Appendix S5: NCBI GenBank Harvest, 4/5/2016

Search strings taxa and gene regions

Taxa

Winteraceae: txid3417[Organism:exp] …
Canellaceae (sister group): txid3424[Organism:exp]

Nuclear gene regions

18S rDNA: … AND 18S [title] NOT "internal" [title]
ITS region of the 35S rDNA: … AND "internal" [title] NOT "4.5S" [title]

Plastid gene regions

*trn*H-*psb*A spacer: … AND "psbA-trnH" [title][1]
*trn*K/*mat*K region: … AND ("trnK" [title] OR matK [gene]) NOT rbcL [gene]
*rps*16 intron: … AND rps16 [gene] NOT rbcL [gene]
*atp*B gene: … AND atpB [gene] NOT rbcL [gene]
*rbc*L gene: … AND rbcL [gene] NOT atpB [gene]
trnTLF region: … AND ("trnL" [gene] OR "trnL-trnF" [title])

[1] Misses one accession (AB331299) covering the entire strand of *rps*12 to *psb*A (including *trn*H) genes.
Note: Search strings avoid capturing of the completely sequenced plastome of *Drimys granadensis* (NC_008456/ DQ887676) in the harvests. Corresponding sequence portions were manually added to the FASTA files as first accession and used as reference for annotation of plastid gene regions.

Processing

Genebank flatfiles transferred into FASTA-format using GBK2FAS (Göker *et al.*, 2009, see batch file for options). Auto-alignments done with MAFFT v. 7.273 (Katoh & Standley, 2013), optimal algorithm chosen by the programme; option "adjustdirection" was used for plastid regions using the strand from the completely sequenced plastome as first sequence and reference.

Dataset	Taxon set	Algorithm	Further remarks
[1] 18S	Winteraceae + Canellaceae	L-INS-i	
[2] ITS	Winteraceae	L-INS-i	
[3] trnH-psbA	Winteraceae	L-INS-i	
[4] trnK/matK	Winteraceae + Canellaceae	FFT-NS-i	
[5] rps16	Winteraceae + Canellaceae	L-INS-i	
[6] atpB	Winteraceae + Canellaceae	L-INS-i	
[7] rbcL	Winteraceae + Canellaceae	L-INS-i	
[8] trnTLF	Winteraceae + Canellaceae	FFT-NS-i	Very limited data on *trn*T-*trn*L spacer

All alignments were inspected by eye, and their ends were trimmed. As a rule we truncated the alignments for less than four accessions on the 5' and 3' ends, with exceptions of the 3' 18S to be able to check for overlap with the start of the ITS sequences. As the main inter-generic relationships are considered to be resolved, we followed the commonly retrieved sequence for our alignments, i.e. *Takhtajania, Tasmannia, Drimys, Pseudowintera,* and *Zygogynum* s.l. (including *Belliolum, Bubbia, Exospermum, Zygogynum*).

Nuclear-encoded 18S rDNA

General—As usual, the 18S is sequentially very conserved between members of the Canellales. Far the most mutation patterns reflect the two families, within-family divergence is essentially non-existent.

This is exemplary illustrated in the sequence of the terminal part of stem 49 at the 3' end of the 18S, where six family-conserved mutations differentiate between the Canellaceae and the Winteraceae. Occasional single-nt indels appear to be due to the age and quality of the sequences rather than representing genuine mutational patterns. The same holds for unique point mutations, observed occasionally in a few sequences (e.g. AAT at pos. 279–281 and GCCC at pos. 320–323 in *Pseudowintera colorata* vs. ATT and GGCC in all other accessions).
Curation—Lacking comparative data we refrained from correcting for putative sequencing/editing artefacts. Each of the indicated sequencing gaps in accession X63313 (*Tasmannia purpurascens*) are in fact at least double as long as indicated, and have been accordingly filled by missing data symbol ("?").

Nuclear-encoded ITS region (ITS1, 5.8S, ITS2)

General—As in many Asteraceae, the spacer ITS1 and ITS2 have a high GC content, which facilitates the identification of potential pseudogenous sequences. In the bird's eye view, the main groups of ITS sequences become obvious: Winteraceae vs. Canellaceae, the latter with much higher intergeneric divergence. Alignment gaps in the auto-alignment are mostly due to the sequential differences between both families, and between the genera of the Canellaceae. Effectively, the sequential difference between Winteraceae and their sister group seen both in the ITS1 and ITS2 is too high to be meaningful for phylogeny estimation. Thus, the alignment was split into two blocks, one for each family. Within the Winteraceae, the most distinct type is that of *Takhtajania perrieri* (AY004129; no voucher information included in GenBank flatfile), which is substantially more AT-rich in the spacer but not the covered 5.8S rDNA gene region (the 5.8S is highly similar to those of all other Winteraceae). *Tasmannia* and *Drimys* are visibly distinct from *Pseudowintera* and the *Zygogynum* alliance (including both purported "paralogues" REF)
Curation—The dataset includes several pairs of ITS1 and ITS2 sequences from the same voucher. These were concatenated. At the start of the alignment, we corrected for the alignment between *Zygogynmum* (CCC) and the remainder (CC) at pos. 21ff; the C-motif shortly after was left-aligned (pos. 24ff). Alignment at pos. 249–263 was adjusted for consistency and to minimise gaps. Pos. 462–466 were right-aligned.
Problematic data—The last c. 50nt of accession EU683899 (*Drimys andina*; Rodriguez al. 2038 CONC/OS) are unlike any other Winteraceae, and MEGABLAST found no other hit. Furthermore, the sequence shows a prominent (sequencing?) gap in the 2nd half of the ITS1. Its ITS2 counterpart (accession EU683900) and the rest of the sequence is inconspicuous and nearly identical to the other two accessions available of the species, providing no additional information. Thus, both sequences were tentatively omitted.

Plastid *trn*H-*psb*A intergenic spacer

General—In the complete reference plastome (*Drimys granadensis*; NC_008456) the *trn*H/*psb*A region is 1557 nt long, with two-third representing the *psb*A gene (1062 nt). The *trn*H gene is 75 nt long. The other sequences cover most (c. 90%) of the spacer, missing ≥ 17 nt at the 5' end, and ≥ 3 nt at the 3' end of the spacer; all accessions are from Winteraceae. The most distinct sequence is that of *Takhtajania perrieri* including 15 unique site mutations and a deletion (pos. 61–74) at otherwise length- and sequence-conserved positions. The remainder is length-conserved, with a three 4-/5-nt long and unique (limited to a single accession) duplications/deletions (pos. 163–166; 225–229; 305–309). The A-dominated motif at pos. 271ff shows some length and sequence divergence within Winteraceae (Table S5.1), which appears to be (currently only limited taxon coverage) genus-specific. Interestingly, a shared general type is found in some *Drimys* and *Zygogynum*. The *Takhtajana* sequence is remarkably different from all other Winteraceae between pos. 264–308, but includes the A-dom. motif.
Curation—Occasionally the sequences show signs of sequencing/editing artefacts (e.g. GKGK/GTGK instead of GTGT in all other accessions at pos. 67–70). This may also apply to mutations unique to *Takhtajania*. Lacking comparative data, all data were kept as-is. The As in the A-dominated length-polymorphic region were left/right-aligned for better visibility.

Table S5.1. Diagnostics of the A-dominated *trn*H-psb*A* motif of extant Winteraceae.

Motif variant	Motif sequence (unique features in bold)	Species (number of accessions)
Takhtajania	A_2T**AG[TAAA]$_3$**A_4 followed downstream by completely different sequence than in other Winteraceae	*T. perrieri* (1)
Drimys variant A	$A_2TA_{12–13}$	*D. granadensis* (3), *D. angustifolia* (1), *D. brasiliensis* (4)
Drimys variant B	A_2TA_4-**GTCAATC**-A_{12}	*D. andina* (2), *D. confertifolia* (1), *D. winteri* (3)
Pseudowintera	A_2T**A$_5$-GTAAAT**-A_6	*P. colorata* (1)
Zygogynum variant A	A_2TA_{11}	*Z. comptonii* (1), *Z. tanyostigma* (1)
Zygogynum variant B	A_2TA_2-**AAAT**-A_6	*Z. baillonii* (1)
Zygogynum variant C	A_2TA_2-**GTAAAT**-A_{12}	*Z. vinkii* (1)

Plastid *trn*K/*mat*K region

General—Five of the Winteraceae and many of the Canellaceae accessions (accession nos KP...) include also the upstream *psb*A-*trn*K intergenic spacer, which were not considered for further analyses. The accessions covering the *trn*K intron downstream of the *mat*K gene terminate ≥ 19 nt before the start of the 3' *trn*K exon. Isolated (typically limited to a single accession, species) length-polymorphism (duplications, rarely deletions) can be found in the *trn*K intron regions; the deletion of part of the AT-motif at pos. 233ff in *Takhtajania,* a few Winteraceae, and all but one Canellaceae accession, appears to be an editing/concatenating artefact (see **Problematic data**). Phylogenetically linked indels may be the duplications at pos. 2484ff (two species of *Tasmannia*) an 2503ff (in all four accessions of the *Zygogynum* alliance), and the lack of a 16-nt long duplication found in all Canellaceae except for *Cinnamosma* at pos. 2204ff.
The *mat*K gene is generally length-conserved, except for a duplication of the 3rd last codon in Canellaceae; and a 2-codon insertion seen in three accessions (out of four) of *Cinnanosma fragrans* (KP407478, …79, and …87) and two labelled as spec. indet. (KP407501/…02)
Curation—The start of accessions including only the downstream intron (accession nos. FJ86xxxx) were not properly aligned with the rest of the data and manually adjusted. Accessions with corresponding intron parts were concatenated. The *mat*K gene was corrected for incomplete codons.
Problematic data—The partial deletion of the AT-motif at pos. 233–255 within the *trn*K intron upstream of the *mat*K gene found in *Takhtajania* (KP407452) is also represented in other Winteraceae accessions by the same authors (*Tasmannia insipida,* KP407453; *Tasmannia lanceolata* KP407454; *Drimys winteri,* KP407457). No such deletion is seen in any other Winteraceae accession, even if the same species was sequenced. For instance, *T. lanceolata* is represented by over 30 accessions, none showing the same deletion than the KP40xxxx accessions. This part is, however, missing in nearly all Canellaceae accessions. The only other *trn*K accession available for Canellaceae is KF408223, which shows no gap but an AT-motif also found in Winteraceae. Hence, we assume that the gap in the KP40xxxx accessions is an editing/concatenating artefact and replaced the purported sequence gap by missing data symbol. Further editing issues can be found: In accession KP407453, a sequence of polymorphic base calls in the *mat*K gene (AARKYYKKTTTTTTWARSGA instead AAAGTCTGTTTTTTTAAGGA), the according part of the sequence was blanked out (i.e. replaced by missing data symbol "?").

Plastid *rps*16 intron

General—The accessions start 35 nt downstream of the *rps*16 5' exon and terminate 66 nt before the start of the 3' exon. Level of length-polymorphism is higher than in the *trn*K intron, but for most part alignment of both families is unproblematic. The two Canellaceae accessions differ from each other to

the same degree than from the Winteraceae, main differences are seen towards the 3' end of the intron. Within Winteraceae, the *rpl*16 intron of *Takhtajania* is visibly the most distinct.
Curation—One accession (FJ554661) includes an overestimated missing portion, others show too few "?" (FJ554660, FJ554662, FJ554663), the length of the missing data portions were adapted based on the reference. Alignment of the A-dominated motif at pos. 171ff was adjusted for consistency. The TGG was moved to the 5' end of the according deletion at pos. 575ff in *Takhtajania.*

Plastid *atp*B gene

General—The accessions cover the entire gene regions (or near so) of all genera included in the harvest. The gene is length-conserved comprising 499 codons according to the reference. Most site variation is linked to point mutations supporting the mutual monophyly of Canellaceae and Winteraceae. Variation is limited within the latter.
Curation—The alignment was trimmed for incomplete codons.
Problematic data—The second *Drimys* accession (AF093425, *Drimys winteri*), in addition to the complete plastome reference, is virtually identical to the two accessions available for *Tasmannia*. The latter differ by several mutation from the family's consensus and the reference sequence of *D. granadensis* (NC_008456). According information compiled for Tropicos.org (2016), *D. winteri* is now a synonym of *D. granadensis* var. *mexicana.* With respect to the phylogenetic position of the two genera (and their geographic disjunctness), it is unlikely that one species/subspecies of *Drimys* should have retained an *atp*B gene shared with *Tasmannia*, whereas the other evolved a different type [Note that *rbc*L of both *Drimys* (sub)species is identical, and differs by several point mutations of that from *Tasmannia* and other Winteraceae.] Thus, the *Drimys winteri atp*B accession in gene bank is probably mislabelled and has been tentatively excluded it from further analyses.

Plastid *rbc*L gene

General—Most accessions cover a part of the gene. Such as the *atp*B gene, the *rbc*L gene is universally length-conserved comprising 479 codons. As for the other gene, most site variation is linked to point mutations supporting the mutual monophyly of Canellaceae and Winteraceae, but intra-family divergence is higher than seen in the *atp*B data.
Curation—The alignment was trimmed for incomplete codons.
Problematic data— One accession of *Cinnamodendron ekmanii* (AJ235776, Qiu_47067_NCU**)** differs from other accession of the species and genus by a number of point mutations and is 100% identical to accessions of *Canella winterana* including one uploaded a year later by the same researcher, AJ131928. An upload error is most likely, the accession was excluded from further analyses.

Plastid trnLLF region

Note on *trn*T-*trn*L spacer—This part is only represented (partly) in up to four accessions (including the reference sequence), sequenced for one Winteraceae genus (*Drimys*) and one Canellaceae (*Canella winterana*), hence, not useful for further analyses. An interesting find is that *Drimys* and *Canella* are very different in this spacer, preventing straightforward alignment for about 50% of the spacer region, which is in strong contrast to the situation in the *trn*L-*trn*F spacer.
General—The final alignment includes the *trn*L intron and 3' exon, the *trn*L-*trn*F intergenic spacer, and most of the *trn*F gene (except for 13 nt). Most gaps are inflicted by combining Winteraceae and Canellaceae in one matrix; nevertheless, the intron and most of the downstream spacer are straightforwardly aligned by MAFFT. Differentiation between both families is best seen in the AT-rich *trn*L intron portion at pos.251–292. The spacer seems to be, however, substantially shorter in Canellaceae than in Winteraceae (c. 100 nt at the 5' end are missing; possibly not a genuine feature but a data artefact). Within-Winteraceae length-polymorphism in the *trn*L-*trn*F spacer is limited to the multi-A and multi-T motives at pos. 587ff and 636ff. As in the case of other gene regions (nuclear and plastid), *Takhtajania* is readily distinct from the remainder of the family in the intron, its spacer is however relatively similar to that of (some) *Tasmannia*.
Curation—Accession AY145347 includes a large missing data portion in the sequence which seems to have no justification and was omitted. Accession KJ412188 terminates with a 72-nt long insertion at the end of the *trn*L-*trn*F spacer, which, for the most part is a duplication of the 3' end of the *trn*L-*trn*F

spacer. This could be a sequencing artefact and the 3' end of the accessions was tentatively blanked out. Very rare, singleton single-nucleotide length-polymorphisms (e.g. C instead of CC in all other accessions) is seen in a few of the newer sequences that appear to be linked to sequence quality issues and were corrected for in KC428458, KJ412185, KJ412184)
Problematic data—Accession KJ412184 (*Bubbia queenslandica*) differs from the other accession of the species and the general sequence type shared by *Zygogynum* and allies. The sequence can be readily identified as *Tasmannia (T. lanceolata*) and was omitted from the data set.

Phylogenetic analyses

Pre-analysis

We inferred fast, simple maximum likelihood (ML) trees and established nonparametric bootstrap (BS) support based on the untransformed alignments for each covered gene region with the Windows-executable of RAxML v. 7.2.8 (Stamatakis, 2006a; Stamatakis *et al.*, 2008). The greatest amount of data are available for the plastid *trn*K/*mat*K and trnLLF regions, the latter showing also a relatively high sequence diversity as illustrated by the number of distinct alignment patterns under ML (Table S5.2). RAxML was set to infer a ML tree under the general-time-reversible substitution model (Rodriguez *et al.*, 1990) allowing for site-specific rate variation modelled using the per-site approximation (Stamatakis, 2006b) of the Gamma (Γ) distribution with 25 per-site rate categories; number of necessary bootstrap pseudoreplicates was determined using the extended majority rule consensus bootstop criterion (Pattengale *et al.*, 2009).

Table S5.2. Matrix and run statistics for the untransformed data.
#taxa = number of taxa (= individual accessions); #char = number of characters; #LD = number of literal duplicates of other accessions; #DAP = number of distinct alignment patterns (unpartitioned analysis); gappyness = proportion of completetly undetermined characters (i.e. missing, gaps); OET = overall execution time (using two parallel threads on a laptop equipped with Intel *Pentium i7* processors) in seconds.

Gene region	Matrix dimensions		#LD	#DAP	Gappy-ness	OET [s]	Includes data on
	#taxa	#char					
18S rDNA	16	1794	None	149	10%	~200	Winteraceae + Canellaceae
ITS	68	730	8	263	13%	~725	Winteraceae
*trn*H-*psb*A	20	420	2	92	8%	~275	Winteraceae
*trn*K/*mat*K	135	2637				~1725	Winteraceae + Canellaceae
*rps*16 intron	21	805	2	147	8%	~250	Winteraceae + Canellaceae
*atp*B	13	1497	2	101	5%	~75	Winteraceae + Canellaceae
*rbc*L	42	1437	8	272	23%	~450	Winteraceae + Canellaceae
trnLLF	106	1020	34	336	20%	~1625	Winteraceae + Canellaceae

The inference results confirm above statements based solely on visual comprehension of the underlying alignments.
18S rDNA—These data are useless for intra-family inferences; the mutual monophyly of Winteraceae and Canellaceae is well represented and accounts for far the most of the distinct alignment patterns.
ITS region—These data support the conservative genus view, supporting five genera, each of which characterised by highly distinct ITS regions (unambiguous BS support under ML: $BS_{ML} \geq 99$), ordered in the following according to their putative phylogenetic sequence (assuming a Winteraceae root recognising *Takhtajania* as the first diverging genus): *(i) Takhtajania, (ii) Tasmannia, (iii) Drimys, (iv) Pseudowintera, (v) Zygogynum* including *Bubbia* and *Exospermum*. The two purported "ITS paralogues" of the latter form respective clades, which do not receive unambiguous support (BS_{ML} = 72 and 54)

trn*H-*psb*A intergenic spacer—Takhtajania*** is most distinct, *Drimys* well separated from the *Zygogynum* alliance. The latter shows little phylogenetic structure with the only sure non-*Zygogynum* accession (*Pseudowintera colorata*, FJ539196) being nested in a *Zygogynum* tree. A potential *Bubbia* (*Bubbia/Zygogynum comptonii*, FJ539199[1]) is placed as sister to the *Zygogynum* s.str. and *Pseudowintera.* Lacking data on the fifth generic lineage, *Tasmannia*, and being much undersampled for *Zygogynum* and allies, the use of this region is so far very limited and was tentatively excluded from further analyses.

***trn*K/*mat*K region—**As commonly observed, this plastid gene region has a high capacity to resolve and recognise plant genera or high plant taxa, and this holds also for the Winterales. Both families are well separated by a long internode (branch) in the tree, and accessions group according to their generic affiliation, the genus' root branches receive high to unambiguous support (BS_{ML} usually ≥ 98, *Warburgia*: $BS_{ML} = 89$, *Bubbia* and *Zygogynum* only represented by singletons, unambiguously supported as sibling accessions). Exception are accessions linked with the Canellaceae genus *Cinnamodendron*, which either share the *trn*K/*mat*K genotype of *Capsicodendron dinisii* (*Cinnamodendron axillare, C.occhionianum,* several accessions not determined to species level: *C.* sp. Barros et al. s.n., *C.* sp. De Assis s.n., *C.* sp. B SM-2015, *C.* sp. DP143) or *Pleodendron* (*Cinnamodendron corticosum, C. cubense, C. ekmanii*[2]*, C.* sp. 'Haiti'/Salazar & Vilmond 2417, *C.* sp. DP136). Data from *Cinnomodendron* was treated accordingly in subsequent analyses (e.g. when computing genus-consensus sequences).

Within the Winteraceae subtree, the phylogenetic sequences as indicated based on the ITS data is fully recovered, but the recognition of *Takthajania* as sister to the remainder receives only low support.

***rps16* intron—**These data show only limited species coverage, and so far no data have made available for *Tasmannia*, so only four of the five main generic lineages in Winteraceae are covered. Canellaceae have just been sampled for two accessions used as outgroups for Winteraceae. Where more than a single accession is available (*Drimys*, *Zygogynum*), the genera show high coherence, with the potential *Bubbia* (*Zygogynum*) *comptonii* nested in a *Zygogynum* clade.

***atp*B gene—**Even though of limited coverage, these data confirm the phylogenetic sequence for the Winteraceae: the mutual monophyly of both families is reflected by a long internode (relatively as prominent as seen in the 18S tree), *Takhtajania* is resolved as first divergent lineage, followed by *Tasmannia,* whereas the single *Drimys* accession is indistinct in this gene region from the two accessions of *Pseudowintera* and the single accession of *Zygogynum.*

***rbc*L gene—**Both Canellaceae and Winteraceae are better represented by *rbc*L data than *rps*16 intron and *atp*B data. The general structure of the tree is in accordance with that inferred based on the *trn*K/*mat*K data, genera are well distinct (with exception of *Cinnamodendron* as in the case of *trn*K/*mat*K), but root branches typically receive lower support.

trnLLF region—Again, accessions of the same genera are placed within the same subtree with the exception of *Cinnamodendron.* The Winteraceae are characterised by substantial generic roots and very low intra-generic divergence. *Pseudowintera* is an exception, accessions of this genus form a soft polytomy in which the *Zygogynum* alliance subtree is rooted; support is low for the root of the *Zygogynum* alliance subtree as well as its common root with *Pseudowintera.* This situation is comparable to some degree with that seen in the comprehensive ITS tree, where *Pseudowintera* shows generally shorter root-tip distances than its siblings of the *Zygogynum* alliance.

Despite resolution issues towards the leaves and deepest intra-family relationships the high phylogenetic coherence of genera and species (the latter never forming more than a single high-supporting clade) allows using a consensus sequence strategy in order to produce alignment with as little missing data as possible, forming the basis for the phylogenetic framework used herein. The phylogenetic sequence well-exhibited in the ITS dataset would be in agreement with all other data, and concatenation of nuclear and plastid data is unproblematic.

---

[1] Accessions of this species covering different gene regions are stored in gene banks either as *Bubbia* or *Zygogynum.*

[2] Not to be confused with the *Pleodendron* species (*P. ekmanii* Urb.) using the same epithet and also included in the harvested data.

Inferring a species and genus-level phylogeny

Using the programme G2CEF (Göker & Grimm, 2008), we computed strict species and genus-level consensus sequences based on the originally harvested matrices.

The basic species consensus matrix covers seven gene regions sequenced for the Winteraceae: *(i)* the 18S rRNA gene (18S rDNA) and ITS region[3] of the nuclear-encoded 35S rDNA cistron (original length in alignment: 2512 nt); and – from the plastome (listed as found in the plastome) – the *(ii) trn*H-*psb*A intergenic spacer (420 nt); *(iii)* the *trn*K intron including the *mat*K gene (2637 nt; including 513 codons); *(iv)* the *rps*16 intron (805 nt); *(v)* the *atp*B gene (1497 nt = 499 codons); *(vi)* the *rbc*L gene (1437 nt = 479 codons); and *(vii)* the trnLLF region including the *trn*L intron, 3' exon, and the *trn*L-*trn*F intergenic spacer (1015 nt). Of these seven gene regions, six are (partly) represented also by species of the Canellaceae (available ITS data not aligned; no *trn*H-*psb*A data available).

The basic genus-level consensus matrix covers the same gene regions and includes the following taxa: five Winteraceae generic lineages: *Drimys, Pseudowintera, Takhtajania, Tasmannia, Zygogynum* s.l. (including *Bubbia, Belliolum, Exospermum,* and *Zygogynum*); in case of the Canellaceae, *Cinnamodendron* accessions were either treated as *Capsicodendron* (= *Cinnamodendron* s.str.[4]) or *Pleodendron* as outlined in the preceding section.

ML trees were inferred and branch support established via nonparametric BS as outlined in section *Pre-analysis.* To cover as much of the model-space as possible, full-partitioned (treating each sequence region and codon position as individual data partitions) and unpartitioned analyses were run. Competing support patterns in the BS pseudoreplicate samples were visualised using bootstrap support networks (Grimm *et al.*, 2006; Schliep *et al.*, 2017) generated with SPLITSTREE v. 4.13.1 (Huson & Bryant, 2006). The results of the analysis are included in the online supporting archive.

[3] Two 18S accessions cover the entire 3' end of the 18S rDNA, a part also included in many ITS accessions. The according overlap was used to clip the 18S subalignment so that it is non-overlapping and directly preceeding the ITS subalignment.

[4] The type species of *Cinnamodendron* is *C. axillare* Endl., published 1840 (Tropicos.org 2016), which shows *Capsicodendron*-type sequences. The name *Cinnamodendron* would have priority over *Capsicodendron* according to the Botanical Code.

References

Göker, M., & Grimm, G.W. (2008) General functions to transform associate data to host data, and their use in phylogenetic inference from sequences with intra-individual variability. *BMC Evolutionary Biology*, **8**, 86.

Göker, M., García-Blázquez, G., Voglmayr, H., Tellería, M.T., & Martín, M.P. (2009) Molecular taxonomy of phytopathogenic fungi: a case study in *Peronospora*. *PLoS ONE*, **4**, e6319.

Grimm, G.W., Renner, S.S., Stamatakis, A., & Hemleben, V. (2006) A nuclear ribosomal DNA phylogeny of *Acer* inferred with maximum likelihood, splits graphs, and motif analyses of 606 sequences. *Evolutionary Bioinformatics*, **2**, 279-294.

Huson, D.H., & Bryant, D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254-267.

Katoh, K., & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.

Pattengale, N.D., Masoud, A., Bininda-Emonds, O.R.P., Moret, B.M.E., & Stamatakis, A. (2009) How many bootstrap replicates are necessary? In: S. Batzoglou (Ed.^Eds.), *RECOMB 2009.* (Vol. 5541, pp. 184–200) Berlin, Heidelberg: Springer-Verlag.

Rodriguez, F., Oliver, J.L., Marin, A., & Medina, J.R. (1990) The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology*, **142**, 485-501.

Schliep, K., Potts, A.J., Morrison, D.A., & Grimm, G.W. (2017) Intertwining phylogenetic trees and networks. *Methods in Ecology and Evolution*, 10.1111/2041-210X.12760

Stamatakis, A. (2006a) RAxML-VI-HPC: Maximum-Likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688-2690.

Stamatakis, A. (2006b) Phylogenetic models of rate heterogeneity: A high performance computing perspective. *Proceedings of 20th IEEE/ACM International Parallel and Distributed Processing Symposium (IPDPS2006), High Performance Computational Biology Workshop.* (ed by, pp. [on CD, no page nos]. Rhodos, Greece, April 2006.

Stamatakis, A., Hoover, P., & Rougemont, J. (2008) A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology*, **57**, 758–771.

Tropicos.org (2016) Tropicos® database. In:Missouri Botanical Garden.