4. Species-level phylogeny

4.1 Molecular evolution

The aligned plastid data included the \textit{rbcL} gene with 1378 base pairs (bp) and the \textit{trnL-F} gene with 1287 bp. The aligned nuclear ITS dataset consisted of 705 bp. Due to ambiguous alignments, portions of the \textit{trnL-F} region had to be excluded (three regions, 300 bp in total). The length of the aligned \textit{trnL-F} dataset for 151 taxa was 890 after ambiguously aligned regions were excluded. Some taxa in the \textit{trnL-F} region had deletions of 397 bp or more (e.g. \textit{Gnidia coriacea}, \textit{G. galpinii} 1, \textit{G. galpinii} 2, \textit{G. humilis}, \textit{G. cf. juniperifolia}, \textit{G. pilosa}, \textit{G. squarrosa}, \textit{G. aff. subulata}, \textit{Pimelea ammoncharis}, \textit{P. argentea}, \textit{P. clavata} 1, \textit{P. clavata} 2, \textit{P. forrestiana}, \textit{P. gigiana}, \textit{P. graniticola}, \textit{P. halophila}, \textit{P. holroydii}, \textit{P. micrantha}, \textit{P. pelinos}, \textit{P. pygmaea}, \textit{P. rara}, \textit{P. spectabilis}, \textit{P. spiculigera} var. \textit{thesioides} 1, \textit{P. spiculigera} var. \textit{thesioides} 2, \textit{P. trichostachya}, \textit{Thecanthes punicea} and \textit{T. sanguinea}). The aligned region of ITS contained the most variable sites (Table 4.1) namely 379 (58.13%) compared to \textit{trnL-F} with 401 (40.75%) and \textit{rbcL} with 340 (24.67%). The number of parsimony informative characters were also higher for ITS (294; 45.09%), than for \textit{trnL-F} (247; 25.10%), and \textit{rbcL} (239; 17.34%). In Table 4.2 ITS had a 1.97 excess of transitions (ts), the consistency index (CI) and retention index (RI) values of ts were less than those for transversions (tv), CI = 0.35 vs. 0.46, RI = 0.72 vs. 0.81, ts and tv, respectively. For \textit{rbcL} there was a 1.52 excess of ts, and the CI and RI values of ts were higher than those for tv: CI = 0.53 vs. 0.50, RI = 0.86 vs. 0.81 (ts and tv respectively). For \textit{trnL-F} there was a 1.03 excess of ts. This more frequent ts performed slightly better, CI = 0.7 vs. 0.63; RI = 0.87 vs. 0.81 (ts and tv respectively). Variable positions in all three regions changed faster for ITS, 4.56 vs. 2.28 (\textit{rbcL}) and 1.96 (\textit{trnL-F}).

4.2 Plastid results

4.2.1 Separate molecular analyses

Separate analyses of \textit{rbcL} and \textit{trnL-F} will not be discussed in detail, due to low resolution. The combined plastid analyses will be the main focus.

4.2.1.1 The coding region \textit{rbcL}

Of the 1378 included characters, 1038 were constant, 340 (24.67%) variable and 239 (17.34%) parsimony informative. Analysis resulted in 47 equally parsimonious trees with a tree length (TL) of 774 steps. The CI was low at 0.52, while the RI was higher at 0.84 (Table 4.1). Figure 4.1A shows the bootstrap consensus tree, with bootstrap values equal to or higher than 50%.
4. Species-level phylogeny
4. Species-level phylogeny
4. Species-level phylogeny

4.2.1.2 The non-coding region trnL-F
Of the 984 included characters, 583 were constant, 401 (40.75%) variable and 247 (25.10%) parsimony informative. Analysis resulted in 27 equally parsimonious trees with a TL of 785 steps. The CI was 0.66, while the RI was higher at 0.84 (Table 4.1). Figure 4.1B shows the bootstrap consensus tree, with bootstrap values equal to or higher than 50%.

4.2.2 Combined molecular analysis (rbcL and trnL-F)
No strongly supported incongruent patterns could be observed between the rbcL and trnL-F analyses thus the datasets were combined directly. Of the 2362 included characters, 1621 were constant, 741 (31.37%) variable and 486 (20.58%) parsimony informative. Analysis resulted in 1630 equally most parsimonious trees with a tree length of 1680 steps. The CI was low at 0.55, while the RI was higher at 0.81 (Table 4.1). Figure 4.2A shows the bootstrap consensus tree.

The bootstrap percentages (BP) in the combined analysis was much higher compared to those of the separate analyses. *Dialyceras* and *Rhopalocarpus* (Sphaerosepalaceae) formed a well supported clade (100BP). Thymelaeaceae were strongly supported as monophyletic with a bootstrap value in the Fitch analysis of 94%. The subfamilies Aquilarioideae and Synandrodaphnoideae were successively sister to Thymelaeoideae, with strong to moderate bootstrap support (89BP and 68BP respectively). Sister to the rest of the Thymelaeaceae was the fourth subfamily Gonystylideae, including two clades: (I) three species of *Octolepis*: *O. dioica*, *O. dioica* f. *oblancoelata* and *Octolepis* sp. (97BP); (II) *Arnhemia*, *Deltaria*, *Gonystylus*, *Solmsia*, and three representatives of *Lethedon*: *L. balansae*, *L. cernua* and *L. aff. salicifolia* (76BP). The Thymelaeoideae were well supported as being monophyletic (81BP) including two well supported clades namely (I) tropical Africa and south-eastern Asia, and (II) non-African and Gondwanan taxa. Clade (I) received strong support (95BP), with *Craterosiphon* and *Synaptolepis* sister (95BP), and *Enkleia* (76BP) as well as *Dicranolepis* sister to this pair. Clade (II), also with strong support (88BP), consisted of many subclades, some of them low in resolution. Next is a relatively well supported subclade (77BP) comprising two strongly supported sister lineages: 1) *Daphne*, *Diarthron* and *Thymelaea* (88BP); 2) *Stellera* and *Wikstroemia* (96BP), with *Edgeworthia* sister to the pair (96BP). Very strong support was obtained for the sister pair *Dais* and *Phaleria* (99BP). The genera *Lachnea* and *Passerina* received moderate support for being monophyletic (71BP and 61BP respectively). The representatives of the genus *Stephanodaphne* formed a clade with moderate support (78BP). The clade consisting of the two species of *Peddiea* received strong support (100BP). The position of *Dirca* and *Ovidia* was not resolved. The largest genus in clade (II), *Gnidia*, was polyphyletic, with at least five lineages.
The first lineage consists of *Drapetes muscoides* sister to some of the *Gnidia* species with very weak support (57BP). These *Gnidia* species, viz. G. cf. *anomala*, *G. denudata*, *G. fastigiata*, *G. geminiflora*, *G. renniana* and G. aff. *renniana* received moderate support (64BP). In the second lineage, *Gnidia pinifolia* and *G. racemosa* were successively sister to representatives of the genus *Struthiola* (100BP). The genus *Struthiola* including the following species: *S. ciliata*, *S. dodecandra*, *S. leptantha*, *S. salters*, *S. striata* and *S. tomentosa*, received moderate support (70BP). The third lineage was well supported (94BP), and consisted of *Gnidia aberrans*, *G. caniflora*, *G. scabrida*, *G. setosa*, *G. singularis* and *G. wikstroemiana*. In the fourth lineage some of the *Gnidia* species were sister to the genus *Pimelea* with very weak support (50BP). The representatives of the genus *Pimelea* received moderate support (72BP), including the genus *Thecanthes* with strong support for *T. punicea* and *T. sanguinea* as a sister pair (84BP). Two clades with no resolution, viz. *Gnidia coriacea*, *G. galpinii*, *G. humilis*, G. cf. *juniperifolia* and G. *pilosa*; and *Gnidia phaeotricha* and *G. squarrosa* received moderate to strong bootstrap support (60BP and 89BP respectively) as being sister to the genus *Pimelea*. Sister to the rest was *Gnidia aff. subulata*. The fifth *Gnidia* lineage included species previously classified under *Lasiosiphon* grouped together with moderate bootstrap support of 79%. This lineage consists of a weakly supported clade (55BP), including, *Gnidia bojeriana*, *G. caffra*, *G. calocephala*, *G. dumetorum*, *G. gilbertae*, *G. glauca*, *G. kraussiana*, *G. madagascariensis*, *G. sericocephala* *Gnidia* sp. and *G. wilmsii*. Sister to the rest was a strongly supported clade with *Gnidia danguyana* and *G. decaryana* (83BP) as well as *G. bakeri*.

### 4.3 Nuclear results (ITS)

Of the 652 included characters, 273 were constant, 379 (58.13%) variable and 294 (45.09%) parsimony informative. Analysis resulted in 33 equally parsimonious trees with a TL of 1727 steps. The CI was very low at 0.39, while the RI was higher at 0.75 (Table 4.1). Figure 4.2B shows the bootstrap consensus tree, with bootstrap values equal to or higher than 50%.

Due to alignment problems as stated in 3.4.1, the same representatives of the Thymelaeoideae found in the plastid analyses could not be included. Two clades could be observed. In clade (I) there was more support than what could be observed for the two sister lineages found in the combined plastid analysis: *Daphne* and *Thymelaea* (85BP); and the two *Wikstroemia* species (82BP) with *Edgeworthia* sister to the pair. In clade (II), *Gnidia* was again paraphyletic, but with at least four lineages. The first lineage contained some of the *Gnidia* species and does not include *Drapetes muscoides* (100BP). *Gnidia pinifolia* and *G. racemosa* grouped with *Struthiola* (96BP), but with two lower support could be observed in the second lineage. In the third lineage, *Gnidia* species were successively sister to the
4. Species-level phylogeny

genera *Pimelea* and *Thecanthes* with much higher support (99BP). The fourth lineage included taxa, previously known as *Lasiosiphon*, grouped with a few *Gnidia* species with relatively strong support (98BP).

The clades including the genera *Dirca* and *Ovidia*, *Stephanodaphne*, and *Peddiea* as well as the third *Gnidia* lineage were successively sister to lineage one and two of *Gnidia*, and both the genera *Lachnaea* and *Passerina* (98BP). The fourth lineage of *Gnidia* and *Dais cotinifolia* were successively sister to the rest of clade (II) of the Thymelaeoideae, with moderate to strong support (62BP and 83BP respectively).

On the specific level, the first lineage does not include the genus *Drapetes* when compared to the combined plastid analysis, *rbcL* and *trnL-F*. The position of *Drapetes muscoides* was unresolved. Stronger support was observed for the small lineage including *Gnidia anomala*, *G. denudata* and *G. renniana* (100BP). In the second lineage, both *Gnidia pinifolia* and *G. racemosa*, with less bootstrap support of 97% and 96% respectively, were successively sister to the stronger supported genus *Struthiola* (94BP). The monophyly of *Passerina* received much higher support (98BP), while the relationship within *Lachnaea* was unresolved. *Lachnaea* received no support for monophyly but instead formed two lineages.

The sister pair *Dirca* and *Ovidia* was relatively well supported with a bootstrap value of 79%. The association among the four species of *Stephanodaphne* as well as the two species of *Peddiea* was very strongly supported with a maximum bootstrap support of 100%.

In lineage three, *Gnidia* species found in lineage four of the combined plastid analysis were grouped with the genera, *Pimelea* and *Thecanthes* with significantly higher support of 99%. A lack in support was observed for the genus *Pimelea* due to low resolution, but included both species of *Thecanthes* with bootstrap support of 100%. Species found in the third *Gnidia* lineage of the combined plastid Fitch tree, were well embedded within *Pimelea*, including *P. rara* (94BP). High bootstrap support (100%) for the clade of *Gnidia coriacea*, *G. humilis*, *G. galpinii*, and *G. aff. subulata* could be observed.

The fourth *Gnidia* lineage included species previously classified under *Lasiosiphon*, with much higher support (98BP) than compared to the fifth lineage of the combined plastid tree. Two well supported sister clades: *Gnidia dumetorum*, *G. bojeriana*, *G. gilbertae*, *G. madagascariensis* and *Gnidia* sp. (100BP); and *G. caffra*, *G. calocephala*, *G. kraussiana* and *G. sericocephala* (97BP) were formed. Sister to the pair (98BP) was *Gnidia bakeri*, *G. danguyana* and *G. decaryana* (90BP).

4.4 Combined molecular analysis (Total evidence)
Trees obtained from the separate analysis of *rbcL*, *trnL-F* and ITS showed no strongly supported incongruent patterns and thus data matrices were directly combined. Of the 3014 included characters, 1886 were constant, 1128 (37.43%) variable and 789 (26.17%)
Species-level phylogeny

The Fitch analysis resulted in 1290 equally parsimonious trees with a TL of 3550 steps. The CI was low at 0.45, while the RI was higher at 0.77 (Table 4.1). SW produced 5710 trees of 2375.3396 weighted steps with a CI = 0.45 and RI = 0.77 (these were a subset of the Fitch trees: Fitch length = 3550 steps). The topology for both the trees in the Fitch and Bayesian analysis were exactly the same, therefore they could be combined and displayed on one tree (Figure 4.3).

The bootstrap values in the combined molecular analysis were much higher compared to the analyses of the combined plastid and the separate nuclear results. Thymelaeaceae were strongly supported as monophyletic (93BP, 92SW). In the Bayesian analysis a PP value of 0.96, almost equal to the bootstrap support in both the Fitch and SW analyses, was observed. *Dialyceras* and *Rhopalocarpus* (Sphaerosepalaceae) formed a well supported clade (99BP, 99SW, PP 1.0). The subfamilies Aquilarioideae and Gilgiodaphnoideae were successively sister to Thymelaeoideae with relatively strong support (89BP, 90SW and 73BP, 71SW respectively). Although the bootstrap support for Aquilarioideae was relatively high in both the Fitch and SW analyses, a much higher support was obtained in the Bayesian analysis with a PP of 1.0. The PP of 0.72 for Gilgiodaphnoideae was equal to the bootstrap support in both the Fitch and SW analyses. Sister to the rest of Thymelaeaceae was the fourth subfamily Gonystyloideae, including two clades: (I) Three species of *Octolepis*, viz. *O. dioica*, *O. dioica* fo. *oblancoelata* and *Octolepis* sp. (99BP, 98SW, PP 0.99); (II) Three representatives of *Lethedon*, viz. *L. balansae*, *L. cernua* and *L. aff. salicifolia* (76BP, 75SW) with more support in the Bayesian analysis (PP 0.98), *Arnhemia*, *Deltaria*, *Gonystylus* and *Solmsia*. Thymelaeoideae was well supported as being monophyletic (82BP, 84SW), and even stronger support was observed in the Bayesian analysis (PP 1.0). This subfamily included three supported clades. Clade (I) and (II) are successively sister to clade (III) (82BP, 84SW and 88BP, 89SW respectively). Both clades (I) and (II) obtained a higher PP of 1.0. Clade (I) received strong support (96BP, 98SW) and a maximum PP of 1.0, with *Craterosiphon* and *Synaptolepis* sister (95BP, 97SW), and *Enkleia* (74BP, 76SW), as well as *Dicranolepis* sister to this pair. Clade (II) was also well supported (94BP, 94SW) and had a higher PP of 1.0. There were two strongly supported sister lineages: *Daphne*, *Diarthron* and *Thymelaea* (88BP, 90SW) with stronger support in the Bayesian analysis (PP 1.0); and *Stellera* and *Wikstroemia* (97BP, 98SW) with a maximum PP of 1.0, with *Edgeworthia* sister to this pair (96BP, 96SW), also with a higher PP of 1.0. In clade (III), the bootstrap support was moderate in the Fitch and SW analyses (70BP, 71SW) and the PP high. The largest genus in this clade, *Gnidia*, was again polyphyletic, with at least four lineages: 1) This clade contains some of the *Gnidia* species sister to *Drapetes muscoides* (71BP, 77SW) with stronger support in the Bayesian analysis (PP 1.0). 2) *Gnidia pinifolia* and *G. racemosa* grouped with *Struthiola* (100BP,
4. Species-level phylogeny
4. Species-level phylogeny

100SW, PP 1.0).  3) *Gnidia pilosa* (previously known as *Englerodaphne pilosa*), together with *Gnidia* species, as sister to the genus *Pimelea* (94BP, 93SW) and stronger support in the Bayesian analysis (PP 1.0).  4) Taxa, previously known as *Lasiosiphon*, grouped together with a few *Gnidia* species with very strong support of 100% for bootstrap values of both the Fitch and SW analysis, and a PP value of 1.0.

The first *Gnidia* lineage as well as the genus *Lachnaea* were successively sister to *Passerina* with low bootstrap support to a very high PP (PP 0.95 and 57BP, 59SW, PP 1.0 respectively).  *Passerina* was strongly supported as monophyletic (90BP, 90SW) whereas *Lachnaea* received weaker bootstrap support (69BP, 70SW).  Both genera received a strong PP of 1.0.  Sister to the rest was the second lineage of *Gnidia* with relatively high bootstrap support (86BP, 87SW) but with no support in the Bayesian analysis.  The position of the third *Gnidia* lineage, the *Peddiea* clade, and the grouping with *Dirca* and *Ovidia* was unresolved.  The monophyly of the representatives of *Peddiea* were very well supported (100BP, 100SW, PP 1.0).  The sister pair *Dirca* and *Ovidia* was relatively well supported with bootstrap values of 71% in both the Fitch and SW analyses and a higher PP of 0.99.

Sister to the rest was the monophyletic *Stephanodaphne* (100BP, 100SW, PP 1.0), the fourth *Gnidia* lineage and the sister pair *Dais* and *Phaleria* with moderate bootstrap support, but a very high PP (PP 0.98; 60BP, 64SW, PP 1.0 and 70BP, 71SW, PP 1.0).  There was strong support for the clade of *Dais* and *Phaleria* (99BP, 99SW) with an even higher PP of 1.0.

On the specific level, very high support (100BP, 100SW, PP 1.0) was received for the monophyly of the *Gnidia* species in the first lineage, including *G. cf. anomala*, *G. denudata*, *G. fastigiata*, *G. geminiflora*, *G. renniana* and *G. aff. renniana*.  2) Strong support was obtained for the monophyly of *Struthiola* (94BP, 96SW) with a higher PP of 1.0, with *Gnidia pinifolia* and *G. racemosa* sister to the genus (100BP, 100SW, PP 1.0 and 100BP, 100SW, PP 1.0 respectively).  3) Moderate support for the monophyly of *Pimelea* (79BP, 80SW, PP 1.0), including the genus *Thecanthes* with very strong support (99BP, 99SW) and a PP of 1.0 for *T. punicea* and *T. sanguinea* as a sister pair.  Although the bootstrap support was very low in both the Fitch and SW analyses (55BP, 53SW), a high PP of 0.98 was obtained for *Gnidia pilosa* being successively sister to *Pimelea*.  Three *Gnidia* clades were also observed within this lineage.  Clade (I) consisting of *Gnidia aberrans*, *G. caniflora*, *G. scabrida*, *G. setosa*, *G. singularis* and *G. wikstroemiana* received high support (96BP, 95SW) and an even higher PP of 0.99.  A low PP of 0.56 and no bootstrap support in both the Fitch and SW analyses was obtained for clade (II), including *Gnidia phaeotricha*, *G. squarrosa* and *G. wilmsii*.  Clade (III) was also highly supported (90BP, 90SW, PP 1.0) and includes *G. coriacea*, *Gnidia galpinii*, *G. humilis*, *G. cf. juniperifolia*, and *G. aff. subulata*.  4) A mixed lineage forming a well supported clade (100BP, 100SW, PP 1.0) with two sister
Species-level phylogeny

clades: *Gnidia bojeriana*, *G. dumetorum*, *G. gilbertae*, *G. madagascariensis* and *Gnidia* sp. (82BP, 83SW) and no support in the Bayesian analysis; and *G. caffra*, *G. calocephala*, *G. kraussiana* and *G. sericocephala* (94BP, 95SW, PP 1.0), with the position of *G. glauca* still unresolved. Sister to the rest was a group consisting of *Gnidia bakeri*, *G. danguyana* and *G. decaryana*, relatively well supported (84BP, 87SW) with higher support in the Bayesian analysis (PP 0.95).