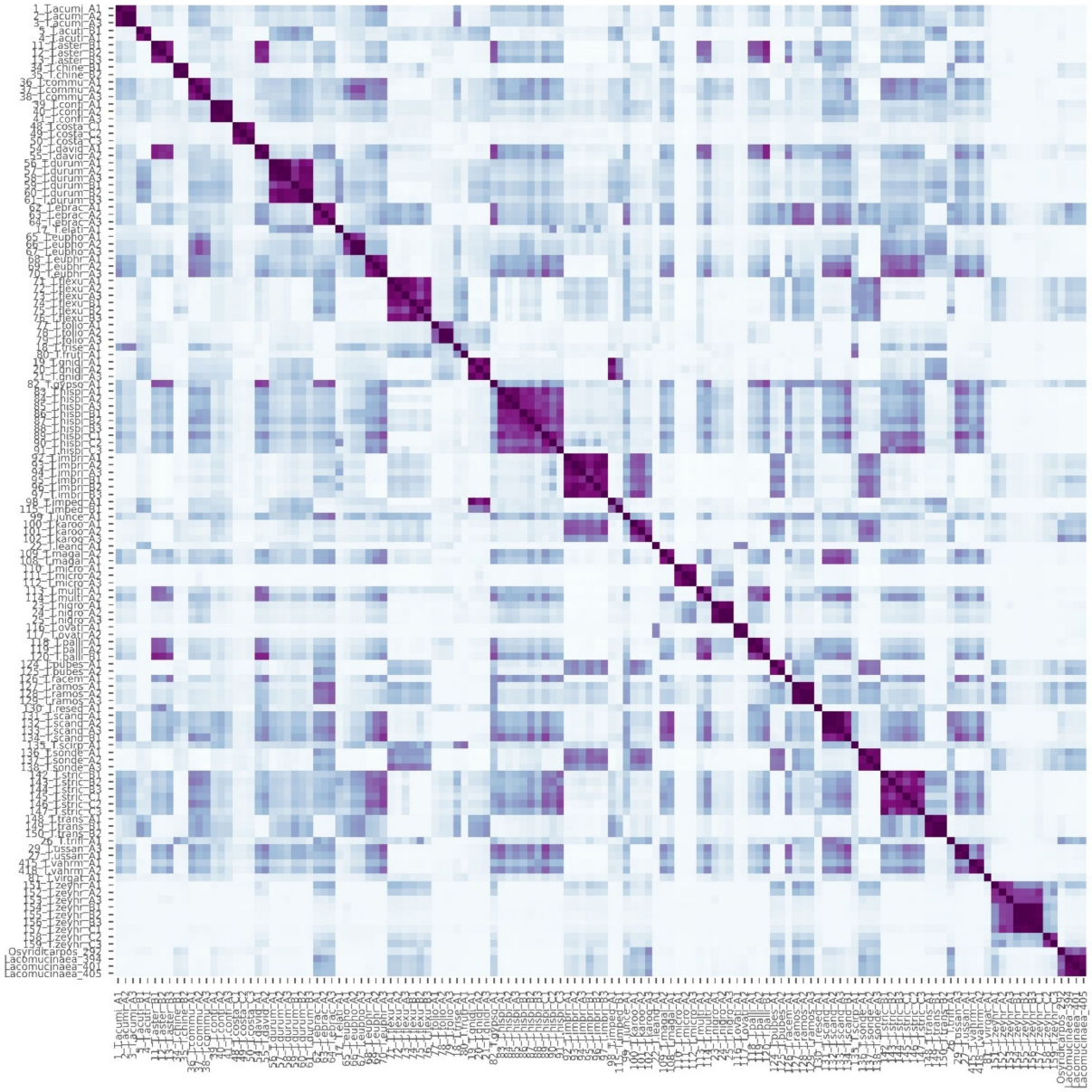

Figure 4. Correlation heatmap showing the similarities among 128 samples of 46 *Thesium* species, as well as sister genera *Osyridicarpos* and *Lacomucinaea*, based on phenolic compound composition. Darker colours indicate higher similarity between samples. Species names corresponding to sample labels are provided in Table 1.

3.5. Chemophenetic Value at Infraspecific Level

Out of the 38 species included in the PCA analysis, 4 species were represented by multiple samples from different populations. These species included *T. durum* Hilliard and B.L.Burtt (two populations), *T. flexuosum* A.DC. (two populations), *T. hispidulum* Lam. (three populations) and *T. imbricatum* (two populations). In *T. durum*, the samples from the same populations grouped together, thereby indicating that the two populations differed (Figure 5A). This difference was driven by slight variances in peak intensities and not dissimilarities in peak composition (Supplementary Figure S1). In the remaining three species (Figure 5B–D), the samples from populations intermixed, indicating that there were no distinguishable differences among populations in peak intensities or peak composition.

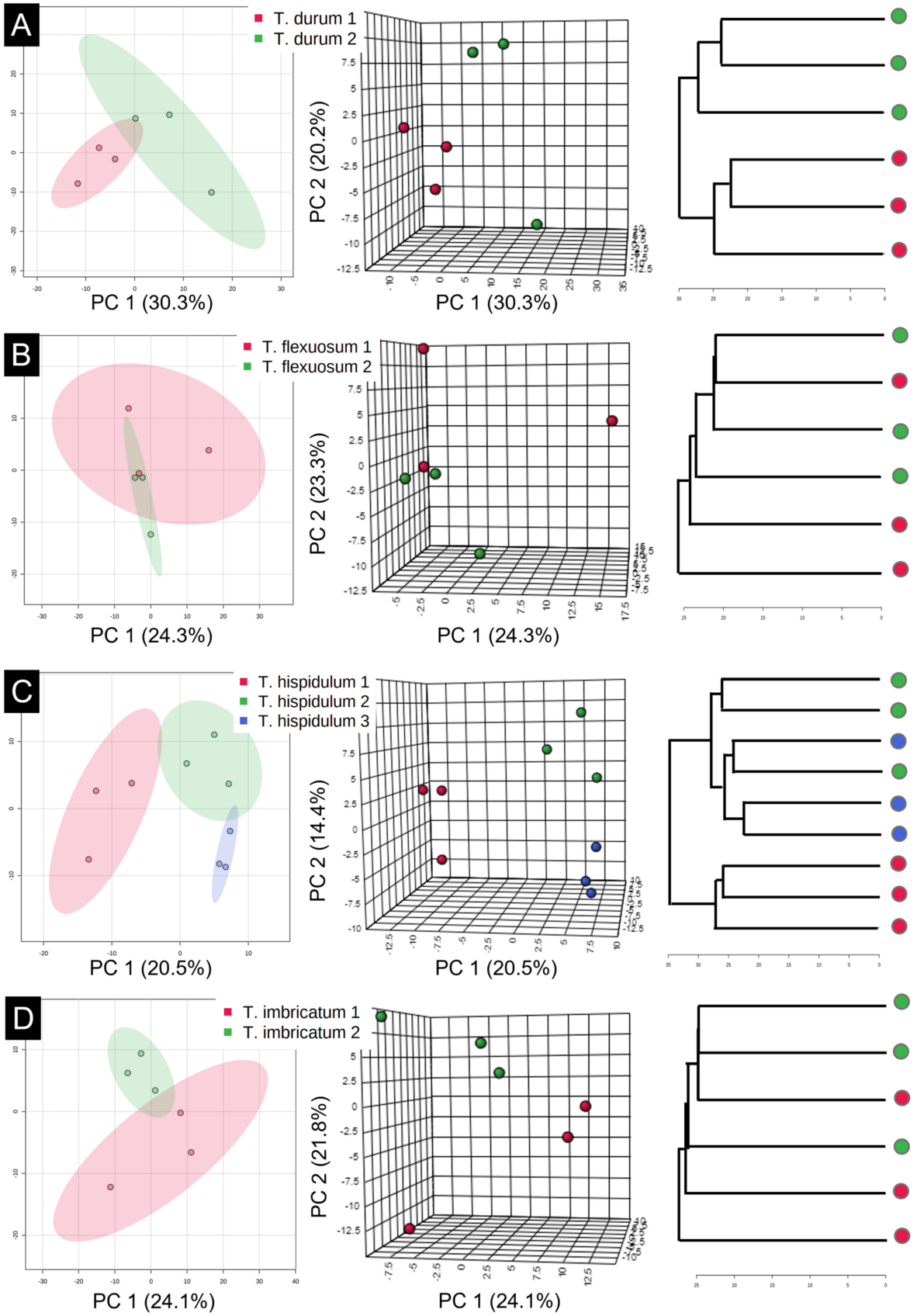

Figure 5. Two- and three-dimensional PCA plots and hierarchical clustering dendrograms comparing the phenolic compound compositions of different populations of four *Thesium* species (**A**). *T. durum*. (**B**). *T. flexuosum*. (**C**). *T. hispidulum*. (**D**). *T. imbricatum*. For each species, different populations are indicated by different colours.

4. Discussion

This study investigated the phenolic compound composition and chemophenetic relationships of 50 *Thesium* species, *L. lineata* and *O. schimperianus*, at the generic, infrageneric, interspecific and infraspecific levels. To our knowledge, it is the first study on patterns in phenolic compound variation in *Thesium*, *Lacomucinaea* and *Osyridicarpos*, and the second and most comprehensive chemophenetic study of *Thesium* to date (after that of De Kock and Rapson in 1938 [14]). With the exception of *T. chinense* and *L. lineata*, the phytochemistry of all of the species included here is studied for the first time.

In line with previous studies (e.g., Refs [3,18–23]), the phenolic compounds found in *Thesium* were mainly flavonols, carboxylic acids, phenolic acids and associated derivatives. Although the phenolic compound diversity in *Thesium* appears to be high (more than 200 prominent peaks were observed), the majority of the compounds are relatively common in the plant kingdom (e.g., citric acid, rutin, chlorogenic acid isomers, quinic acid derivatives and flavonoid glycosides). To our knowledge, several compounds, including cryptochlorogenic acid, neochlorogenic acid, coumaroylquinic acid, feruloylquinic acid, eriodictyol and quercetin are reported here for the first time in *Thesium* (see Ref [3]). No flavones were detected as major compounds, even though glycosides of apigenin, luteolin and chrysoeriol were previously reported from *T. chinense*.

Overall, the phenolic and organic acid composition seems to have some chemophenetic value at the generic level but limited value at the infrageneric level. Preliminary results show distinctive phenolic and organic acid differences between *Lacomucinaea* and *Osyridicarpos,* which support molecular data to retain them as separate genera [1]. While both genera have leaves with distinct petioles and smooth fleshy fruit, there are noteworthy differences in, among others, the shape of the leaves and the structure of the flowers [1]. In *Thesium,* however, no compound or group of compounds could be found to universally differentiate all *Thesium* species from *Lacomucinaea* or *Osyridicarpos*. Nevertheless, the majority of *Thesium* species have at least one of six compounds (rutin, citric acid, isorhamnetin O-glucoside O-rhamnoside, kaempferol O-rutinoside, quinic acid and cryptochlorogenic acid) at relatively high intensities. One exception is *T. microcarpum*, in which only very low levels of citric acid and rutin were detected. *Thesium microcarpum* forms part of Subgenus *Hagnothesium,* which was previously recognised as a separate genus, *Thesidium* Sond. [7,24]. Species of Subgenus *Hagnothesium* are characterised by, among others, their dioecious reproductive system and mostly 4-merous flowers (versus monoecious reproductive system and predominantly 5-merous flowers in the other subgenera of *Thesium*) [7]. Although molecular studies have shown *Thesidium* to be embedded in *Thesium* [4,7], it is interesting to note that its phenolic and organic acid composition differs somewhat from the main trend observed in *Thesium*. Not only are the main compounds found in the majority of *Thesium* species almost absent from *T. microcarpum*, but it also contains several main peaks not present in the other species investigated. More extensive studies of Subgenus *Hagnothesium*, as well as *Lacomucinaea* and *Osyridicarpos,* are needed to further explore these apparent patterns.

Contrary to the work conducted on a phlobatannin and/or an unidentified volatile oil glycoside [13], no discernible infrageneric groupings were present among the species investigated. Although *T. transvaalense* was slightly further removed from the other species in the PCA analysis, and *T. zeyheri* and *T. transvaalense* in the dendrogram, no significance could be linked to these placements in terms of morphology, distribution or molecular relationships. The separation of *T. zeyheri* and *T. transvaalense* was likely driven by an unknown saccharide (F7) and rutin, as the majority of variation in the first principal component of the PCA analysis was attributed to these two compounds (Supplementary Figure S4). It is possible that the study of additional *Thesium* species may alter the preliminary trends and relationships observed in this study. Future studies might also investigate patterns in the imino acid *cis*-4-Hydroxy-L-proline [25] and acetylenic seed oils [26], as these compounds have shown chemophenetic potential at the generic level in Santalaceae.

The value of phenolic compound composition apparently lies at the interspecific level, as many *Thesium* species, as well as *L. lineata* and *O. schimperianus*, are recognisable to some

degree by their total ion chromatograms and might therefore aid species identification when used in combination with morphological characters. However, this approach should be used with caution, as the chemistry of the majority of *Thesium* species has not yet been studied. The combination of peaks and their relative intensities can also be helpful in distinguishing between morphologically similar species. For instance, *T. triflorum* L.f. and *T. scandens* are closely related based on both morphology [6] and DNA data [7]. It is often difficult to distinguish them from one another—to such a degree that they might be considered conspecific based on morphological species concepts [27]. The main morphological difference between them is found in their leaves: fleshy, more or less terete leaves without clearly visible veins in *T. scandens* and flat, linear to linear-lanceolate leaves with one to three (five) main veins in *T. triflorum* (Figure 6). However, there seems to be a continuum of variation in leaf morphology between *T. scandens* plants, which occur in the Cape regions of South Africa, and *T. triflorum* plants, which occur in subtropical and tropical Africa. The total ion chromatograms of these two species, collected about 800 km apart, had several similarities (Figure 6). They shared all peaks except for neochlorogenic acid and cryptochlorogenic acid, which was present at high intensities in *T. triflorum* but absent from *T. scandens*. The presence or absence of these compounds can be a useful tool to distinguish between *T. triflorum* and *T. scandens* and illustrates the potential of phenolic total ion chromatograms to aid future systematic studies of this taxonomically difficult genus.

Trends in the infraspecific variation are less clear and require further study. The majority of species where multiple individuals were studied (38 species) showed almost no variation in peak composition and limited variation in peak intensity (Figures 2, 4 and 5; Supplementary Figure S1). This indicates that phenolic total ion chromatograms can be used with some confidence to distinguish among species while keeping in mind that the majority of *Thesium* species have not yet been studied. Four exceptions were *T. asterias*, *T. costatum*, *T. strictum* and *T.* cf. *ussanguense,* where substantial variation in peak composition and peak intensity was present. This variation might be attributed to several factors, including laboratory-induced differences between sample runs, widely varying local habitat conditions, possible lateral transfer of phenolic compounds, misidentified samples and erroneous species concepts. It has been shown that some parasitic plants have very similar phenolic profiles to their host plants [17] and that they may contain different major flavonoids and xanthones depending on their hosts [28]. This poorly explored aspect deserves further study but will also require in-depth studies of the host diversity and host specificity of *Thesium* species. In the case of *T. asterias*, however, there is evidence that the species delineation needs to be reconsidered. Closer examination of the two populations studied (roughly 700 km apart) revealed noteworthy morphological differences (Figure 7). Plants from Limpopo, South Africa, had thickened, perennial rootstocks and virgate, parallel branches, while plants from KwaZulu-Natal, South Africa, had slender, possibly annual rootstocks and with branches angled at about 45° from the central stem (Figure 7). Further taxonomic investigation will likely prove these two populations to be different species, one of which may be new to science. This example shows how phenolic total ion chromatograms may reveal inconsistencies in current species concepts.

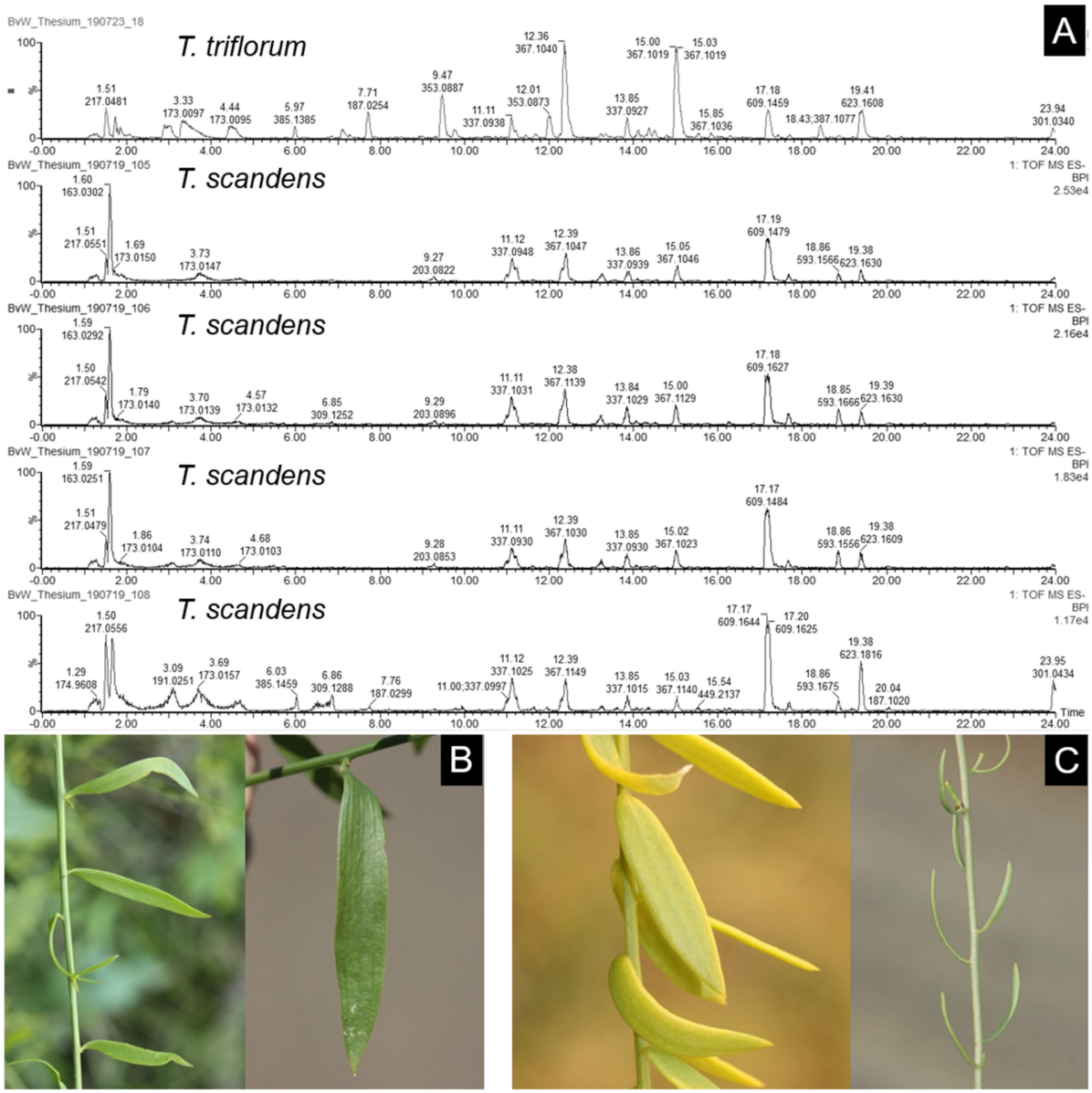

Figure 6. Comparison of the phenolic compound composition and morphology between two closely related *Thesium* species, *T. triflorum* and *T. scandens*. (**A**). Total ion chromatograms of a *T. triflorum* plant from near St. Lucia, KwaZulu-Natal, South Africa (**top**) and four *T. scandens* plants from Eastern Cape, South Africa (**bottom**). (**B**). Flat linear to linear-lanceolate leaves of *T. triflorum* with one to three (five) main veins. (**C**). Fleshy and often terete leaves of *T. scandens* lacking clearly visible veins.

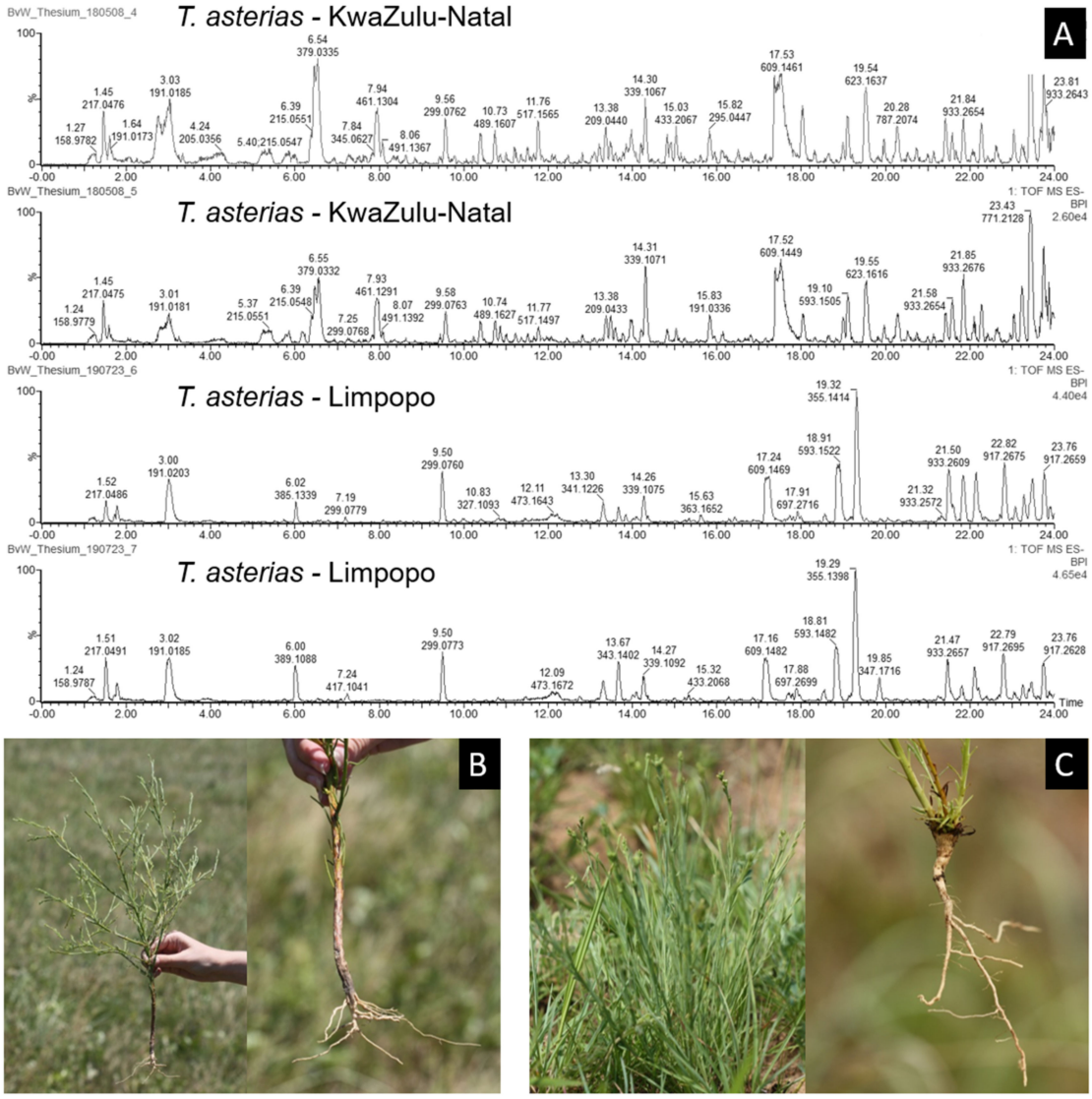

Figure 7. Differences in phenolic compound composition and morphology between two populations of *Thesium asterias*. (**A**). Total ion chromatograms of two *T. asterias* plants from a single population near St. Lucia, KwaZulu-Natal, South Africa (**top**) and two *T. asterias* plants from a single population near Lekgalameetse Nature Reserve, Limpopo, South Africa (**bottom**). (**B**). Habit, with branches spreading at ±45 °, and slender rootstock of *T. asterias* plants from KwaZulu-Natal. (**C**). Habit, with parallel and virgate branches, and thickened rootstock of *T. asterias* plants from Limpopo.

5. Conclusions

This study provides valuable preliminary insights into the phenolic diversity and chemophenetic relationships in *Thesium*, *Lacomucinaea* and *Osyridicarpos*. The phytochemistry of 50 *Thesium* species and *Osyridicarpos* is studied for the first time, and the first chemophenetic evaluation of the genus in more than 80 years is provided.

LCMS analysis revealed a wide diversity of compounds in *Thesium* (more than 200 prominent peaks), mainly flavonols, carboxylic acids, phenolic acids, organic acids and associated derivatives. This result is in line with previous studies on the genus.

The phenolic compounds in *Thesium* seem to have some chemophenetic value at the generic level but limited value at the infrageneric level. While *Lacomucinaea* and *Osyridicarpos* appear to be distinguishable from one another based on the presence and absence of several main peaks, this result should be confirmed with a larger sampling of individuals and populations. The true value of phenolic compound variation in *Thesium* lies at the species level. Due to the wide variety of observed peaks, the total ion chromatograms

of the species are largely distinct from one another and may therefore assist with species identification and help tease apart unclear species boundaries. While infraspecific chemical variation seems limited in the majority of species, studies of multiple plant populations are needed to verify the degree of infraspecific variation. Further examination of about 300 *Thesium* species that remain phytochemically unknown will not only enhance our knowledge of phenolic diversity in the genus but might also alter the preliminary trends and relationships observed in this study.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/d14080590/s1. Supplementary Table S1. Phenolic compounds detected in *Thesium* responsible for the most variation in Principal Components one to four. Supplementary Figure S1. Total ion chromatograms of 156 sample extracts of 50 *Thesium* species (1–50), *Osyridicarpos schimperianus* (51) and *Lacomucinaea lineata* (52). Supplementary Figure S2. Dendrogram of phenolic relationships showing the geographic distributions of 46 *Thesium* species, *Lacomucinaea lineata* and *Osyridicarpos schimperianus*. Supplementary Figure S3. Dendrogram of phenolic relationships among 46 *Thesium* species, *Lacomucinaea lineata* and *Osyridicarpos schimperianus*, accompanied by the distribution of characters states of nine taxonomically important morphological characters: A. Root type. B. Maximum plant height. C. Plant surface. D. Stem type. E. Leaf type. F. Inflorescence type. G. Reproductive system. H. Inner corolla lobe surface. I. Style length. Supplementary Figure S4. Loading plots for Principal Components 1–3 (A–C) with the most significant contributors indicated. Supplementary Figure S5. Heatmap of the main phenolic peaks detected in the extracts of 128 samples, representing 46 *Thesium* species, as well as sister genera *Osyridicarpos* and *Lacomucinaea*.

Author Contributions: Conceptualisation, B.-E.V.W., N.L. and M.M.L.R.; sample collection, B.-E.V.W., N.L. and M.M.L.R.; formal analysis, H.R., N.L. and M.A.S.; writing—original draft preparation, N.L.; writing—review and editing, B.-E.V.W., M.A.S., H.R. and M.M.L.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Foundational Biodiversity Information Programme, small grant number 104931, National Research Foundation of South Africa, grant number 84442, the Botanical Education Trust, University of Johannesburg and the South African National Biodiversity Institute.

Data Availability Statement: Not applicable.

Acknowledgments: The following people and organisations are thanked: the University of Johannesburg, South African National Biodiversity Institute and University of Stellenbosch for hosting this study; Bronwynn Egan, Daniel Nickrent, Kate and Graham Grieve, John Burrows, Magda Nel, Reuhl Lombard, Sedzani Simali, Teodor van Wyk and Wiam Haddad for assistance in the field; Zigmantas Gudzinskas for providing fresh material of *T. ebracteatum* and *T. ramosum* from Lithuania; Daniel Zhigila for assistance with species identifications; Malcolm Taylor from Central Analytical Facilities, University of Stellenbosch, for assistance with data analysis and compound identifications.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Nickrent, D.L.; García, M.A. *Lacomucinaea*, a new monotypic genus in Thesiaceae (Santalales). *Phytotaxa* **2015**, *224*, 173–184. [CrossRef]
2. The Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* **2016**, *181*, 1–20. [CrossRef]
3. Lombard, N.; Van Wyk, B.-E.; le Roux, M.M. A review of the ethnobotany, contemporary uses, chemistry and pharmacology of the genus *Thesium* (Santalaceae). *J. Ethnopharmacol.* **2020**, *256*, 112745. [CrossRef] [PubMed]
4. García, M.A.; Nickrent, D.L.; Mucina, L. *Thesium nautimontanum*, a new species of Thesiaceae (Santalales) from South Africa. *PhytoKeys* **2018**, *109*, 41–51. [CrossRef]
5. De Candolle, A. Santalaceae. In *Prodromus Systematis Naturalis*; De Candolle, A., Ed.; Treuttel et Würtz: Paris, France, 1857; pp. 619–692.
6. Hill, A.W. Order CXX. Santalaceae. In *Flora Capensis, Volume 5*; Thiselton-Dyer, W.T., Ed.; L. Reeve & Co. Ltd.: London, UK, 1925; pp. 135–212. [CrossRef]
7. Zhigila, D.A.; Verboom, G.A.; Muasya, A.M. An infrageneric classification of *Thesium* (Santalaceae) based on molecular phylogenetic data. *Taxon* **2020**, *69*, 100–123. [CrossRef]

8. Linnaeus, C. *Species Plantarum: Exhibentes Plantas Rite Cognitas, Ad Genera Relatas, Cum Differentiis Specificis, Nominibus Trivialibus, Synonymis Selectis, Locis Natalibus, Secundum Systema Sexuale Digestas*; Tomus I; Laurentius Salvius: Stockholm, Sweden, 1753; pp. 207–208.
9. Hill, A.W. The genus *Thesium* in South Africa, with a key and descriptions of new species. *Bull. Misc. Inf.* **1915**, *1*, 1–43. [CrossRef]
10. Pilger, R. Santalaceae. In *Die Natürlichen Pflanzen Familien*, 2nd ed.; Engler, A., Ed.; W. Engelmann: Leipzig, Germany, 1935; pp. 52–91.
11. Bobrov, E.G. Family XLIX. Santalaceae R.Br. In *5. Flora of the U.S.S.R*; Komarov, V.L., Ed.; Israel Program for Scientific Translations: Jerusalem, Israel, 1970; pp. 325–340.
12. Hendrych, R. The natural history and systematic of the genus *Thesium* L. *Acta Univ. Carol. Biol.* **1970**, *4*, 293–358.
13. Stander, M.A.; Brendler, T.; Redelinghuys, H.; Van Wyk, B.-E. The commercial history of Cape herbal teas and the analysis of phenolic compounds in historic teas from a depository of 1933. *J. Food Compos. Anal.* **2019**, *76*, 66–73. [CrossRef]
14. De Kock, P.C.; Rapson, W.S. A correlation between the chemical constituents and morphological structure of certain *Thesium* species. *Nature* **1938**, *142*, 1078–1079. [CrossRef]
15. Moore, T.E.; Verboom, G.A.; Forest, F. Phylogenetics and biogeography of the parasitic genus *Thesium* L. (Santalaceae), with an emphasis on the Cape of South Africa. *Bot. J. Linn. Soc.* **2010**, *162*, 435–452. [CrossRef]
16. Willems, J.L.; Khamis, M.M.; Mohammed-Saeid, W.; Purves, R.; Katselis, G.; Low, N.H.; El-Aneed, A. Analysis of a series of chlorogenic acid isomers using differential ion mobility and tandem mass spectrometry. *Anal. Chim. Acta* **2016**, *933*, 164–174. [CrossRef] [PubMed]
17. Assanga, S.B.I.; Luján, L.M.L.; Ruiz, J.C.G.; McCarty, M.F.; Cota-Arce, J.M.; Espinoza, C.L.L.; Gil Salido, A.A.; Ángulo, D.F. Comparative analysis of phenolic content and antioxidant power between parasitic *Phoradendron californicum* (toji) and their hosts from Sonoran Desert. *Results Chem.* **2020**, *2*, 100079. [CrossRef]
18. Lu, Y.; Wang, J. Study on chemical constituents of *Thesium chinense*. *Chin. Tradit. Herb. Drugs* **2004**, *35*, 491–493.
19. Parveen, Z.; Deng, Y.; Saeed, M.K.; Dai, R.; Ahamad, W.; Yu, Y.H. Anti-inflammatory and analgesic activities of *Thesium chinense* Turcz. extracts and its major flavonoids, kaempferol and kaempferol-3-O-glucoside. *Yakugaku Zasshi* **2007**, *127*, 1275–1279. [CrossRef] [PubMed]
20. Lee, I.K.; Kim, K.H.; Choi, S.U.; Lee, J.H.; Lee, K.R. Phytochemical constituents of *Thesium chinense* Turcz. and their cytotoxic activities in vitro. *Nat. Prod. Sci.* **2009**, *15*, 246–249.
21. Liu, Y.; Zahida, P.; Deng, Y.L.; Zhang, J.H.; Dai, R.J. Study on the flavonoids compounds of *Thesium chinense*. *Zhong Yao Cai* **2009**, *32*, 518–520.
22. Liu, C.; Li, X.-T.; Cheng, R.-R.; Han, Z.-Z.; Yang, L.; Song, Z.-C.; Wang, Z.-T. Anti-oral common pathogenic bacterial active acetylenic acids from *Thesium chinense* Turcz. *J. Nat. Med.* **2018**, *72*, 433–438. [CrossRef] [PubMed]
23. Zou, Y.; Hong, M.; Yang, X. Isolation of chemical components from *Thesium chinense*. *Zhongguo Shiyan Fangjixue Zazhi* **2016**, *22*, 74–77.
24. Forest, F.; Manning, J.C. The minor genera *Kunkeliella* and *Thesidium* included in *Thesium*. *Bothalia* **2013**, *43*, 214–216. [CrossRef]
25. Kuttan, R.; Pattabhiraman, K.S.V.; Radhakrishnan, A.N. Possible chemotaxonomic significance of the occurrence of *cis*-4-hydroxy-L-proline in Santalaceae. *Phytochemistry* **1974**, *13*, 453–454. [CrossRef]
26. Hopkins, C.Y.; Chisholm, M.J.; Cody, W.J. Fatty acid components of some Santalaceae seed oils. *Phytochemistry* **1969**, *8*, 161–165. [CrossRef]
27. Hilliard, O.M. Thesium. In *Flora Zambesiaca, Vol. 9, Part 3: Polygonaceae and Myriaceae*; Pope, G.V., Polhill, R.M., Martins, E.S., Eds.; Royal Botanic Gardens, Kew: Richmond, UK, 2006; pp. 206–241.
28. Jadhav, R.B.; Anarthe, S.J.; Surana, S.J.; Gokhale, S.B. Host-hemiparasite transfer of the C-glucosyl xanthone mangiferin between *Mangifera indica* and *Dendrophthoe falcate*. *J. Plant Interact.* **2005**, *1*, 171–177. [CrossRef]