Structural Plastome Evolution in Holoparasitic Hydnoraceae with Special Focus on Inverted and Direct Repeats

Matthias Jost [1,*], Julia Naumann[1], Jay F. Bolin[2], Carlos Martel [3,4], Nicolás Rocamundi[5], Andrea A. Cocucci[5], Darach Lupton[6,7], Christoph Neinhuis[1], and Stefan Wanke [1,8,*]

[1]Institut für Botanik, Technische Universität Dresden, Dresden, Germany
[2]Department of Biology, Catawba College, Salisbury, NC, USA
[3]Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK
[4]Instituto de Ciencias Ómicas y Biotecnología Aplicada, Pontificia Universidad Católica del Perú, Lima, Peru
[5]Laboratorio de Ecología Evolutiva y Biología Floral, IMBIV, CONICET and Universidad Nacional de Córdoba, Córdoba, Argentina
[6]Oman Botanic Garden, Seeb, Sultanate of Oman
[7]National Botanic Gardens, Glasnevin, Ireland
[8]Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City, Mexico

*Corresponding authors: E-mails: matthias.jost@tu-dresden.de; stefan.wanke@tu-dresden.de.

Accepted: 19 May 2022

Abstract

Plastome condensation during adaptation to a heterotrophic lifestyle is generally well understood and lineage-independent models have been derived. However, understanding the evolutionary trajectories of comparatively old heterotrophic lineages that are on the cusp of a minimal plastome, is essential to complement and expand current knowledge. We study Hydnoraceae, one of the oldest and least investigated parasitic angiosperm lineages. Plastome comparative genomics, using seven out of eight known species of the genus *Hydnora* and three species of *Prosopanche*, reveal a high degree of structural similarity and shared gene content; contrasted by striking dissimilarities with respect to repeat content [inverted and direct repeats (DRs)]. We identified varying inverted repeat contents and positions, likely resulting from multiple, independent evolutionary events, and a DR gain in *Prosopanche*. Considering different evolutionary trajectories and based on a fully resolved and supported species-level phylogenetic hypothesis, we describe three possible, distinct models to explain the Hydnoraceae plastome states. For comparative purposes, we also report the first plastid genomes for the closely related autotrophic genera *Lactoris* (Lactoridaceae) and *Thottea* (Aristolochiaceae).

Key words: *Hydnora*, *Prosopanche*, holoparasite, heterotrophy, minimal plastome, structure.

Significance

Plastome condensation during adaptation to a heterotrophic lifestyle is generally well documented and lineage-independent evolutionary models have been derived. Controversially discussed are the putative evolutionary trajectories resulting in structures of plastomes close to a proposed "minimal plastome." Three evolutionary models are discussed to explain the large repeat content of newly assembled, highly condensed plastomes of holoparasitic Hydnoraceae. One of the models assumes the regain of the quadripartite plastome structure after ancestral loss, which has only been reported for two other seed plant lineages to date. An alternative model discusses lineage-specific inverted repeat condensation and translocation instead. Furthermore, we report the loss of all but one intron in the Hydnoraceae plastomes.

Introduction

The plastid chromosome of land plants is, on average ~150 kb in length (Daniell et al. 2016; Weng et al. 2017) with extreme size differences between the two ends of the size spectrum. The largest plastid genome to date has been reported for *Pelargonium* (~242 kb, Weng et al. 2017) and the smallest for the heterotrophic genus *Pilostyles* (~11 kb, Bellot and Renner 2016). Typical plastid chromosomes show a quadripartite structure. Two single-copy regions, namely the large (LSC) and small (SSC) single-copy region are separated by two copies of an inverted repeat (IR) (Palmer 1985; Mower and Vickrey 2018). Around 100–120 unique genes are encoded on plastomes of autotrophic plants, which are mostly related to photosynthesis, carbon fixation, or involved in their transcription and translation (Wicke et al. 2011). Plastomes are considered highly conserved with respect to gene order, content and structure (Palmer 1985), with many extremes and deviations from the norm reported for mycoheterotrophic, and parasitic plants including: (1) record-setting smallest known plastomes sizes exclusively found in parasitic and mycoheterotrophic plants (Bellot and Renner 2016; Petersen et al. 2018; Su et al. 2019), (2) the most extreme plastome A + T nucleotide bias reported in parasitic *Balanophora* (Balanophoraceae) with only ~12% G + C nucleotides (Su et al. 2019) compared with an average of 34–39% (Cai et al. 2006), or (3) the plastome of *Pilostyles aethiopica* (Apodanthaceae) with a total of only five encoded genes (Bellot and Renner 2016). Hypothetically culminating in a complete loss of the plastid chromosome in endoparasitic Rafflesiaceae (Molina et al. 2014; Cai et al. 2021).

Apart from the abovementioned differences in genome size, nucleotide composition, and gene content, structural diversity such as changes of gene order and IR loss can be observed. The loss of one IR copy is not uncommon in highly condensed plastomes (Wicke et al. 2013; Bellot and Renner 2016; Arias-Agudelo et al. 2019; Su et al. 2019; Jost et al. 2020), but not exclusive to the heterotrophic lifestyle (Palmer et al. 1987; Lavin et al. 1990; Wicke et al. 2011; Zhu et al. 2016; Jin et al. 2020). Comparative research on IR-containing and IR-lacking plastomes suggests a positive influence of the IR presence on structural plastome stability (Birky and Walsh 1992; Maréchal and Brisson 2010). IR losses or extreme reduction, on the other hand, seem to correlate with an increased number of rearrangements in some lineages (Palmer and Thompson 1982; Palmer et al. 1987; Wu et al., 2011; Wu and Chaw, 2016). However, not all IR-lacking plastomes show a higher degree of structural changes (Jansen et al. 2008; Sanderson et al. 2015; Blazier et al. 2016), casting doubts on IRs being a major factor preventing plastomes from structural changes. Current studies suggest a putative involvement of short repeat sequence accumulation as an additional factor driving homologous recombination (Wu et al. 2011; Mower and Vickrey 2018). As much as the adaptation to a heterotrophic lifestyle seems to follow lineage-independent patterns (Barrett and Davis 2012; Wicke et al. 2013; Naumann et al. 2016; Wicke et al. 2016; Graham et al. 2017), differences among the spectrum of highly condensed plastomes can be observed.

Understanding heterotrophic lineages and exploring evolutionary events at the cusp of a putative minimal plastome is essential to complement and expand current knowledge. One of these families is Hydnoraceae, with only two plastomes published to date (Naumann et al. 2016; Jost et al. 2020). The plastomes of *Hydnora visseri* and *Prosopanche americana* are among the most reduced genomes with only 24 and 26 unique genes, respectively (Naumann et al. 2016; Jost et al. 2020). Their gene order is mostly conserved, along with their genome size and G + C content (Jost et al. 2020). However, *P. americana* lacks a typical IR but instead contains a direct repeat (DR) with a gene set similar to the IR in *H. visseri*, indicating putative homology (Jost et al. 2020). In both cases, the repeat gene contents represent a highly condensed version of what is commonly found in IRs of autotrophic flowering plants (Zhu et al. 2016).

Hydnoraceae is a family of holoparasitic root parasites consisting of the two genera *Hydnora* Thunb. (Thunberg 1775) and *Prosopanche* de Bary (de Bary 1868) (fig. 1). *Hydnora* and its eight known species (Bolin et al. 2018) are found exclusively in the Old World, in Central and South Africa and Madagascar (Musselman and Visser 1989). The seven species of genus *Prosopanche* are found in the New World, in Central and South America (Gómez-Pignataro and Gómez-Laurito 1981; Musselman and Visser 1989; De Carvalho et al. 2021). *Hydnora* and *Prosopanche* have an estimated crown group age of ~55 Ma (Naumann et al. 2013), but could potentially be as old as the separation of the African and South American continent. To date, little is known about the infrageneric relationships of *Hydnora* and *Prosopanche*. Tree reconstruction on a set of eight Hydnoraceae species using two phylogenetic markers (ITS, *rpo*B) recovered distinct clades (Bolin 2009), yet one of the two markers (*rpo*B) was later shown to be missing from the complete plastomes (Naumann et al. 2016; Jost et al. 2020).

Here, we compare complete plastomes of seven species of *Hydnora*, six newly sequenced, and three species of *Prosopanche*, two newly sequenced. For comparative reasons we additionally assembled plastomes of close photosynthetic relatives (*Thottea sumatrana*, Aristolochiaceae and *Lactoris fernandeziana*, Lactoridaceae). Using comparative approaches, we studied the extent of plastome condensation and structural variation. In combination with a resolved species-level phylogenetic hypothesis of Hydnoraceae, we explore putative plastome evolutionary trajectories, in particular the inverted and DR evolution.

Fig. 1.—The only aboveground plant structures of Hydnorraceae are the flowers. Flowers of (*A*) *Prosopanche americana*, (*B*) *Prosopanche bonacinae*, (*C*) *Hydnora visseri*, and (*D*) *Hydnora triceps*. The latter species is one of the only two flowering plants from which it is known that they exclusively flower below ground.

Results

Size and Gene Content of Hydnoraceae Plastomes and Their Relatives

The plastomes of all ten studied species of Hydnoraceae show a high degree of similarity with respect to gene content as well as genome size (table 1, supplementary fig. 1, Supplementary Material online). The plastomes of the *Prosopanche* species are on average larger than the ones of the *Hydnora* species. Within *Prosopanche*, *P. panguanensis* has the largest (28,262 bp) and *P. americana* the smallest (28,191 bp) plastome. Plastomes of *Hydnora* species vary to a higher degree with respect to genome size, with *H. esculenta* having the smallest (24,479 bp) and *H. triceps* having the largest (28,658 bp) (table 1).

The two Hydnoraceae genera share a set of genes, consisting of 15 genes coding for ribosomal subunits (*rps*, *rpl*), one gene coding for the beta subunit of the acetyl CoA carboxylase (*acc*D), two genes involved in protein translocation through plastid membranes (*ycf*1, *ycf*2), four ribosomal RNAs (*rrn*4.5, *rrn*5, *rrn*16, *rrn*23), as well as three tRNAs (*trn*E-UUC, *trn*I-CAU, *trnf*M-CAU). *Prosopanche* species contain two additional tRNAs, that is, *trn*W-CCA and *trn*Y-GUA. The Hydnoraceae plastomes have lost the genes generally containing group IIa introns with the exception of *rpl*2, *rpl*16, and *rps*12. However, the only remaining intron in the Hydnoraceae plastomes is intron 1 of the transspliced *rps*12.

All Hydnoraceae protein-coding genes show open reading frames (ORFs), except for *rps*19 in *P. americana* and *rps*7 in *H. visseri*. For the latter, however, a site was identified where potential posttranscriptional editing could create an in-frame start codon. Generally, the *Hydnora rps*7 5′-region is highly divergent between accessions regarding length and nucleotide composition. In contrast, the 3′ end is highly conserved, including *H. visseri*. In total, the *rps*7 ORF in *Hydnora* varies greatly in length (546–1,065 bp from *H. triceps* to *H. longicollis*, respectively) with the majority of length differences stemming from the 5′ region.

Table 1

Plastome Size and Nucleotide Composition of Hydnoraceae in Comparison to Outgroup Lineages

Accession	Total Length (bp)	LSC (bp)	SSC (bp)	IR (bp)	Total GC (%)	GC LSC (%)	GC SSC (%)	GC IR (%)
Aristolochia fimbriata	160,529	90,080	19,583	25,433	38.5	36.9	33.3	43.4
Thottea sumatrana	158,610	88,665	19,053	25,446	38.3	36.3	33.4	43.5
Lactoris fernandeziana	159,739	81,254	18,529	29,978	38.5	37.1	32.6	42.2
Prosopanche bonacinae	28,246	—	—	—	23.9	—	—	—
Prosopanche americana	28,191	—	—	—	20.4	—	—	—
Prosopanche panguanensis	28,262	—	—	—	20.8	—	—	—
Hydnora esculenta	24,479	—	—	—	24.0	—	—	—
Hydnora arabica	24,790	15,734	6,740	1,158	23.1	24.4	20.1	23.0
Hydnora abyssinica	24,672	15,793	6,947	966	24.1	25.5	21.2	23.0
Hydnora africana	26,957	22,233	1,386	1,669	22.3	23.0	22.1	17.4
Hydnora triceps	28,658	22,126	1,420	2,556	21.8	23.1	21.3	16.2
Hydnora longicollis	26,714	22,329	1,539	1,423	23.4	23.9	23.3	19.7
Hydnora visseri	27,233	22,751	1,550	1,466	23.4	24.0	23.0	19.0

NOTE.—Size and %GC differences are estimated based on complete plastomes as well as on plastome regions (LSC, SSC, and IR).

The longest ORFs are more than twice the size of autotrophic relatives (1,065 bp in *H. longicollis* compared with 468 bp in *Aristolochia fimbriata*, *L. fernandeziana*, and *T. sumatrana*). The *rps*19 gene shows a high degree of sequence similarity in *Prosopanche* although no ORF could be identified in *P. americana* owing to a missing in-frame start codon. In *Prosopanche bonacinae*, the ORF is expanded, partially overlapping with the *ycf*1 gene.

The secondary structure predictions for the Hydnoraceae tRNA genes result in typical cloverleaf folds for all tRNAs (supplementary fig. 3, Supplementary Material online). Additionally, the anticodon as predicted by sequence similarity is seen in all cases, except for *trn*W-CCA in *P. americana* which showed ACA instead.

The nucleotide composition of the Hydnoraceae plastomes is highly biased toward A and T nucleotides, compared with outgroup taxa (table 1). The average GC content (%) of the latter is 38.4, compared with an average of 23.1 in *Hydnora* (21.8 in *H. triceps*—24.1 in *Hydnora abyssinica*) and 21.7 in *Prosopanche* (20.4 in *P. americana*—23.9 in *P. bonacinae*).

For comparative reasons, we assembled plastomes of two close photosynthetic relatives of Hydnoraceae. *Thottea sumatrana* (Aristolochiaceae) and *L. fernandeziana* (Lactoridaceae) plastomes show the typical quadripartite structure consisting of two single-copy regions (LSC and SSC) separated by two copies of an IR. For details, refer to table 1 and supplementary fig. 2, Supplementary Material online.

Relationships Within Hydnoraceae

Constrained and unconstrained phylogenetic tree reconstructions using maximum likelihood recovered identical relationships within holoparasitic Hydnoraceae, including full support (100 bootstrap support, BS) for monophyletic *Prosopanche* and *Hydnora* (fig. 2 and supplementary figs. 4 and 5, Supplementary Material online). The constrained reconstruction is based on generic relationships published in Jost et al. (2021). Within *Prosopanche*, *P. bonacinae* is recovered as sister to Prosopanche clade I (*P. americana*, *P. panguanensis*) with full BS support. The sister relationship within the latter clade scored BS 93–100. Within *Hydnora*, *H. esculenta* is recovered as sister to all other species (BS 100). Hydnora clade I (*H. africana*, *H. triceps*, *H. longicollis*, *H. visseri*) is recovered sister to Hydnora clade II (*H. arabica*, *H. abyssinica*). Within Hydnora clade I (BS 99–100), *H. longicollis* plus *H. visseri* are sister to *H. africana* plus *H. triceps* (BS 100, fig. 2, supplementary figs. 4 and 5, Supplementary Material online). Extremely elevated branch lengths compared with the outgroup sampling are found leading to Hydnoraceae as well as to *Hydnora* and *Prosopanche* specifically (supplementary fig. 4, Supplementary Material online). The longest branch lengths within Hydnoraceae are found leading to Hydnora clade II, whereas they are generally much longer between all parasitic members compared with their autotrophic relatives (supplementary figs. 4 and 5, Supplementary Material online).

Structural Hydnoraceae Plastome Comparison

Structurally, all species of *Hydnora* contain an IR with varying gene content (fig. 3), except for *H. esculenta*, which has lost its IR (fig. 3, supplementary fig. 1, Supplementary Material online). The three sampled *Prosopanche* species do not contain an IR, instead two copies of a DR were identified (fig. 3, supplementary fig. 1, Supplementary Material online). The latter repeats contain a complete copy of *trn*I along with a partial copy of *rpl*2 and, in *P. bonacinae* additionally a partial *ycf*2 copy (figs. 3 and 4). The latter repeat is identical to the IR gene content found in Hydnora clade I (*ycf*2-*trn*I-*rpl*2). The IR of *H. africana*, unlike the other Hydnora clade I repeats, additionally contains the 3′-portion of *trn*E (figs. 3 and 4). In general, the Hydnoraceae repeat content is a highly condensed version

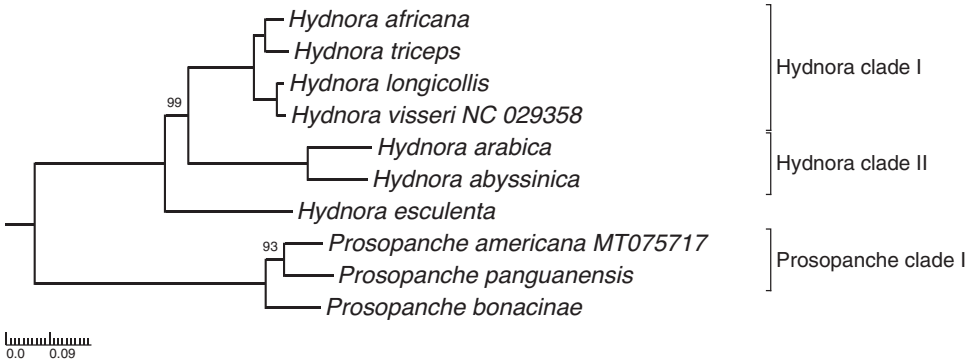

Fig. 2.—Tree reconstruction recovers distinct clades within Hydnoraceae. Phylogenetic tree reconstruction using RAxML shown as phylogram. Only Hydnoraceae taxa are shown. For full phylogenetic relationships, refer to supplementary figs. 4 and 5, Supplementary material online. The tree reconstruction is based on an 83 gene matrix, applying a gene partition approach and without restriction of relationships. Support values are based on 1000 bootstrap replicates and are shown for nodes with <BS 100.

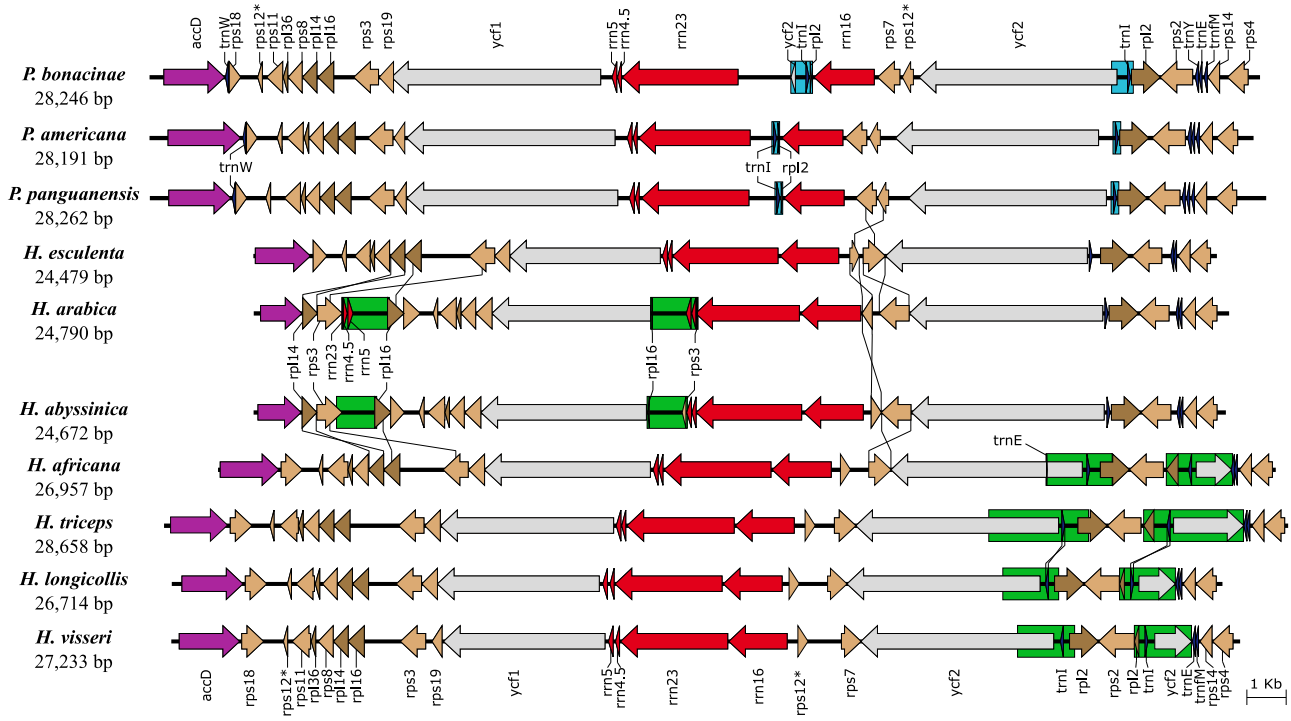

Fig. 3.—Linear plastome comparison highlights structural similarities and dissimilarities within Hydnoraceae (*Hydnora* [H.] and *Prosopanche* [P.]). Linear comparison of plastome diversity is shown with genes depicted as arrows, with their direction indicating their orientation in relation to one another. The colors correspond to different gene groups: *acc*D (purple), tRNA (blue), *rps* (light brown), *rpl* (dark brown), *ycf* (gray), and rRNA (red). IRs are highlighted by green boxes, DRs by light blue boxes. Synteny between neighboring plastomes shows rearrangements.

of what is found in Aristolochiaceae and Lactoridaceae (fig. 4, supplementary fig. 2, Supplementary Material online). PCR and Sanger sequencing have confirmed the IRs and their positioning for *H. abyssinica* and *H. arabica*.

Multiple events of gene inversion and translocation are observed in Hydnoraceae. The *rps*7–*rps*12 region in *Prosopanche* is inverted compared with the majority of *Hydnora* species (*H. esculenta* and Hydnora clade I) as well as the outgroup (figs. 3 and 4). In Hydnora clade II, *rps*7 is found in opposite orientation compared with the remainder of *Hydnora* accessions. Lastly, the *rps*12 exon 2 is inverted in *H. abyssinica* compared with *H. arabica*, but identical with the other Hydnoraceae species (fig. 3). For the sister group *H. longicollis* and *H. visseri*, *trn*I is inverted

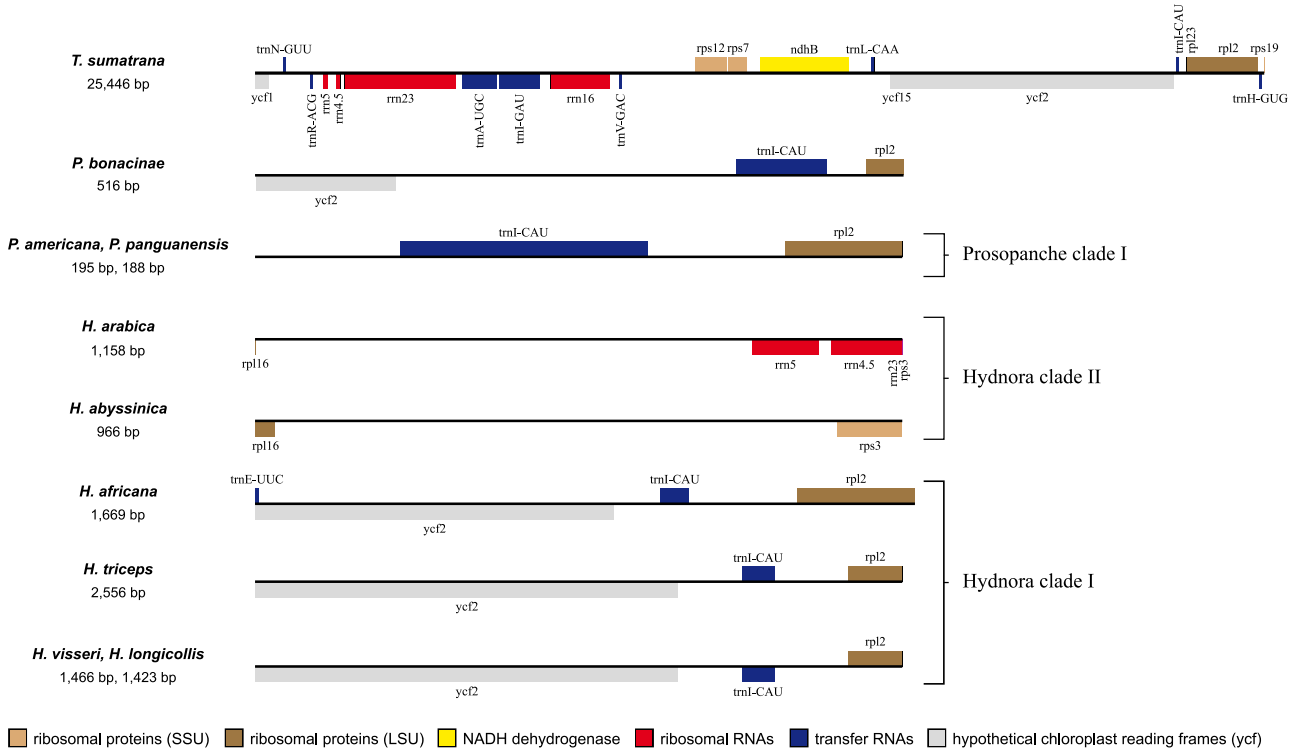

Fig. 4.—Hydnoraceae repeats show varying gene content and size. Gene content of the *Hydnora* IR and *Prosopanche* DR copies in comparison to the IR of autotrophic *Thottea sumatrana* as representative for the autotrophic sister lineages. Taxa with identical repeat gene content are grouped. Repeats are not drawn to size; repeat lengths are displayed as number of base pairs.

compared with the other sampled members of Hydnoraceae (fig. 3) as well as to outgroup taxa *L. fernandeziana* and *T. sumatrana*. Additional structural differences to the remainder of the sampling can be found in Hydnora clade II (fig. 3). Here, the *rpl*14 gene is translocated and inverted, placing it in the *acc*D and *rps*3 spacer. *Rpl*16–*rps*3 has been subjected to a similar event, whereas also containing the second IR copy in Hydnora clade II. Although neighboring one another, both events are to be considered independent from each other, based on gene order.

Simple and Tandem Repeat Analyses

Mononucleotide repeats were the most common type in all studied taxa, accounting for at least half of the total SSRs (supplementary table 1, Supplementary Material online). In Hydnoraceae, the more complex repeat types account for much less of the total number of SSRs, with penta- and hexanucleotide repeats either missing or having the smallest share. In *Hydnora arabica*, no tetra-, penta, or hexanucleotide repeats have been found (supplementary table 1, Supplementary Material online).

Two regions containing tandem repeats in the Hydnoraceae plastomes are noteworthy. First, the shared tandem repeats in the *rps*3–*rpl*16 spacer. Some of which seem to be homologous based on sequence alignments, but often vary in length and copy number, others being unique for specific taxa. With Hydnoraceae having lost the *rpl*16 intron, the length of the spacer in *Hydnora* often equals the length between *rps*3 and *rpl*16 exon 2 in Aristolochiaceae and Lactoridaceae. In *Prosopanche*, the spacer is much shorter, yet tandem repeats were also identified here. Second, tandem repeats in the *Hydnora rps*7 5′-region have been identified. Based on sequence similarity, the repeat sequences seem to be homologous, yet their exact length and copy number varies between accessions, resulting in different *rps*7 ORF lengths. In *Prosopanche*, *rps*7 contains two distinct tandem repeats, both in the center of the gene as well as near the 3′ end. The repeat copy numbers here are more conserved, resulting in less variation in gene length. Generally, the number of genes affected by tandem repeats is slightly higher in Hydnoraceae compared with Aristolochiaceae and Lactoridaceae.

In nearly all accessions, tandem repeats make up a significantly larger portion of the plastomes compared with SSRs (with the exception of *L. fernandeziana* and *H. esculenta*, supplementary table 1, Supplementary Material online).

Discussion

In Hydnoraceae, a high degree of plastome similarities within and among the two genera are reported for many

categories such as genome size, gene order, as well as gene and GC content, despite the split of *Prosopanche* and *Hydnora* ~55 Ma (crown age 36–74 Ma, Naumann et al. 2013). However, major differences are observed with respect to inverted and DR structure and content among Hydnoraceae species.

Possible Evolutionary Trajectories Resulting in the Current Plastome States

A large fraction of genes retained in the Hydnoraceae plastomes is commonly found in the IRs of autotrophic relatives (supplementary figs. 1 and 2, Supplementary Material online, Naumann et al. 2016; Jost et al. 2020). Although this fraction is generally high among highly condensed plastomes (Bellot and Renner 2016; Lim et al. 2016; Roquet et al. 2016; Arias-Agudelo et al. 2019; Su et al. 2019; Yudina et al. 2021), comparatively those genes appear to occupy the most space in the Hydnoraceae plastomes. In contrast to the studied species of *Hydnora* (excluding *H. esculenta*), plastomes of other highly condensed holoparasitic plants generally lack their quadripartite structure, due to IR loss (Bellot and Renner 2016; Roquet et al. 2016; Arias-Agudelo et al. 2019; Su et al. 2019). The plastomes of mycoheterotrophic *Thismia* on the other hand contain two IR copies, occupying a large portion of the total plastomes sequence (Lim et al. 2016; Yudina et al. 2021). Currently, there are no highly condensed plastomes for which *Prosopanche*-like DR structures have been reported. Mapping evolutionary events to the Hydnoraceae phylogeny leads to three possible scenarios (models) for plastome evolution (fig. 5), based on the presence or absence of distinct repeats, their gene content, as well as their differing positioning (fig. 6). Comparison of gene order and orientation between Hydnoraceae and the outgroup taxa (*Thottea* plastome shown as representative) reveals that the same repeat copy (hereafter named IRb/DRb) is the only one affected by either translocation or loss, whereas the second copy (hereafter named IRa/DRa) positioning is homologous to autotrophic relatives (fig. 6). Assuming homology between the plastomes of the outgroup and Hydnoraceae, one would expect IRb/DRb to be located between *rps*19 and *ycf*1 (considering gene losses in comparison to the outgroup, fig. 6). However, the repeats are found at distinct positions for each clade, resulting from independent events of translocation (fig. 6).

The contraction model (figs. 5 and 6) assumes the Hydnoraceae repeats (IR or DR) to be condensed copies of the IR present in autotrophic relatives. A total of three independent events of repeat contraction and a single repeat loss have to be assumed in Hydnoraceae (fig. 5), considering the different gene content of the Hydnora clade I and II IRs (figs. 4 and 6). Additionally, this model requires three independent events of repeat copy translocation (IRb/DRb) among Hydnoraceae taxa. The *Prosopanche* DRs therefore result from the inverse insertion of the translocated IRb copy (fig. 6). Small fluctuations of the inverted repeat borders are not uncommon (Zhu et al. 2016), however, the degree of contraction as assumed under this model has not been reported so far to our knowledge.

The second model, which requires fewer steps compared with the previous model and is, therefore, more parsimonious, is called the ancestral loss model. This model has to assume an IR loss (IRb) before Hydnoraceae diversification (fig. 5), followed by three independent, secondary repeat regain events (DRa' in *Prosopanche* and two IRa' in *Hydnora* depending on the clade, fig. 6). These events have to be considered independent based on the significantly different plastome structures and repeat gene contents. One regain is postulated in *Prosopanche* (DR) and one in each *Hydnora* clade (IR2 and IR3) (fig. 5). In this model, the repeat gene content and the repeat position, for each independent regain event, are simply by chance. In case of duplication events, chances are high that parts of the originally IR-located sequences are duplicated, given that ~40–55% of the Hydnoraceae plastomes comprise content that is commonly found in IRs of autotrophic relatives (fig. 6). In general, secondary IR gain after an ancestral loss is extremely rare and has only been reported in two lineages so far. Twice in legumes (Choi et al., 2019; Wu et al. 2021) and once in heterotrophic *Parasitaxus* (Qu et al. 2019). Mechanisms underlying the regain have not been studied extensively, although the involvement of double-strand break repair mechanisms has been discussed. The repair of double-strand breaks via formation and solution of Holliday junctions and synthesis-dependent strand annealing are proposed to trigger potential IR regain (Choi et al. 2019).

The following repeated loss model is as parsimonious as the latter model (fig. 5). In contrast to the aforementioned ancestral loss model, the repeat content in Hydnora clade I (IR) and *Prosopanche* is assumed homologous as the independent gain of repeat structures with (nearly) identical gene content is unlikely. In both lineages, the repeat consists of *trn*I, and partial copies of *ycf*2 and *rpl*2 (fig. 4). However, Prosopanche clade I lost the partial *ycf*2 copy and the IR in *H. africana* contains an additional partial copy of *trn*E (fig. 4). An additional translocation step is required to explain the inverted second repeat copy and positioning in *Prosopanche* (fig. 6 and Jost et al. 2020). The model further requires the IR loss on the branch leading to *H. esculenta* and a loss leading to Hydnora clade II. The latter is followed by a secondary IR gain in Hydnora clade II as the IRs of Hydnora clades I and II are clearly distinct from one another (fig. 5). The IR loss in Hydnora clade II could have been the trigger for the observed structural alternations (*rps*3–*rpl*16, *rpl*14, *rps*7, fig. 5).

Under parsimony criteria, the ancestral loss and repeated loss model are to be favored as they imply fewer steps than

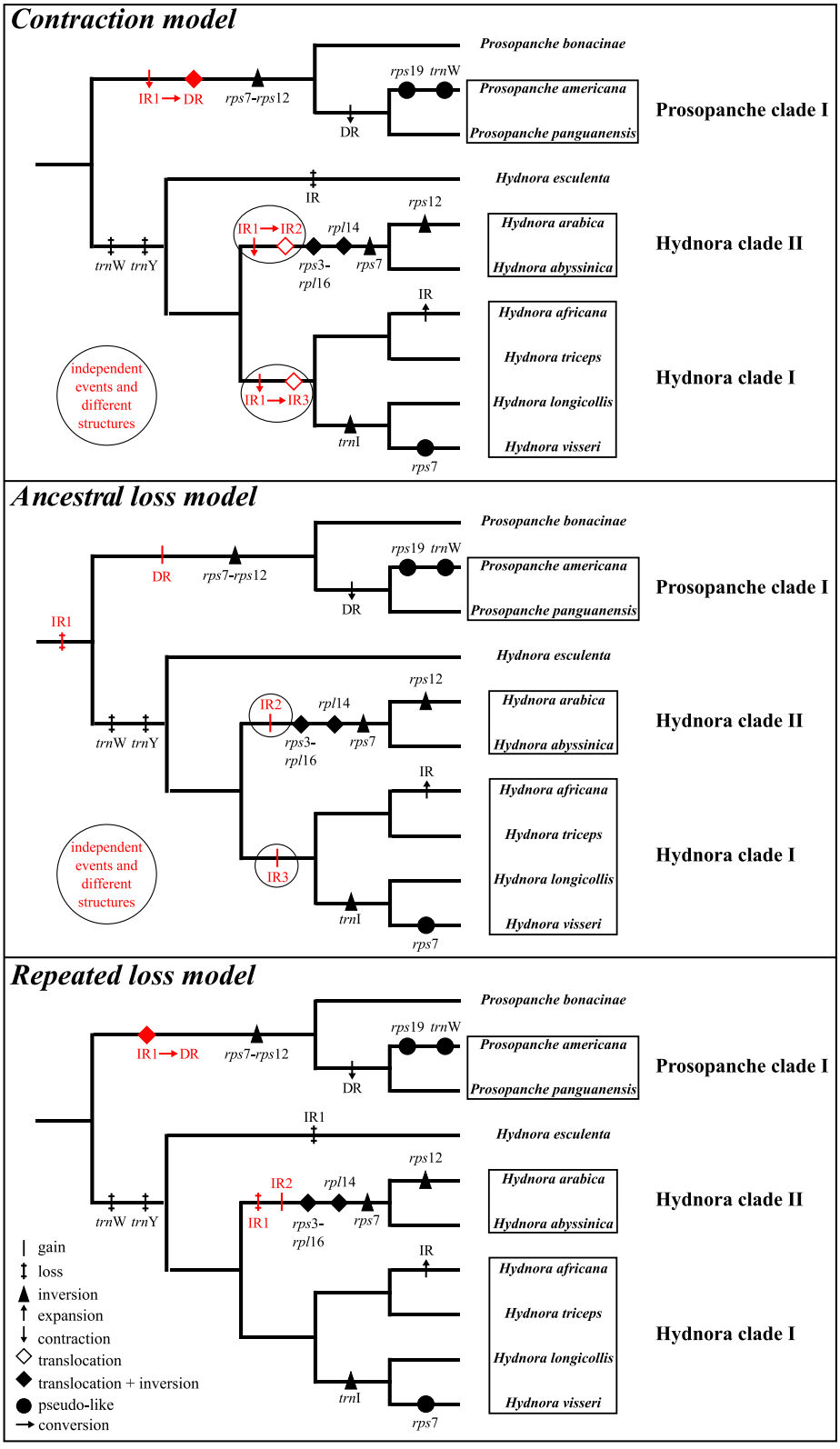

Fig. 5.—Three models to explain repeat (IR, DR) presence in Hydnoraceae. Phylogenetic tree reconstruction of the Hydnoraceae with mapped evolutionary events, separated by model. IR1 corresponds to the putative ancestral IR of the Hydnoraceae MRCA, IR2 to the IR found in Hydnora clade II, and IR3 to the IR of Hydnora clade I. DR corresponds to the DR found in *Prosopanche*. Model-dependent events on the branches are highlighted in red.

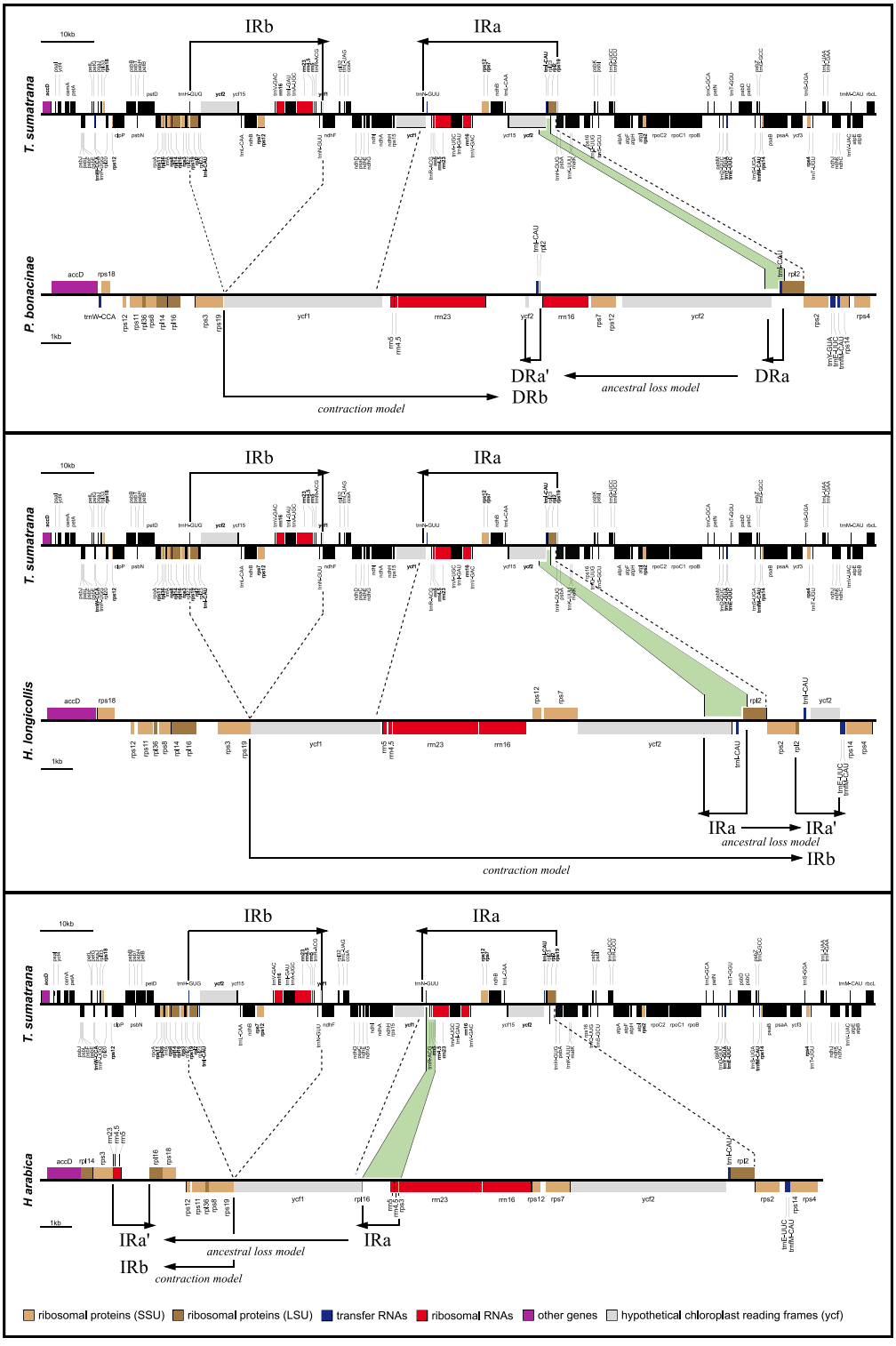

Fig. 6.—Linear comparison of Hydnoraceae repeat content and positions compared with autotrophic *Thottea*. Linear comparison of the distinct repeat sequences of *P. bonacinae*, *H. longicollis* (exemplary for Hydnora clade I) and *H. arabica* (exemplary for Hydnora clade II) compared with *Thottea* as an example for the autotrophic relatives. Repeat copies are named based on the *Thottea* IRs and their putative origin, for example, *Hydnora* IRb is assumed homologous to the *Thottea* IRb, whereas *Hydnora* IRa′ is a repeat directly derived from *Hydnora* IRa. Dotted lines represent Hydnoraceae plastome sequence corresponding to the *Thottea* IRs, as well as their expected positioning based on gene order. The green area highlights the actually repeated sequence in the Hydnoraceae plastomes and their corresponding region on the *Thottea* plastomes. Arrowheads indicate repeat copy orientations. Black boxes on the *Thottea* plastomes indicate genes missing from the Hydnoraceae plastomes.

the contraction model. However, the contraction model might be intuitively favored, given that the Hydnoraceae repeat content could represent highly condensed versions of the IRs of autotrophic relatives. Although the second repeat copies are never in the location expected for an IRb contraction (fig. 6).

A model-independent observation is related to the Hydnora clade II IR copy (IRa′/IRb, fig. 6) contained within translocated *rps*3–*rpl*16 region. This has possible implications for the IR creation within this clade. If the second repeat copy was integrated into this cluster before its translocation, it was so as a DR. Translocation and subsequent inversion could therefore have created the Hydnora clade II IR based on an ancestral DR. In general, we did not find any other structural rearrangements near the translocated repeat copies (fig. 3).

Features Putatively Affecting Plastome Stability

Hydnora and *Prosopanche* share an identical set of genes, which is slightly expanded in *Prosopanche* (additional *trn*W and *trn*Y copies). Considering the evolutionary time since the switch to full heterotrophy (assumed before lineage speciation), nonessential genes are assumed lost from the plastome. Therefore, the retained set of genes represents the assumed lineage-dependent essential gene set of Hydnoraceae. Lineage-dependent essential gene sets encode for gene products that cannot be otherwise compensated via nuclear import (Schelkunov et al. 2015; Wicke et al. 2016). Retention of tRNAs W and Y in *Prosopanche* could represent lineage-specific special needs. However, the functionality of *trn*W in *P. americana* cannot be assumed, as the secondary structure prediction shows an ACA instead of the expected CCA anticodon (supplementary fig. 3, Supplementary Material online). Whether the putatively nonfunctional status of *trn*W can be translated to the other *Prosopanche* species is uncertain. The loss of tRNAs Y and W in *Hydnora*, as opposed to a gain in *Prosopanche*, is supported by the general consensus that gene loss during heterotrophic plastome condensation is presumed to be irreversible and therefore one-directional (Barrett and Davis 2012; Graham et al. 2017; Yudina et al. 2021).

Apart from the varying large repeats (IRs, DRs), Hydnoraceae reveal a strikingly low degree of structural rearrangements (fig. 3). We did not find a correlation between IR loss and an increase of the latter (*Prosopanche*, *H. esculenta*, fig. 3), as was shown for certain other lineages (Palmer and Thompson 1982; Palmer et al. 1987). Whether the DR in *Prosopanche* exhibits the same criteria as known for typical IRs is uncertain. In IR-lacking, autotrophic lycophytes, however, IR-like copy-dependent DNA repair is acting upon the retained DR copies, along with lower substitution rates of genes contained within (Mower et al. 2019). IR/DR presence in combination with a lack of large dispersed repeats could help explain the structural conservation as seen in Hydnoraceae. Large dispersed repeats have been shown to coincide with more frequent rearrangements (Mower and Vickrey 2018). Structural plastome stasis, a hypothesized phase of decelerated evolution that followed the phase of accelerated evolution during the major phase of plastome condensation and adaptation to parasitic lifestyle may be considered as additional driver in maintaining the genomic stability (Naumann et al. 2016; Wicke et al. 2016).

Hydnoraceae Plastomes have Lost All But One Intron

Three genes generally known to contain intron sequences are retained on the highly reduced plastomes of Hydnoraceae (*rpl*2, *rpl*16, and *rps*12, Jansen et al. 2007). The loss of the *rpl*2 intron in a single *Hydnora* and *Prosopanche* species was previously reported (Naumann et al. 2016; Jost et al. 2020). Intron losses are not exclusive to plants following a heterotrophic lifestyle, but are generally rather rare (Downie et al. 1991, 1996; McPherson et al. 2004; McNeal et al. 2009; Su et al. 2019). With the additional Hydnoraceae accessions sequenced, we are now confident to report that the remaining two genes have also lost their group IIa introns. As a result, only the trans-spliced *rps*12 intron remains. The intron losses did not affect the presence of ORFs for these genes, indicating putative functionality. The loss of all group IIa introns is in line with the loss of maturase K, which is essential for their splicing (Liere and Link 1995; McNeal et al. 2009).

Materials and Methods

Plant Material, DNA Extraction, and Sequencing

Genomic DNA of Hydnoraceae as well as *T. sumatrana* (Aristolochiaceae) were isolated and sequenced via whole genome shotgun sequencing (supplementary table 2, Supplementary Material online). Hydnoraceae taxa included: *H. abyssinica*, *H. africana*, *H. arabica*, *H. esculenta*, *H. longicollis*, *H. triceps*, *P. bonacinae*, and *P. panguanensis*. All DNAs were extracted from silica dried plant material using the protocol of Doyle and Doyle (1987), modified to include RNAse A (Thermo Scientific, Waltham, MA, USA) treatment (10 mg/ml). DNA concentration and quality were measured using a Qubit 3 Fluorometer (Thermofisher Scientific, Waltham, MA, USA) and the Agilent Technologies 12-capillary Fragment Analyzer (Agilent Technologies, Santa Clara, CA, USA) using the Genomic DNA 50 kb kit. Hydnoraceae sequencing was carried out on rapid mode HiSeq 2,500 flow cells (Illumina, San Diego, CA, USA) as 150 bp paired-end reads for 2 × 150 cycles targeting about 50 million reads per sample.

Sequencing of *Thottea* aimed for 5 million 300 bp paired-end reads instead.

Data Mining from Public Repositories

To expand the sampling within Piperales, as well as to sample outgroup lineages from magnoliids and ANA grade taxa for phylogenetic tree reconstruction, the NCBI nucleotide data base was mined for complete plastomes. In addition, sequence raw data for *L. fernandeziana* (Lactoridaceae) from the Kew Tree of Life data release 0.1 were downloaded from the Sequence Read Archive (ERX4143470). The complete sampling list for this study can be found in supplementary table 2, Supplementary Material online.

Raw Data Assembly and Plastome Reconstruction

Raw read data were assembled using the de novo assembly function in CLC Genomics Workbench (v. 11.0, Qiagen, MD, USA), allowing for automatic calculation of optimal word and bubble sizes for each sample. The assembly was blasted (BLASTn, default settings with an *E*-value cut-off at 1e−10) for contigs of potential plastid origin, using a database containing >40 complete plastomes representing major angiosperm lineages. The latter contigs were imported into Geneious (v. 11.1.5, Biomatters, Ltd., New Zealand) and the correctness of the assembly verified by analyzing the corresponding read-mapping using Tablet (v. 1.21, Milne et al. 2013).

To verify specific plastome regions of interest, and bridge assembly gaps, primers were derived from the flanking regions (supplementary table 3, Supplementary Material online). Both strands of the resulting PCR products were then Sanger sequenced. Manual scaffolding to create complete, circular plastomes was done in Geneious (v. 11.1.5, Biomatters, Ltd., New Zealand). Gene annotation was done using Geneious (v. 11.1.5, Biomatters, Ltd., New Zealand) based on published, closely related plastomes of *H. visseri* (NC_029358, Naumann et al. 2016), *P. americana* (MT075717, Jost et al. 2020), and *Piper nigrum* (NC_034692, Zhang et al. unpublished). Annotations were manually inspected and adjusted where necessary based on ORFs. Circular and linear plastomes were visualized using OGDraw (Greiner et al., 2019), the latter additionally using EasyFig (v. 2.2.5, Sullivan et al. 2011). Secondary structure of tRNAs was predicted using tRNAscan-SE (v. 2.0, Chan et al. 2021) (supplementary fig. 3, Supplementary Material online). For the plastome visualizations, the genome start for all Hydnoraceae was set to the *rps*4-*acc*D spacer.

Phylogenetic Analyses

Single gene alignments for 83 plastid loci (79 protein-coding genes, 4 ribosomal RNAs) were created using the MAFFT algorithm (v. 7.450, Katoh et al. 2002; Katoh and Standley 2013), implemented in Geneious (v. 11.1.5, Biomatters, Ltd., New Zealand) and manually adjusted in AliView (v. 1.20, Larsson 2014). All genes were concatenated using SequenceMatrix (v. 1.8, Vaidya et al. 2011) and inferences were calculated using RAxML (v. 8.2.12, Stamatakis 2014) implemented on Cipres Science Gate (Miller et al. 2010) applying a gene partitioning approach, after estimating the best-fitting nucleotide substitution model using PartitionFinder2 (Lanfear et al. 2014, 2017). In addition to the 83 plastid gene matrix, inferences were estimated based on a plastid matrix reduced to the genes present in the Hydnoraceae plastomes (21 gene matrix). Node support was calculated by 1,000 BS replicates. Inferences for the concatenated 83 loci data set were estimated twice, once without predefined tree topology and once by restricting the relationships of Piperales genera to the ones estimated in Jost et al. (2021). Tree files were then visualized using TreeGraph 2 (Stöver and Müller 2010).

Repeat Analysis

Two types of repeat sequences were studied in the Hydnoraceae plastomes as well as in autotrophic relatives (*A. fimbriata*, *L. fernandeziana*, *T. sumatrana*) (supplementary table 1, Supplementary Material online). Simple sequence repeats were analyzed using the MIcroSAtellite identification tool (MISA, Beier et al. 2017) setting the minimum number of repetitions to ten for mononucleotide repeats, to five for dinucleotide repeats, to four for trinucleotide repeats, and to three for tetra-, penta-, and hexanucleotides. Tandem repeats were identified using Tandem Repeats Finder (Benson 1999) with the recommended settings. Repeat analysis was done on complete plastomes, excluding the second IR copy.

Additionally, we studied dispersed repeats using Geneious (v. 11.1.5, Biomatters, Ltd., New Zealand), setting the minimum repeat length to 40 bp and allowing for up to 10% of mismatches between the repeat copies.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We like to thank the authorities of Peru (Peruvian research permit N° 96-2015-SERFOR/DGGSPFFS granted by the Servicio Forestal y de Fauna Silvestre [SERFOR]), Namibia (Namibian MET Permits No. 1013/2006 and 1350/2009), Oman (permit number 7/2014 granted by the Ministry of Environment and Climate Affairs, Sultanate of Oman), and Madagascar (Parcs Nationaux Madagascar Research Permit: 291/07) for permission to collect material. We

would also like to thank Shireen Harris (South African National Biodiversity Institute and Karoo Desert National Botanical Garden) and Donovan Kirkwood (Stellenbosch University Botanical Garden) for material exchange and assistance in plant material collection. Additionally, we thank Jakob Wegener for help with laboratory work and Claude W. dePamphilis for feedback on an earlier version of the manuscript. This work was supported by the Graduate Academy of TU Dresden.

Data Availability

The newly assembled plastomes were uploaded to NCBI and are available using following identifiers: Hydnoraceae (OM782325–OM782332), *Lactoris fernandeziana* (BK061146), *Thottea sumatrana* (OM686845).

Literature Cited

Arias-Agudelo LM, González F, Isaza JP, Alzate JF, Pabón-Mora N. 2019. Plastome reduction and gene content in New World *Pilostyles* (Apodanthaceae) unveils high similarities to African and Australian congeners. Mol Phylogenet Evol. 135:193–202.

Barrett CF, Davis JI. 2012. The plastid genome of the mycoheterotrophic *Corallorhiza striata* (Orchidaceae) is in the relatively early stages of degradation. Am J Bot. 99(9):1513–1523.

Beier S, et al. 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics 33(16):2583–2585.

Bellot S, Renner SS. 2016. The plastomes of two species in the endoparasite genus *Pilostyles* (Apodanthaceae) each retain just five or six possibly functional genes. Genome Biol Evol. 8(1):189–201.

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27(2):573–580.

Birky CW Jr, Walsh JB. 1992. Biased gene conversion, copy number, and apparent mutation rate differences within chloroplast and bacterial genomes. Genetics 130(3):677–683.

Blazier JC, et al. 2016. Variable presence of the inverted repeat and plastome stability in *Erodium*. Ann Bot. 117(7), 1209–1220.

Bolin J F. 2009. Ecology and molecular phylogenetics of Hydnora (Hydnoraceae) in Southern Africa. [Ph.D. dissertation]. Norfolk (VA): Old Dominion University.

Bolin JF, Lupton D, Musselman LJ. 2018. *Hydnora arabica* (Aristolochiaceae), a new species from the Arabian Peninsula and a key to *Hydnora*. Phytotaxa 338(1):99.

Cai Z, et al. 2006. Complete plastid genome sequences of *Drimys*, *Liriodendron*, and *Piper*: implications for the phylogenetic relationships of magnoliids. BMC Evol Biol. 6(1):77.

Cai L, et al. 2021. Deeply altered genome architecture in the endoparasitic flowering plant *Sapria himalayana* Griff.(Rafflesiaceae). Curr Biol. 31(5):1002–1011.

Chan PP, Lin BY, Mak AJ, Lowe TM. 2021. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res. 49:9077–9096.

Choi I-S, Jansen R, Ruhlman T. 2019. Lost and found: return of the inverted repeat in the legume clade defined by its absence. Genome Biol Evol. 11(4):1321–1333.

Daniell H, Lin C-S, Yu M, Chang W-J. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biol. 17(1):1–29.

de Bary A. 1868. Prosopanche burmeisteri, eine neue Hydnoree aus Süd-Amerika. HW Schmidt.

De Carvalho JDT, et al. 2021. *Prosopanche cocuccii* (Hydnoraceae): a new species from Southern Brazil. Phytotaxa 521(3):177–192.

Downie SR, et al. 1991. Six independent losses of the chloroplast DNA *rpl*2 intron in dicotyledons: molecular and phylogenetic implications. Evolution 45(5):1245–1259.

Downie SR, Llanas E, Katz-Downie DS. 1996. Multiple independent losses of the *rpo*C1 intron in angiosperm chloroplast DNA's. Syst Bot. 21:135–151.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19:11–15.

Biomatters, Ltd., New Zealand. Geneious. https://www.geneious.com/.

Gómez-Pignataro LD, Gómez-Laurito J. 1981. A new species of Prosopanche (Hydnoraceae) from Costa Rica. Una especie nueva de Prosopanche (Hydnoraceae) de Costa Rica. Phytologia 49(1):53–55.

Graham SW, Lam VK, Merckx VS. 2017. Plastomes on the edge: the evolutionary breakdown of mycoheterotroph plastid genomes. New Phytol. 214(1):48–55.

Greiner S, Lehwark P, Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 47(W1): W59–W64.

Jansen RK, et al. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci U S A. 104(49): 19369–19374.

Jansen RK, Wojciechowski MF, Sanniyasi E, Lee S-B, Daniell H. 2008. Complete plastid genome sequence of the chickpea (*Cicer arietinum*) and the phylogenetic distribution of *rps*12 and *clp*P intron losses among legumes (Leguminosae). Mol Phylogenet Evol. 48(3):1204–1217.

Jin D-M, et al. 2020. The loss of the inverted repeat in the putranjivoid clade of Malpighiales. Front Plant Sci. 11:942.

Jost M, Naumann J, Rocamundi N, Cocucci AA, Wanke S. 2020. The first plastid genome of the holoparasitic genus *Prosopanche* (Hydnoraceae). Plants 9(3):306.

Jost M, Samain M-S, Marques I, Graham SW, Wanke S. 2021. Discordant phylogenomic placement of Hydnoraceae and Lactoridaceae within Piperales using data from all three genomes. Front Plant Sci. 12:642598.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30(14):3059–3066.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

Lanfear R, Calcott B, Kainer D, Mayer C, Stamatakis A. 2014. Selecting optimal partitioning schemes for phylogenomic datasets. BMC Evol Biol. 14(1):82.

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol Biol Evol. 34(3). 772–773.

Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30(22):3276–3278.

Lavin M, Doyle JJ, Palmer JD. 1990. Evolutionary significance of the loss of the chloroplast-DNA inverted repeat in the Leguminosae subfamily Papilionoideae. Evolution 44(2), 390–402.

Liere K, Link G. 1995. RNA-binding activity of the *mat*K protein encodecd by the chloroplast *trn*K intron from mustard (*Sinapis alba* L.). Nucleic Acids Res. 23(6):917–921.

Lim GS, Barrett CF, Pang C-C, Davis JI. 2016. Drastic reduction of plastome size in the mycoheterotrophic *Thismia tentaculata* relative to that of its autotrophic relative *Tacca chantrieri*. Am J Bot. 103(6): 1129–1137.

Maréchal A, Brisson N. 2010. Recombination and the maintenance of plant organelle genome stability. New Phytol. 186(2):299–317.
McNeal JR, Kuehl JV, Boore JL, Leebens-Mack J, Depamphilis CW. 2009. Parallel loss of plastid introns and their maturase in the genus *Cuscuta*. PLoS One 4(6):e5982.
McPherson MA, Fay MF, Chase MW, Graham SW. 2004. Parallel loss of a slowly evolving intron from two closely related families in Asparagales. Syst Bot. 29(2):296–307.
Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: 2010 Gateway Computing Environments Workshop (GCE), p. 1–8.
Milne I, et al. 2013. Using tablet for visual exploration of second-generation sequencing data. Brief Bioinform. 14(2):193–202.
Molina J, et al. 2014. Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). Mol Biol Evol. 31(4):793–803.
Mower JP, et al. 2019. Lycophyte plastid genomics: extreme variation in GC, gene and intron content and multiple inversions between a direct and inverted orientation of the rRNA repeat. New Phytol. 222(2):1061–1075.
Mower JP, Vickrey TL. 2018. Structural diversity among plastid genomes of land plants. Adv Bot Res. 85, 263–292.
Musselman LJ, Visser JH. 1989. Taxonomy and natural history of *Hydnora* (Hydnoraceae). Aliso: J Syst Evol Bot. 12(2):317–326.
Naumann J, et al. 2013. Single-copy nuclear genes place haustorial Hydnoraceae within Piperales and reveal a Cretaceous origin of multiple parasitic angiosperm lineages. PLoS One 8(11):e79204.
Naumann J, et al. 2016. Detecting and characterizing the highly divergent plastid genome of the nonphotosynthetic parasitic plant *Hydnora visseri* (Hydnoraceae). Genome Biol Evol. 8(2):345–363.
Palmer JD. 1985. Comparative organization of chloroplast genomes. Annu Rev Genet. 19(1):325–354.
Palmer JD, Osorio B, Aldrich J, Thompson WF. 1987. Chloroplast DNA evolution among legumes: loss of a large inverted repeat occurred prior to other sequence rearrangements. Curr Genet. 11(4): 275–286.
Palmer JD, Thompson W.F. 1982. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell 29(2):537–550.
Petersen G, Zervas A, Pedersen HÆ., Seberg O. 2018. Genome reports: contracted genes and dwarfed plastome in mycoheterotrophic *Sciaphila thaidanica* (Triuridaceae, Pandanales). Genome Biol Evol. 10(3):976–981.
Qu X-J, Fan S-J, Wicke S, Yi T-S. 2019. Plastome reduction in the only parasitic gymnosperm *Parasitaxus* is due to losses of photosynthesis but not housekeeping genes and apparently involves the secondary gain of a large inverted repeat. Genome Biol Evol. 11-(10):2789–2796.
Roquet C, et al. 2016. Understanding the evolution of holoparasitic plants: the complete plastid genome of the holoparasite *Cytinus hypocistis* (Cytinaceae). Ann Bot. 118(5):885–896.
Sanderson MJ, et al. 2015. Exceptional reduction of the plastid genome of saguaro cactus (*Carnegiea gigantea*): loss of the *ndh* gene suite and inverted repeat. Am J Bot. 102(7):1115–1127.
Schelkunov MI, et al. 2015. Exploring the limits for reduction of plastid genomes: a case study of the mycoheterotrophic orchids *Epipogium aphyllum* and *Epipogium roseum*. Genome Biol Evol. 7(4):1179–1191.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312–1313.
Stöver BC, Müller KF. 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinform. 11(1):7.
Su H-J, et al. 2019. Novel genetic code and record-setting AT-richness in the highly reduced plastid genome of the holoparasitic plant *Balanophora*. Proc Natl Acad Sci U S A. 116(3):934–943.
Sullivan MJ, Petty NK, Beatson SA. 2011. Easyfig: a genome comparison visualizer. Bioinformatics 27(7):1009–1010.
Thunberg CP. 1775. Beskrifning paa en ganska besynnerlig och obekant svamp, *Hydnora africana*. Konglia Vetenskaps Akademiens Handlingar. 36:69–75.
Vaidya G, Lohman DJ, Meier R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27(2):171–180.
Weng M-L, Ruhlman TA, Jansen RK. 2017. Expansion of inverted repeat does not decrease substitution rates in *Pelargonium* plastid genomes. New Phytologist. 214(2):842–851.
Wicke S, et al. 2013. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. Plant Cell. 25(10):3711–3725.
Wicke S, et al. 2016. Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants. Proc Natl Acad Sci U S A. 113(32):9045–9050.
Wicke S, Schneeweiss GM, Depamphilis CW, Müller KF, Quandt D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol. 76(3): 273–297.
Wu S, et al. 2021. Extensive genomic rearrangements mediated by repetitive sequences in plastomes of *Medicago* and its relatives. BMC Plant Biol. 21(1):1–16.
Wu C-S, Chaw S-M. 2016. Large-scale comparative analysis reveals the mechanisms driving plastomic compaction, reduction, and inversions in conifers II (Cupressophytes). Genome Biol Evol. 8(12): 3740–3750.
Wu C-S, Lin C-P, Hsu C-Y, Wang R-J, Chaw S-M. 2011. Comparative chloroplast genomes of Pinaceae: insights into the mechanism of diversified genomic organizations. Genome Biol Evol. 3:309–319.
Yudina SV, et al. 2021. Comparative analysis of plastid genomes in the non-photosynthetic genus *Thismia* reveals ongoing gene set reduction. Front Plant Sci. 12:602598.
Zhang N, Ramachandran P, Ottesen AR, Timme RE, Handy SM, Wen J. Plastid genomes of illicium. unpublished.
Zhu A, Guo W, Gupta S, Fan W, Mower JP. 2016. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. New Phytol. 209(4): 1747–1756.

Associate editor: Daniel Sloan