Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften
(Dr. rer. nat.) an der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

NATURAL HISTORY, TAXONOMY, BIOGEOGRAPHY AND GENOME EVOLUTION OF THE WORLDWIDE ENDOPARASITE FAMILY APODANTHACEAE (CUCURBITALES)

Sidonie Bellot

München, 25 Nov. 2014

PREFACE

Statutory declaration

Erklärung

Diese Dissertation wurde im Sinne von §12 der Promotionsordnung von Prof. Dr. Susanne S. Renner betreut. Ich erkläre hiermit, dass die Dissertation nicht einer anderen Prüfungskommission vorgelegt worden ist und dass ich mich nicht anderweitig einer Doktorprüfung ohne Erfolg unterzogen habe.

Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt wurde.

Sidonie Bellot, 25 November 2014
(Unterschrift)

1. Gutachter: Prof. Dr. Susanne S. Renner
2. Gutachter: Prof. Dr. Jochen Heinrichs

Datum der mündlichen Prüfung: 17. April 2015

Note

In this thesis, I present the results from my doctoral research, carried out in Munich from November 2010 to November 2014 under the guidance of Prof. Dr. Susanne S. Renner. My thesis resulted in five manuscripts, presented in Chapters 2 to 6, of which four have been published (Chapters 2, 3, 4 and 6), and one is in review (Chapter 5). I also gave the conference talks listed below. I generated all data and conducted all analyses myself. Writing and discussion involved collaboration with Prof. Dr. Susanne S. Renner.

Sidonie Bellot (Signature)	Prof. Susanne S. Renner (Signature)

List of publications

Peer-reviewed journal articles

BELLOT, S., AND S. S. RENNER. In review. Intracellular and horizontal transfer of plastid genes to the chondriome – endoparasitism as the precondition for plastome loss. *Molecular Biology and Evolution*.

BELLOT, S., AND S. S. RENNER. 2014. Exploring new dating approaches for parasites: the worldwide Apodanthaceae (Cucurbitales) as an example. *Molecular Phylogenetics and Evolution* 80: 1–10.

BELLOT, S., AND S. S. RENNER. 2014. The systematics of the worldwide endoparasite family Apodanthaceae (Cucurbitales), with a key, a map, and color photos of most species. *Phytokeys* 30: 41–57.

BELLOT, S., AND S. S. RENNER. 2013. Pollination and mating systems of Apodanthaceae and the distribution of reproductive traits in parasitic angiosperms. *American Journal of Botany* 100: 1083–1094.

Book chapter

RENNER, S. S., AND S. BELLOT. 2012. Horizontal gene transfer in eukaryotes, focusing on fungi-to-plant and plant-to-plant transfers. Pp. 223-235 in R. Bock & V. Knoop, eds., Genomics of Chloroplasts and Mitochondria. Springer, Heidelberg.

Oral presentations

BELLOT, S. The evolutionary retention of plastid genomes in non-photosynthetic plants: A comparative approach centred on the endoparasitic Apodanthaceae. *Third symposium on the biology of non-weedy parasitic plants*. Namur, Belgium, September 12–15, 2013.

BELLOT, S. Reconstruction of the chloroplast genome of *Spartina maritima*, and molecular dating of the divergence between the two genomes reunited in the recent allopolyploid *S. anglica*. *XVIII International Botanical congress IBC 2011*. Melbourne, Australia, July 23–30, 2011.

BELLOT, S. The complete chloroplast genome of *Spartina maritima* (Pocaeae, Chloridoideae) using 454 sequencing. *BioSystematics Berlin 2011*. Berlin, Germany, February 21–27, 2011.

Poster

RENNER, S. S., AND S. BELLOT. Evolution of the widely distributed endoparasites Apodanthaceae. *Radiations 2014*. Zürich, Switzerland, June 12 – 15, 2014.

Field work

- Australia, near Perth, August 2011
- Zimbabwe, near Harare and Mutare, February 2012
- Iran, near Teheran and Isfahan, June 2012

Funding

• My research was supported by the Deutsche Forschungs Gemeinschaft (DFG RE 603/16-1).

• Travel to attend the 2011 *BioSystematics* conference in Berlin was funded by an EESLMU Travel Grant.

CONTENTS

SUMMARY

This dissertation deals with three key issues in plant evolutionary biology, namely the co-evolution of biological traits, genome evolution, and molecular dating methods. These topics are of particular interest in parasites, which coevolve with their hosts, have highly modified or reduced genomes, often high rates of nucleotide substitutions, and usually no fossil record. To study these topics, I selected the flowering plant family Apodanthaceae (Cucurbitales). Apodanthaceae are endoparasites that live embedded in the stems of their host (Fabaceae or Salicaceae), emerging only in the form of 3-12 mm-long flowers. I revised the taxonomy of the two genera and ten species (Chapter 2) and studied the sexual systems and pollination biology of two species during fieldwork in East Africa and Iran. I also placed key ecological traits of Apodanthaceae in the context of all parasitic angiosperms (focusing on growth form, longevity, pollination, dispersal, and sexual systems) to detect correlations between reproductive traits and parasitism (Chapter 3). I used mitochondrial and nuclear DNA sequences, and molecular clocks, to infer since when Apodanthaceae might have occupied their ranges in North and South America, Africa, Iran, and Australia. This provided an opportunity to explore different relaxed clock models and calibration approaches in the absence of a fossil record (Chapter 4). Lastly, I assessed the consequences of parasitism on the plastid genomes of two species, focusing on the organellar location of autochthon and horizontally transferred plastid genes (Chapters 5 and 6).

My taxonomic work involved the allocation of the 36 names published in the family, using 407 herbarium collections on loan from 11 relevant herbaria, along with geo-referencing of collecting sites, flower dissection, and DNA isolation from representative specimens. I generated a complete phylogeny for the family, a key to all species, an inventory of each species' hosts, and the first specimen-based map of the family's distribution. My fieldwork revealed that *Pilostyles aethiopica* and *P. haussknechtii* are pollinated by flies (Calliphoridae and Ulidiidae) and are dioecious, with one host individual usually bearing only one sex of the parasite. My review of reproductive traits and parasitism in angiosperms showed the predominance of animal pollination among parasites as well as their tendency to be dioecious.

High nucleotide substitution rates (compared to their autotrophic relatives) characterize many parasites. I explored two relaxed clock models with different ways of handling rate heterogeneity, namely the Uncorrelated Lognormal and the Random Local Clock models, both implemented in the dating software BEAST, using Bayes Factors to

assess model fit. These experiments were possible because the small species number of Apodanthaceae permitted reaching stationarity of the Markov chain Monte Carlo runs. Perhaps because of the small size of my data matrix, the statistical tests did not always result in clear model preference. The results imply that Apodanthaceae are old, with a stem age between 65 and 81 million years.

The last two chapters of this dissertation deal with the genomic consequences of parasitism, with a focus again on the African *P. aethiopica* but also the Australian *P. hamiltonii*, for both of which I analyzed high-throughput genome sequencing data. I expected them to have reduced plastid genomes, but both species turned out to have completely lost their plastid genome and to retain only non-functional plastome fragments in their mitochondrial genomes. This contrasts with the "minimal plastid genome" retained in the 19 exo-parasitic angiosperms so far investigated. *Rafflesia lagascae*, the only other endoparasite studied besides my two Apodanthaceae, has also completely lost its plastome (Molina et al., Molecular Biology and Evolution 31: 793-803, 2014). These results point to a qualitative difference in plastome evolution between endo- and exo-parasites. That most plastid fragments found in the two Apodanthaceae are located in their mitochondrial genome matches a previously reported sink-like behaviour of this genomic compartment in Cucurbitales. Using BLAST and maximum-likelihood phylogenetic methods, I was also able to show that a large part of the Apodanthaceae plastid-like fragments are of host origin, adding to the many instances of horizontal gene transfer reported between hosts and parasites, a topic that I reviewed in the last Chapter of this thesis.

Chapter 1

GENERAL INTRODUCTION

Type and phylogenetic occurrence of parasitism in the angiosperms

Parasitic plants obtain nutrients and water from another plant, the so-called host. They either live entirely inside the host (endoparasitism; below) or instead penetrate it from the outside (exoparasitism). By definition, parasites damage their host, but they rarely lead to the host's death. Hemiparasites are still able to photosynthesize, while holoparasites have lost the ability to photosynthesize and must attach to a host in order to survive. As a rule of thumb, green parasites are hemiparasites, whereas white or pale, chlorophyll-less plants are holoparasites.

There are 13 clades of parasitic angiosperms (Barkmann et al., 2007: 11, but 13 suspected; Bellot and Renner, 2013: 13), namely Apodanthaceae (10 species) in Cucurbitales, Balanophoroaceae (43 spp.) in basal Santalales, *Cassytha* (21 spp., Lauraceae) in Laurales, *Cuscuta* (200 spp., Convolvulaceae) in Solanales, Cytinaceae (10 spp.) in Malvales, Hydnoraceae (perhaps 9 spp.) in Piperales, Krameriaceae (18 spp.) in Zygophyllales, Lennoaceae (4 spp.) in Boraginales, Mitrastemonaceae (2 spp.) in Ericales, Orobanchaceae (>2000 spp., excluding the basal *Lindenbergia*) in Lamiales, Rafflesiaceae (34-36 ssp.) in Malpighiales, about 2100 species in higher Santalales (excluding Coulaceae, Erythropalaceae and Strombosiaceae), and Cynomoriaceae (1 or 2 species). The latter may belong to Saxifragales (Nickrent et al., 2005; Maia et al., 2014; Renner, Cusimano, and Bellot, DFG proposal, 2014) or Rosales (Zhang et al., 2009). The focus of the present thesis is the worldwide endoparasite family Apodanthaceae.

As expected from their repeated origins, parasitic angiosperm lineages show a great variation in morphology and ecology. Some lineages, such as *Cuscuta*, Orobanchaceae, and parasitic Santalales, comprise hundreds of species and include facultative or obligate hemiparasites as well as holoparasites with narrow or broad host ranges. The most detrimental parasitic weed in terms of economic damage may be *Striga hermonthica* (Orobanchaceae), thought to be responsible for billions of USD in crop loss in Asia and Africa (Parker, 2009). Host ranges can be narrow as in the 34-36 species Rafflesiaceae, all of which only grow on species of the genus *Tetrastigma* (Vitaceae), or broad as in the 1-2 species of Cynomoriaceae, which parasitize species of Chenopodioideae, Plumbaginaceae, Tamaricaceae, and Frankeniaceae (all Caryophyllales), Cistaceae (Malvales), Fabaceae (Fabales), Asteraceae (Asterales), and Nitrariaceae (Sapindales) (Cui et al., 2013). The Apodanthaceae on which I have worked include just 10 species, yet have a worldwide distribution and parasitize species in two families, Fabaceae and Salicaceae, as presented in

Chapter 2 of this thesis. More information about their morphology and ecology is provided in the last section of this *General Introduction*.

Biological traits of parasitic angiosperms

A review of the biological traits of parasitic angiosperms published some 20 years ago (Molau, 1995) found no shared biological trait setting parasites apart from non-parasitic angiosperms. In the same year, however, an analysis of the reproductive systems (and other traits) of angiosperms found that 4% of the genera containing dioecious species also contain parasitic species, whereas only 1% of the genera without dioecious species contain parasites (Renner and Ricklefs, 1995). This apparent over-representation of dioecy among parasites matches the finding that 17% of the non-photosynthetic species occurring in the flora of North and South Carolina are dioecious, exceeding the percentage of dioecious species among non-parasitic species in the Carolinas (Conn et al., 1980). More recently, a diversification analysis by Hardy and Cook (2012) found a positive correlation between parasitism and species extinction. To place my work on the Apodanthaceae in this kind of context, I reviewed the morphological and mating system data now available on traits of parasitic angiosperms to test if they really share no common traits. I did not re-do the diversification analysis of Hardy and Cook, however, because few new data on the stem ages and species numbers of parasitic plants have become available since their work (except that they failed to include Apodanthaceae, which were first studied phylogenetically and in terms of divergence times in the present thesis).

To test traits that might characterize most or all parasitic angiosperms, I databased the 4010 accepted species names of parasites along with information available in the literature for species' flower morphology (whether unisexual or bisexual), population sexual system(s), longevity (annual, perennial), pollinator(s), and seed and fruit dispersal. I also surveyed sister groups of the parasitic clades for the same biological traits. This resulted in data for 5058 species. I completed this review by fieldwork on the pollination and sexual system of two species, *Pilostyles aethiopica* in Zimbabwe and *P. haussknechtii* in Iran. The results of my review of parasite ecological traits (and documentation of the pollinators of the two species) are presented in Chapter 3 of this thesis.

Substitution rates in the genomes of parasitic angiosperms, and their handling in molecular clock models

Placing parasitic plants in molecular trees has been problematic because of the difficulty to isolate clean parasite DNA (Nickrent et al., 2004), to amplify it with standard primers (Bellot and Renner, 2014a, b), and to properly align the often highly divergent sequences (Barkman et al., 2004). In many parasites, only flowers are available for DNA isolation because leaves may be lacking, stems may also be absent or may comprise much dead tissue, and fruits are rarely collected and preserved in suitable condition for DNA isolation. In endoparasites, the flowers are intimately connected to the host from whose bark they emerge (colour photos on Figs. 5 and 6 in Chapter 2). Contamination by host DNA during the grinding of flowers, in the case of Apodanthaceae just a few millimetres large, is therefore difficult to avoid. A second issue is horizontal gene transfer (HGT) between parasite and host (cf. *Internal and horizontal gene transfer in parasitic angiosperms*). Contamination or HGT have sometimes led to contradictory placements of parasitic plants with different genes and different laboratories (e.g., Nickrent et al., 2004; Zhang et al., 2009; Barkman et al., 2007).

I encountered the "formidable contamination issues" mentioned by the parasite specialist Dan Nickrent (in Nickrent et al., 2004; p. 12) in my own work: even when using only the uppermost parts of Apodanthaceae flowers where no host tissue could be distinguished under the binocular, I sometimes amplified host DNA along with fungal DNA. This problem may be amplified by high sequence divergence causing poor primer binding. It is often necessary to design primers specific to a particular parasite group, which requires at least one clean, long reference sequence of the respective gene or shotgun sequencing data (Bellot and Renner, 2014a). A further difficulty when working with holoparasites (which of course include all endoparasites) is that plastid markers cannot be used because they have been lost or are too divergent to be amplified with standard primers (Nickrent et al., 1997). Studies aiming at placing parasites in the angiosperm phylogeny therefore have relied on mitochondrial and nuclear markers, mostly *mat*R and 18S (Nickrent et al., 2002; Nickrent et al., 2005; Filipowicz and Renner, 2010).

Once the proper parasite sequence has been amplified, high divergence from other angiosperms can make alignment problematic. To facilitate proper alignment some phylogenetic studies of parasites have used codon-based alignments (Filipowicz and Renner, 2014) or have incorporated information from the secondary structure of ribosomal RNAs

(Bellot and Renner, 2014b). A combination of these approaches allowed me to build matrices for Apodanthaceae and up to 12 outgroups from Fagales, Cucurbitales and Rosales, which was necessary to compare nucleotide substitution rates in Apodanthaceae with those in their photosynthetic relatives.

Substitution rates in holoparasitic species have been assessed by comparing non-synonymous and synonymous substitutions in parasites and their sister clades (Bromham et al., 2013). It appears that the plastid rRNA *rrn*16 and the plastid genes *mat*K and *rbc*L have higher substitution rates in holoparasites than in their photosynthetic relatives, except for Mitrastemonaceae (Bromham et al., 2013). The nuclear ribosomal 18S region also has a higher rate in the holoparasites *Rafflesia keithii* and *Rhizanthes infanticida* (Rafflesiaceae), *Bdallophyton americanum* and *Cytinus hypocistis* (Cytinaceae), *Pholisma arenarium* (Lennoaceae), *Cuscuta gronovii* (Convolvulaceae), and *Mitrastemon yamamotoi* (Mitrastemonaceae) compared to their autotrophic sister taxa (Lemaire et al., 2011). For the mitochondrial *matR* gene, this holds for Rafflesiaceae compared to Euphorbiaceae (Davis et al., 2007), and for some other parasite/non-parasite pairs the rate difference also holds for mitochondrial *cox*1, *atp*1, and/or *nad*1 (Bromham et al., 2013). The nuclear *Leafy* gene, however, appears to evolve similarly in Balanophoraceae and their autotrophic relatives (Su and Hu, 2012). Attempts to explain such branch length differences between parasites/non-parasites by herbaceous growth, small population sizes, or short generation times (Bromham et al., 2013) have not been conclusive, which matches my finding that parasites as a group have none of these traits in common (Bellot and Renner, 2013).

Regardless of the eventual explanation of possible genome-wide high substitution rates in parasitic plants, any large rate jumps need to be taken into account in molecular-clock dating studies that include parasitic and non-parasitic lineages. Such dating is helpful for understanding disjunct geographic ranges (as in Apodanthaceae; see geographic map in Fig. 2 in Chapter 2) or for inferring the likely time of evolution of biological traits. To obtain absolute times (not just relative times) from calibrated genetic distances (Sarich and Wilson, 1967) workers today usually rely on Bayesian approaches because they provide well-understood measures of confidence.

The most common way to handle heterogeneous substitution rates in clock models is to use an uncorrelated lognormal clock model as implemented in the dating software BEAST (Drummond et al. 2006, 2012; Bouckeart et al., 2014). The uncorrelated lognormal (UCLN) model assumes different rates of substitution in adjacent branches, and its prior distribution of rate variation assigns a low probability to rates deviating considerably from the mean rate.

This model is not well suited for phylogenies in which rates are identical or near-identical among close relatives, but may "jump" among distant relatives (Rannala and Yang, 2007). To cope with this problem, Drummond and Suchard (2010) implemented the Random Local Clock (RLC) model in BEAST, starting from version 1.5.4. This allows the co-occurrence in a phylogeny of autocorrelation among adjacent branches and rate jumps among more distant branches. The RLC model should be useful for clades, such as parasites and their non-parasitic sister species, where one expects a large rate jump between the branch leading to the parasitic clade but not necessarily within clades (say closely-related parasite species).

The RLC model has not been much used, probably because of its huge parameter space, which makes it difficult to reach convergence. Two studies so far have compared RLC and UCLN models using simulated or empirical data (Dornburg et al., 2012; Rothfels and Schuettpelz, 2014). The former study analysed two empirical matrices of 32 and 51 taxa and, respectively, 13 mitochondrial and 4 nuclear markers, totalling thousands of aligned nucleotides. The two clock models yielded somewhat, but not drastically, different node ages, with the UCLN model being more sensitive to rate shifts and to changes in the calibration. Dornburg et al. (2012) also carried out simulations on a matrix of 1000 nucleotides x 32 taxa and found that the RLC model recovered true ages better than did the UCLN model in cases of localized strong rate heterogeneity. The more recent study, by Rothfels and Schuettpelz (2014), analysed a matrix of 26 taxa and hundreds of aligned nucleotides (6 markers from all genomic compartments) under both models and obtained highly divergent node ages, but could not decide which model fit their data better. Neither of the two studies used statistical tests to assess model fit, and both studies had trouble reaching convergence with the RLC model. For Rothfels and Schuettpelz (2014, p. 38), "some individual runs failed to converge" but they could reach convergence by combining runs, whereas Dornburg et al. (2012, p. 732) "encountered severe problems with the mixing of the Markov chains when using the RLC model" leading to different estimates from individual non-stationary runs.

Since the UCLN and the RLC models are not nested, their fit to data can only be compared by the ratio of their marginal likelihoods (Bayes factor, Kass and Raftery, 1995) rather than from their likelihoods as in the classic likelihood ratio test (Felsenstein, 1981). For estimating marginal likelihoods, path sampling and stepping stone sampling are the most accurate methods (Baele et al., 2012), and both tests have been available in the BEAST software since 2012 (v. 1.7 Drummond et al., 2012). The drawback of Bayes factor comparisons is that with current computers they are too memory craving to be used on datasets with hundreds of taxa. Apodanthaceae comprise only 10 species, so I could apply

Bayes Factors to compare the fit of different UCLN and RLC rate change models to my data. I also explored the influence that different priors on the number of rate changes had on the posterior rate distribution and age estimates obtained with the RLC model. The results of my clock model experiments are presented in Chapter 4 of this dissertation.

The plastid genome of parasitic angiosperms

Plastids, of which the green chloroplasts are one form, are the remnants of cyanobacteria that have been acquired by the plant unicellular eukaryote ancestor *via* endosymbiosis (Palmer, 1991; Keeling, 2013). Since the original endosymbiosis event about 1 billion years ago (Shih and Matzke, 2013), most of the bacterial genes have been lost or transferred to the host cell's nuclear genome (Martin et al., 2002; Yoshida et al., 2014). Land plant plastids today have genomes of only ca. 150 kb in size, with ca. 100 genes and transfer RNAs (Wicke et al., 2011) of which the majority code for proteins involved in photosynthesis or in the transcription, translation, or maturation of these proteins. In addition to photosynthesis, chloroplasts are physically involved in other metabolic pathways, such as lipid or pigment biosynthesis (Neuhaus and Emes, 2010; Jensen and Leister, 2014). This works via the import of nuclear proteins from the cytoplasm (Jarvis and Lopez-Juez, 2013), except for the plastid genes *acc*D and *clp*P; the former is involved in fatty-acids biosynthesis, the latter encodes a protease also involved in protein import (Krause 2012). The function of two plastid genes, *ycf1* and *ycf2*, is still unclear; *ycf*1 may be involved in the transfer of proteins from the plastid intermembrane space to the stroma (Kikuchi et al., 2013). The plastid-encoded *trn*E gene, finally, is involved in the biosynthesis of haem components of mitochondrial cytochromes (Barbrook et al., 2006).

Due to their functions in pathways other than photosynthesis, the just discussed five genes have been suspected to form a minimal plastome, essential even for non-photosynthetic plants (Wolfe et al. 1992; Krause, 2012). Indeed, these genes are present in the plastomes of most mycoheterotrophs and holoparasites (Logacheva et al., 2011, 2014), although *acc*D, *clp*P and *ycf*1 are pseudogenized in some Orobanchaceae (Li et al., 2013; Wicke et al., 2013). In some non-parasitic Poales, *acc*D has been transferred to the nuclear genome (Harris et al., 2013), and *ycf*1, *ycf*2 and *clp*P have been lost from the plastomes of *Passiflora* and other non-parasitic flowering plants (Downie et al., 1994; Wicke et al., 2011). The possible relocation of these genes to the nucleus is unknown. A minimal plastome might thus be composed of only the *trn*E gene (Barbrook et al., 2006).

The loss of photosynthesis-related genes might happen quickly or slowly, and it may or may not be proportional to the time that elapsed since the evolution of parasitism. As of 10 November 2014, the plastomes of 19 exo-holoparasitic and holomycotrophic angiosperms from 13 genera and 4 families have been sequenced. They are two species of *Cuscuta* (Funk et al., 2007; McNeal et al., 2007), the myco-heterotrophic orchids *Rhizantella gardneri* (Delannoy et al., 2011), *Neottia nidus-avis*, (Logacheva et al., 2011), *Corallorhiza striata* (Barrett and Davis, 2012), *C. maculata* ssp. *maculata*, *C. maculata* ssp. *occidentalis* and *C. mertensiana* (Barrett et al., 2014), the myco-heterotroph Petrosaviaceae *Petrosavia stellaris* (Logacheva et al., 2014), and 10 Orobanchaceae (Wolfe et al., 1992; Wicke et al., 2013; Li et al., 2013).

In the most recent of these studies, Barrett et al. (2014) extend the minimal genome hypothesis of Barbrook et al. (2006) to 11 "housekeeping" genes (including *ycf*2), seven transfer RNAs (including *trn*E) and four ribosomal RNAs, on the empirical observation that these genes are conserved in all plastomes of all non-photosynthetic land plants sequenced so far. When I began my research on Apodanthaceae plastomes in late 2010, very few other parasite plastomes had been studied (the publication dates of the just-cited studies mostly postdate 2010). The surge in studies of parasite organelle genomes is due to next-generation sequencing methods, which have made it possible to obtain data without locus-specific primers. These genomic data, together with phylogenetic and clock-dating studies involving parasitic plants, now provide the framework in which to detect possible patterns in plastid or mitochondrial evolution of parasitic plants. (Clock dating is required to have an absolute time frame for gene loss; see the previous section of this *Introduction*). The specific expectations I had at the outset of my work were that parasites would always loose most or all genes involved in photosynthesis and that this might be proportional to the time elapsed since the acquisition of parasitism.

Two recent studies maintain that the reduction of the plastid genome is neutral and may just be a matter of time (Smith and Lee, 2014; Molina et al., 2014). From high-coverage shotgun sequencing of the genomic DNA of the four species of the holoparasitic alga genus *Polytomella*, Smith and Lee(2014) failed to recover any plastome fragment, and transcriptomic screening showed no expression of the nuclear genes involved in the maintenance of plastid DNA, leading Smith and Lee to conclude that *Polytomella* has completely lost its plastome. Illumina sequencing applied to the genomic DNA of the holoparasitic angiosperm *Rafflesia lagascae* revealed only traces of plastome-like sequences (Molina et al., 2014). Molina and colleagues suggest that "given the low sequencing read

coverage for these gene fragments (<2x coverage), it is likely that these remnant plastid sequences are located in the nuclear genome" (Molina et al., 2014, p. 799). However, drawing such interpretations only from coverage information can be misleading especially in non-photosynthetic plants (Isono et al., 1997) for which we do not have a clear expectation of plastome coverage as we do for photosynthetic *Arabidopsis* or *Oryza*.

To infer the genomic location of the data derived from high-throughput sequencing of Apodanthaceae, I analysed the flanking regions of all plastid-like sequences captured in the total genomic data. This circumvents the problem of not being able to rely on coverage for extracting the plastome, and it is the approach I took in my research to analyse the plastomes of the endoparasites *Pilostyles aethiopica* and *P. hamiltonii*. The focus of my research was thus partly on developing bioinformatic pipelines. The results of my work on plastome evolution are shown in Chapter 5 of this thesis.

Internal and horizontal gene transfer in parasitic angiosperms

The discovery by Molina et al. (2014) of plastid sequences in the nuclear and mitochondrial genomes of *Rafflesia lagascae* adds to the many documented occurrences of intra-cellular genetic transfer (IGT) in angiosperms (Stern and Lonsdale, 1982; Richly and Leister, 2004; Kubo and Newton, 2008). In the same species, these authors also document horizontal gene transfers (HGT), namely plastid sequences of the host *Tetrastigma* in the chondriome of *Rafflesia*. Most of the known HGT events involve parasitic angiosperms (Bergthorsson et al., 2003; Won and Renner, 2003; Mower et al., 2010; reviewed by Renner and Bellot, 2012, Chapter 6 of this thesis; Gao et al., 2014), and experimental work shows that whole organelles and even entire nuclei can be transferred between grafted individuals of different species (Stegemann and Bock, 2009; Stegemann et al., 2012; Fuentes et al., 2014). Since parasitic plants have direct cell contact with their hosts via plasmodesmata, horizontal gene transfer is expected.

Among angiosperm species for which HGT has been reported at a genomic scale is *Rafflesia cantleyi* in which Xi et al. (2012, 2013) counted dozens of mitochondrial and nuclear genes acquired from its *Tetrastigma* host. Another angiosperm with genomic-scale HGT is *Amborella trichopoda*, which is not a parasite, but may have been a host for parasites in the past as suggested by traces of Santalales DNA in its mitochondrial genome (Rice et al., 2013). This understory shrub is often wounded by falling tree branches, and its stems, branches, and leaves are commonly covered by mosses; moss hyphae growing into small

wounds may explain the discovery of long stretches of moss-like DNA in the mitochondrial genome of *Amborella trichopoda* (Rice et al., 2013).

The distinction between HGTs and autochthon sequences can be difficult for short sequences or when the donor and the receiver species have similar genomes. This problem is increased in plastid sequences of parasitic plants because of their high rate of evolution compared to their photosynthetic relatives. This may result in molecular convergence with distantly related taxa that could be misinterpreted as an indication of HGT. Using custom scripts, I developed an empirical BLAST-based approach to detect HGT from their Fabaceae hosts in the genome of Apodanthaceae, with empirical thresholds to discard false negatives and positives. This approach and its results are described in Chapter 5 of this dissertation.

Background on Apodanthaceae and main research questions

To explore plastome evolution and horizontal gene transfer in parasites, I selected the family Apodanthaceae, which is one of only four families of endoparasitic angiosperms (the others being Cytinaceae, Mitrastemonaceae, and Rafflesiaceae). Apodanthaceae comprise 10 species of which one occurs in East Africa, one in Iran, three in Australia, one in North America, and 4 in Central and South America (taxon names and their authors are available in Table S1 of Bellot and Renner, 2014a, here included as Chapter 2). They have 3 to12 mm long flowers and parasitize species of Fabaceae and Salicaceae in whose wood they grow as a network of cells (called endophytic system; Amaral and Ceccantini, 2011). Apodanthaceae are not photosynthetic: they do not have green tissues and appear to lack chloroplasts (Dell et al., 1982; Rutherford, 1970, Bellot, personal observation in Zimbabwe and Iran, 2012). They instead obtain all nutrients from their hosts, sometimes damaging them heavily (Gomes and Fernandes, 1994, Bellot personal observation in Zimbabwe and Iran, 2012). The only modern taxonomic work on Apodanthaceae species was by Ida de Vattimo (1950, 1955, 1971, 1973) who recognized 23 species in three genera (one not validly published). Most studies deal with only one or few species (Dell and Burbidge, 1981; Thiele et al., 2008; Gonzáles and Pabón-Mora, 2014a,b), but Blarer et al. (2004) compared the morphology of three species from different genera and continents. Prior to my work there were next to no reports on ecological traits, such as pollination or sexual systems.

I therefore began my study by observing living plants and collecting material in Australia, Iran and Zimbabwe. In the latter two countries, I documented sexual systems and pollinators, among the first such data for the family. To update the taxonomy of

Apodanthaceae was a pre-condition for my work on their phylogeny, biogeography and ecology. I analysed specimen labels and flowers of 407 collections of Apodanthaceae, including numerous collections made since Vattimo (1950, 1955, 1971, 1973). I also sequenced *matR* and 18S from dry specimens and my own collected material to create a complete phylogeny of the family (Chapters 2 and 4). To place my newly acquired knowledge on pollination and sexual systems of Apodanthaceae in a broader context and to assess the existence of correlations between a parasitic way of life and other biological traits, I included my results on Apodanthaceae in a review of sexual systems, pollination, dispersal modes, and growth forms of all parasitic angiosperms and their sister groups (Chapter 3). Secondly, I used my complete phylogeny of the family to test different clock models for molecular dating and to infer the time frame in which Apodanthaceae achieved their disjunct continental distribution (Chapter 4). Finally, to explore the consequences of endoparasitism on organelle genomes, I analyzed Illumina shotgun sequencing data from the genomic DNA of the Australian *P. hamiltonii* and the African *P. aethiopica* (Chapter 5). During those genomic studies, I also investigated the extent of internal and horizontal gene transfer, phenomena expected in parasitic plants (and a topic that I review in Chapter 6)

Chapter 2

THE SYSTEMATICS OF THE WORLDWIDE ENDOPARASITE FAMILY APODANTHACEAE (CUCURBITALES), WITH A KEY, A MAP, AND COLOR PHOTOS OF MOST SPECIES

Bellot, S. and S.S. Renner

Phytokeys 30: 41–57

PhytoKeys 36: 41–57 (2014)
doi: 10.3897/phytokeys.36.7385
www.phytokeys.com

RESEARCH ARTICLE

A peer-reviewed open-access journal
PhytoKeys
Launched to accelerate biodiversity research

The systematics of the worldwide endoparasite family Apodanthaceae (Cucurbitales), with a key, a map, and color photos of most species

Sidonie Bellot[1], Susanne S. Renner[1]

1 *Systematic Botany and Mycology, University of Munich (LMU), Menzinger Strasse 67, 80638 Munich, Germany*

Corresponding author: *Sidonie Bellot* (sido.bellot@gmail.com)

Academic editor: *H. De Boer* | Received 26 February 2014 | Accepted 2 April 2014 | Published 30 April 2014

Citation: Bellot S, Renner SS (2014) The systematics of the worldwide endoparasite family Apodanthaceae (Cucurbitales), with a key, a map, and color photos of most species. PhytoKeys 36: 41–57. doi: 10.3897/phytokeys.36.7385

Abstract

Using morphological, nuclear, and mitochondrial data, we here revise the taxonomy of Apodanthaceae and allocate the 36 names published in the family to ten biological species in two genera, *Apodanthes* and *Pilostyles*. All species are endo-parasites that live permanently inside trees or shrubs of the families Salicaceae or Fabaceae and that only emerge to flower. Because of this life history, Apodanthaceae are among the least known families of flowering plants. Nevertheless, the World's herbaria as of 2013 hold at least 785 collections that, in combination with DNA phylogenies, permit well-founded species circumscriptions and geographic range maps. We also provide a key to all species, discuss the newly accepted or synonymized names, and make available color photos of six of the ten species.

Keywords

Apodanthaceae, genus circumscriptions, mitochondrial DNA sequences, nuclear DNA sequences, parasitic plants, species circumscriptions

Introduction

Apodanthaceae Tiegh. ex Takht. (Cucurbitales) is a family of endoparasites that live entirely in their host's stems and only become visible once the strictly unisexual flowers have burst through the bark. This life style, added to the small size of the flowers and

patchy occurrence of the apparently mostly dioecious populations, has made it difficult to collect good and complete herbarium material (including both sexes and flowering and fruiting specimens). While populations once identified may be recollected at the same time year after year, usually only local botanists will have the opportunity to carry out such recollections. Apodanthaceae are disjunctly distributed in North and South America, mainland Africa, Iran, and Australia. They occur in arid as well as humid tropical environments. Two genera have been validly described, the worldwide *Pilostyles*, and *Apodanthes* from Central and South America.

The taxonomy of the genera and species of Apodanthaceae has not been studied since the work of Ida de Vattimo-Gil (Vattimo-Gil 1950, 1955, 1971, 1973). Modern molecular-phylogenetic work based on representatives of most of the so-far named species (Bellot and Renner in review), together with study of collections deposited in the World's herbaria since the end of the 19th century, has revealed the need to synonymize many superfluous names, a task that we carry out here. We also up-date the circumscription of the family and its two genera, and clarify that they have specialized on different hosts, namely Salicaceae (mainly *Casearia*) and Fabaceae.

To achieve a better understanding of species boundaries and relationships, and to clarify the species' geographic and host ranges, we compared loans from numerous herbaria, dissected flowers, and isolated DNA from multiple collections. Molecular markers useful for these obligatory holoparasites are the nuclear 18S ribosomal RNA region and mitochondrial *matR* (Barkman et al. 2004; Bellot and Renner in review), and we show here that these markers can be used to place incomplete collections (for example, those of only one sex or only with fruits) in the correct species. Lastly, we provide an annotated key to all species that we recognize, and brief descriptions of their diagnostic traits along with color images and comments on their geographic and host ranges.

Methods

Plant material, DNA extraction and sequencing, phylogenetic analyses

We enlarged the DNA data matrix of Bellot and Renner (in review) by extracting DNA from additional specimens representing either unusual individuals or potential new species. No DNA sequences could be obtained from *P. stawiarskii*, known only from two collections in R, and *P. holtzii,* the only collection of which was destroyed in World War II. Suppl. material 1 shows species names and their authors, herbarium vouchers, and GenBank accession numbers. In total, 10 sequences (3 of 18S and 7 of *matR*) were newly generated for this study.

Total genomic DNA was extracted from herbarium specimens using the commercial plant DNA extraction Invisorb® Spin Plant Mini Kit (Stratec molecular, Berlin, Germany). The mitochondrial *matR* and the nuclear 18S genes were amplified using the primers listed in Bellot and Renner (in review). PCR products were purified with the ExoSAP or FastAP clean-up kits (Fermentas Life Sciences, St. Leon-Rot, Germany), and sequencing relied on the Big Dye Terminator v. 3.1 cycle sequencing kit (Applied

Biosystems, Foster City, CA, USA) and an ABI 3130-4 automated capillary sequencer. Chromatograms were checked and sequences were edited using Geneious R7 (Biomatters, available from http://www.geneious.com), and contigs were then blasted against GenBank to rule out contamination. Alignments of the clean sequences were performed using the program MAFFT v. 7 (Katoh 2013) resulting in matrices of 1626 and 1727 aligned nucleotides for *matR* and 18S, respectively. We failed to amplify the gene *matR* from the African *Pilostyles aethiopica* and from the Iranian *P. haussknechtii*. Phylogenetic reconstructions relied on maximum likelihood (ML) as implemented in RAxML-7.2.8-ALPHA (Stamatakis 2006), using the GTR + G model of nucleotide substitution with 100 bootstrap replicates under the same model. Trees were rooted on *Corynocarpus laevigatus* (Corynocarpaceae; Cucurbitales), based on Filipowicz and Renner (2010).

Morphological data and assessment of the host ranges of Apodanthaceae

We geo-referenced locality data from 785 herbarium collections on loan from the herbaria B, G, C, GH, K, M, MO, MSB, W, NA, PMA, and SI and added data from the Global Biodiversity Information Facility (GBIF Backbone Taxonomy, 2013-07-01, http://www.gbif.org/species/7279680). We also recorded host names, up-dating their taxonomy as relevant. All label information was compiled in a database using the Botanical Research and Herbarium Management System (BRAHMS, http://herbaria.plants.ox.ac.uk/), and maps were produced using DIVA-GIS 7.5 (http://www.diva-gis.org). Collections were sorted by geography, flowering specimens were sexed to evaluate sexual dimorphism, and a representative number of flowers were then dissected under a stereoscope. For each dissected flower, the first author recorded the number, arrangement and size of the tepals, shape and ornament of the pistil/central column, number of pollen sacs, presence of hairs and presence of a nectary at the base of the flower. Pictures of representative organs were taken using a Dino-Lite USB microscope model AM413ZT (Dino-Lite Europe) and the DinoCapture Imaging software version 2.0 of the same company.

Results and discussion

Genus and species boundaries in Apodanthaceae

The dissections showed that species have characteristic flower sizes, number of tepals, tepal cilia, and number of anthers rings. For the American species, we use these differences in the key (below). Suppl. material 2 shows measurements and counts from the 123 dissected flowers. Six collections could not reliably be assigned to these groups because their flowers were slightly unusual: *R. Callejas et al. 8062*, a male plant from Colombia identified as *Apodanthes casearia*e by A. Idarraga in 2002; *Y. Mexia 4540*, a female plant from Brazil that is the type of the name *A. minarum*; *H. S. Irwin et al. 20350*, a female plant from Brazil identified as *Pilostyles ulei* by Ida de Vattimo

in 1975; *H.S. Irwin 31560*, a male plant identified as *P. blanchetii* by the first author but parasitizing an uncommon host (*Dioclea*, Fabaceae); *J. Rzedowski 11303*, a female plant from Mexico identified by the collector as *P. thurberi*; and *F. Chiang 9034*, a female plant from Mexico identified as *P. thurberi* by J. Henrickson in 1972.

The 18S and *matR* molecular trees show the *Pilostyles* collections that we wanted to identify (in red on Fig. 1) grouped with *P. thurberi* or *P. blanchetii*. The collections *R. Callejas et al. 8062* and *Y. Mexia 4540* grouped with two undoubted representatives of *A. casearіae*. *R. Callejas et al. 8062* is a male plant and comes from the border with Panama, a country where *A. casearіae* has been repeatedly collected. The host of *R. Callejas et al. 8062* was originally identified as *Trema* (Cannabaceae), but a partial *matR* sequence of this host BLASTed to *Casearia nitida*, making it likely that the host was in fact a *Casearia*. If that is the case, this would suggest that the collection represents an *Apodanthes*. The few male flowers of *Apodanthes casearіae* that have so far been dissected (Suppl. material 2) do not allow assessing the full morphological variability of the male flowers of this species. Therefore we had to rely on DNA for identification. In terms of its *matR* (Fig. 1A) *R. Callejas et al. 8062* was embedded among other sequences of *A. casearіae*, while in terms of its 18S (Fig. 1B), it was sister to them. We identified the specimen as *A. casearіae*. Other *matR* and 18S sequences in the *A. casearіae* clade are from the type of the name *A. minarum* (*Mexia 4540*) from Brazil. Its host was a *Casearia* and its (female) flowers match those of *A. casearіae* (Suppl. material 2). We therefore synonymize *A. minarum* under *A. casearіae* (an action carried out below).

In combination, the present morphological and molecular results show that Apodanthaceae comprise at least ten biological species that can be allocated to two mutually monophyletic genera. In the Americas, these are *Pilostyles thurberi* in the southern United States of America and Mexico, *P. mexicana* in Mexico, Guatemala and Honduras, the widely distributed *P. blanchetii* from Panama to Jamaica to Brazil and Uruguay, and *P. berteroi* in Chile and Argentina. The Americas also harbor *A. casearіae* from Guatemala to Brazil (Fig. 2). Australia has three species, *P. coccoidea*, *P. collina*, and *P. hamiltonii*; Iran has *P. haussknechtii,* and Africa has *P. aethiopica*. The second African species, *P. holtzii* has not been recollected since 1907, when its type collection was made. Another species, the southern Brazilian *P. stawiarskii,* is only known from two specimens (one of them the type) collected at the same locality in Jan./Feb. 1948 and Dec. 1949; morphologically it resembles *P. blanchetii* (Vattimo, 1950). The host ranges of our accepted genera and species do not overlap. *Apodanthes* parasitizes only Salicaceae, whereas *Pilostyles* parasitizes only Fabaceae. As seen on Figure 3, there is a correspondence, although not perfect, between the phylogenies of host genera and parasitic species, and host specialization may have played a role in speciation of Apodanthaceae. At the species level, Table 1 shows that species of Apodanthaceae can grow on one or up to thirteen host species. As seen in Figures 2 and 3, our species concepts are corroborated by geographic and host ranges, except in the case of *Apodanthes casearіae* and *Pilostyles blanchetii*, both widespread in Brazil. These two species have different sized flowers (see below), and parasitize phylogenetically distantly related hosts (Fig. 3).

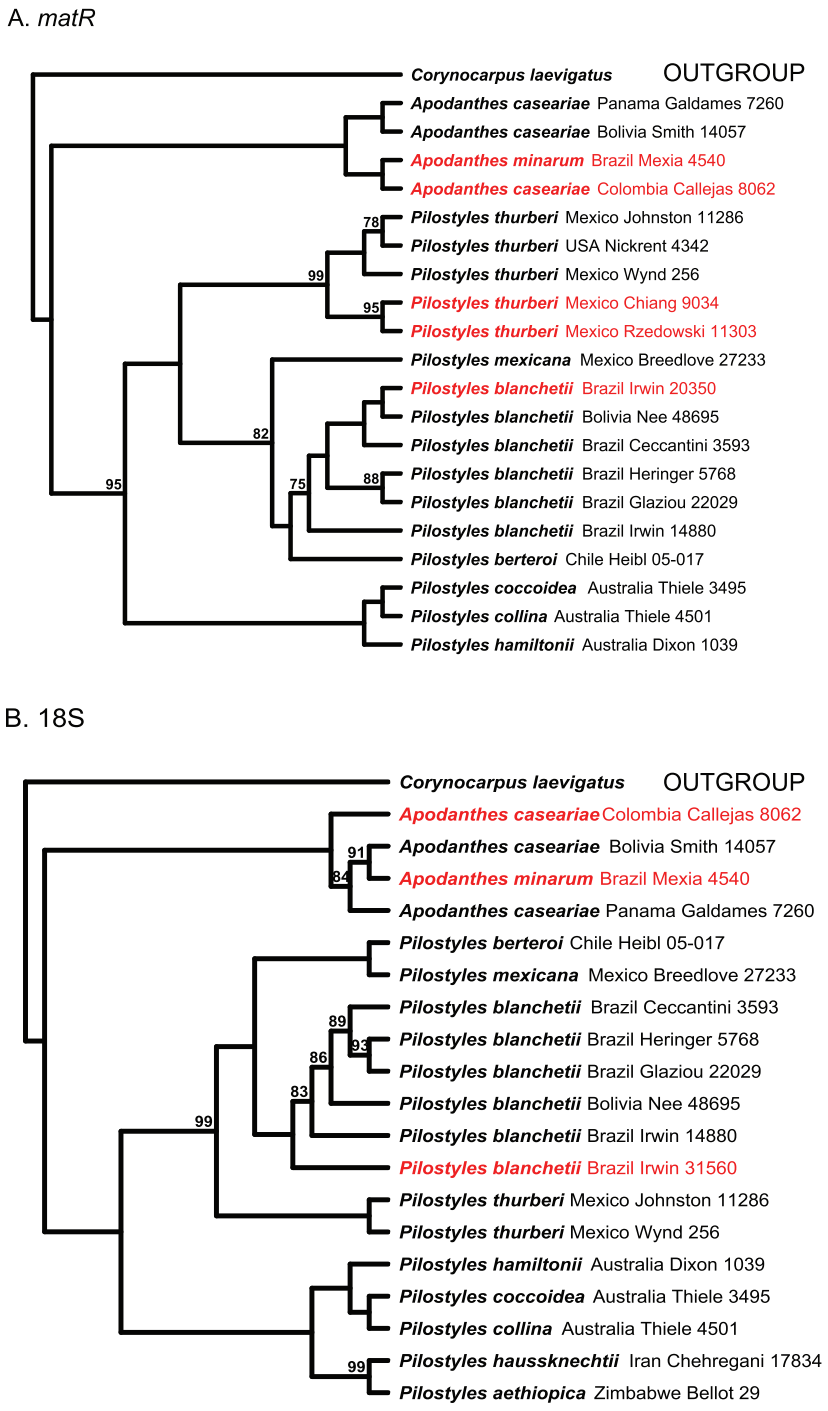

Figure 1. Phylogenetic relationships in Apodanthaceae obtained under maximum likelihood from the mitochondrial gene *matR* (**A**) and the nuclear ribosomal RNA gene 18S (**B**). Shown in red are the specimens we wanted to identify to species. Numbers indicate bootstrap support >70%.

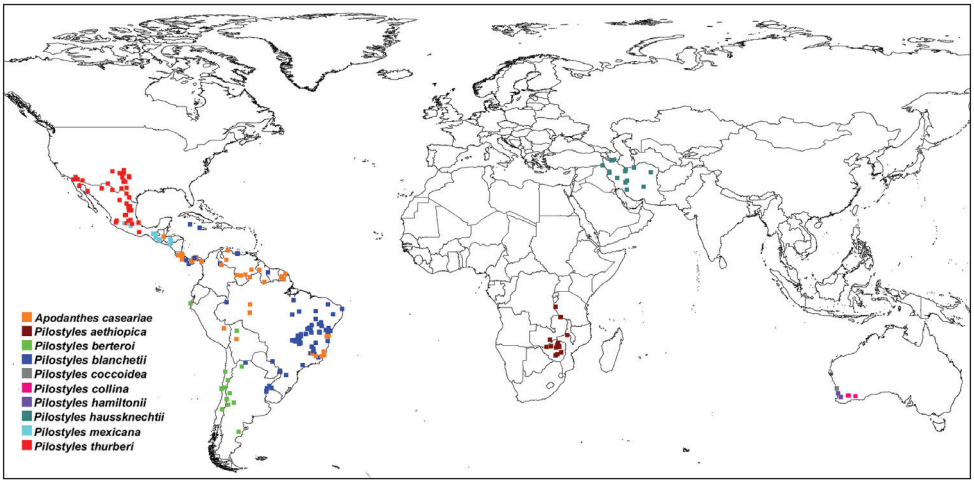

Figure 2. Geographic distribution of Apodanthaceae based on label information from 785 herbarium collections.

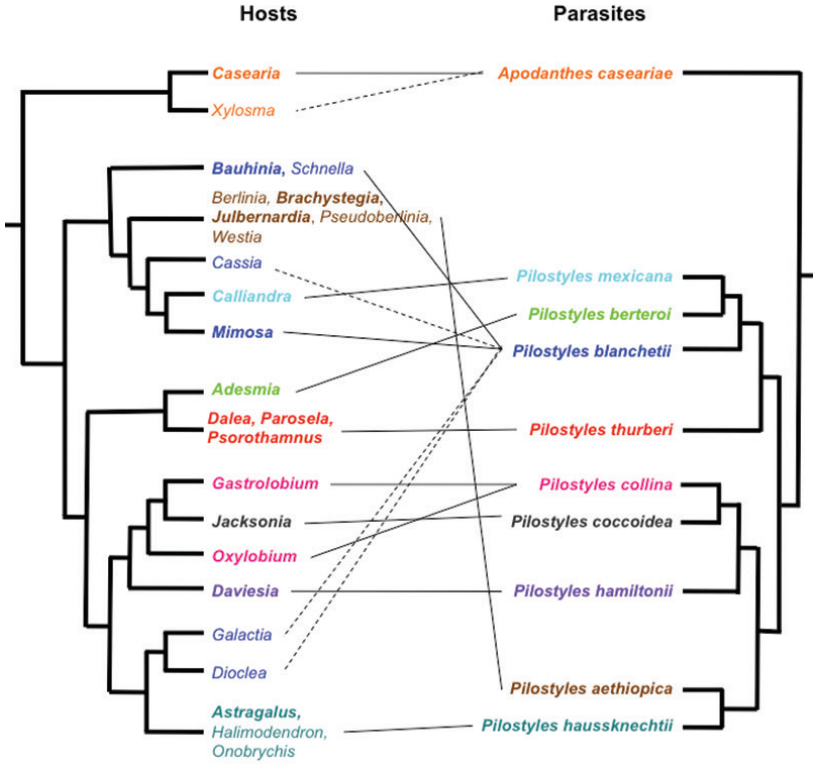

Figure 3. Phylogenetic relationships among the hosts of Apodanthaceae (legume relationships from Wojciechowski et al. 2006) and among the species of Apodanthaceae (from Bellot and Renner, in review). Identical colors link parasite species and their host(s) and are also used in Figure 2. Dashed lines represent associations with rarely reported hosts; hosts in bold are the most common ones.

Table 1. Hosts of Apodanthaceae based on label information from 785 herbarium collections. Upper case numbers refer to the references below the table.

Parasite	Host genera	Host species
P. aethiopica	*Berlinia*[1], *Brachystegia, Julbernardia*, *Pseudoberlinia*[1], *Westia*[1]	*Brachystegia boehmii* Taub., *Brachystegia glaucescens* x *spiciformis*, *Brachystegia spiciformis* Benth., *Brachystegia taxifolia* Harms., *Julbernardia globiflora* (Benth.) Troupin
P. berteroi	*Adesmia*	*Adesmia arborea* Bert. ex Savi, *Adesmia* aff. *spinosissima* Meyen, *Adesmia obovata* Clos, *Adesmia bedwellii* Skottsb., *Adesmia miraflorensis* Remy, *Adesmia uspallatensis* Gill ex H. & A., *Adesmia gracilis* Meyen ex Vogel, *Adesmia microphylla* H. & A., *Adesmia monosperma* Clos, *Adesmia pinifolia* Gillies, *Adesmia trijuga* Gillies
P. blanchetii	*Bauhinia, Cassia, Dioclea, Galactia*[2], *Mimosa, Schnella*	*Bauhinia candicans* Benth., *Bauhinia divaricata* L., *Mimosa claussenii* Benth., *Mimosa cyclophylla* Taub., *Mimosa* aff. *setosa* Benth., *Mimosa maguirei* Barneby, *Mimosa scabrella* Benth., *Mimosa setosissima* Taub., *Mimosa uraguensis* H. & A., *Mimosa* cf. *xanthocentra* Martius, *Schnella cumanensis* Britton & Rose
A. caseariae	*Casearia, Xylosma*	*Casearia aculeate* Jacq., *Casearia arborea* Urb., *Casearia decandra* Jacq., *Casearia grandiflora* Cambessèdes, *Casearia guianensis* Urb., *Casearia hirsute* Swartz, *Casearia nitida* Jacq.
P. coccoidea	*Jacksonia*	
P. collina	*Gastrolobium, Oxylobium*	*Gastrolobium euryphyllum* Chandler & Crisp
P. hamiltonii	*Daviesia*	*Daviesia angulata* Benth., *Daviesia decurrens* Meissner, *Daviesia pectinata* Meissner, *Daviesia preissii* Lindley
P. haussknechtii	*Astragalus, Halimodendron, Onobrychis*	*Astragalus brachycalyx* Fisch., *Astragalus brachystachys* DC., *Astragalus cephalanthus* DC., *Astragalus chalaranthus* Boiss. & Hausskn., *Astragalus compactus* Reiche, *Astragalus floccosus* Boiss., *Astragalus gossypinus* Fisch., *Astragalus microcephalus* Willd., *Astragalus rhodosemius* Boiss. & Hausskn., *Astragalus spinosus* Muschler, *Astragalus susianus* Boiss., *Astragalus verus* Olivier, *Halimodendron halodendron* (Pall.) Druce
P. mexicana	*Calliandra*	*Calliandra houstoniana* (Miller) Standley
P. thurberi	*Dalea, Psorothamnus, Parosela*[3]	*Dalea bicolor* Humb. & Bompl. in Willd., *Dalea formosa* Torrey, *Dalea frutescens* Gray, *Psorothamnus emoryi* (Gray) Rydberg

[1]Verdcourt, B., 1998. Flora of tropical East Africa - Rafflesiaceae. Flora of tropical East Africa 175, 1–2. CRC Press.

[2]Ule, E., 1915. Rafflesiaceae. Notizblatt des Königl. botanischen Gartens und Museums zu Berlin-Dahlem 6, 292–293.

[3]Rose, J. N., 1909. Studies of Mexican and Central American Plants n°6. Contributions from the United States National Herbarium 7, 26–265.

Description of the family

Stem-endoparasites, non-photosynthetic. No leaves, stem or roots, instead an endophytic system of cells inside the stem parenchyma of the host, flowers bursting through the host bark. Flowers unisexual, plants dioecious or monoecious, a point still insufficiently known; flowers of both sexes on the same host or not. Pollination by flies and bees, possibly also wasps (Bellot and Renner 2013; Sipes et al. 2014), based on the fruit color and size, dispersal is probably by birds. Flowers white or yellow (*Apodanthes*), or white, pink, orange, red, purple or brown (*Pilostyles*),

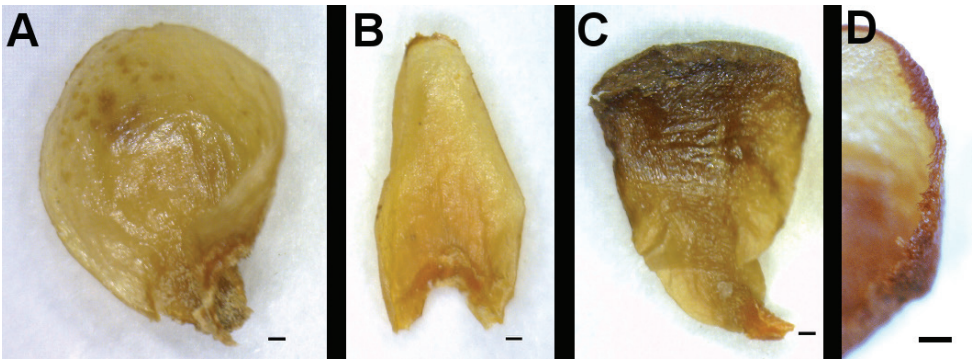

Figure 4. Tepals of Apodanthaceae. **A** Tepal of the outer whorl of *Apodanthes casearіae* **B** Tepal of the middle whorl of *A. casearіae* **C** Tepal of the inner whorl of *A. casearіae* **D** Tepal margin of *A. casearіae*. The scale bar corresponds to 0.2 mm.

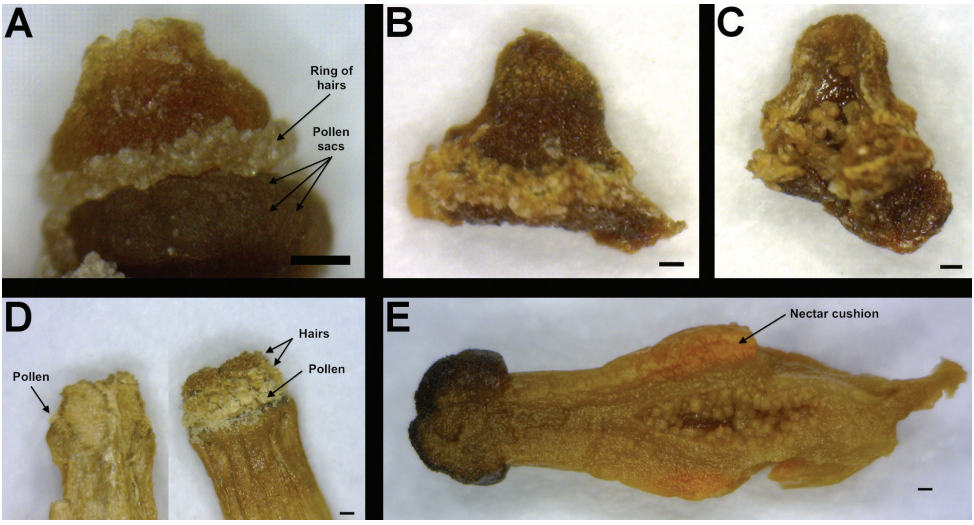

Figure 5. Sexual organs of Apodanthaceae from rehydrated herbarium material. **A** Androecium of *Pilostyles haussknechtii*, note the two rings of pollen sacs topped by a ring of hairs **B** Style and stigma of *P. haussknechtii* **C** Ovary locule and ovules of *P. haussknechtii* **D** Androecium of *Apodanthes casearіae* after bursting of the pollen sacs, note the hairs covering the column apex **E** Gynoecium of *A. casearіae*. The scale bar corresponds to 0.2 mm.

aggregated on the host stems, minute (1.5 to 15 mm long when dried), usually with radial symmetry. Perianth composed of 2 or 3, rarely 4, whorls of tepals (Fig. 4A–C), the latter sometimes with hairs along their margins (Fig. 4D), or a hair cushion at their basis (Blarer et al. 2004). In male flowers, the staminal filaments completely fused and forming a tube around a central column that is usually fused to the column (Fig. 5A, D), the up to 72 pollen sacs arranged in 1–4 rings around the column apex (Fig. 5A), the column apex dome-shaped and circled or covered by single-celled hairs (Fig. 5A, D), a basal nectar cushion in both sexes (Fig. 5E). Female flowers without staminodes and with a single thick style topped by the

dome-shaped stigma (Fig. 5B, C, E). Ovary semi-inferior, placentation parietal with 50–300 ovules (Fig. 5C, E). Seeds ca. 0.5 mm long (Bouman and Meijer 1994). Fruit a fleshy berry.

Key to the genera and species of Apodanthaceae

1a From the Neotropics, parasitizing Salicaceae, tepals always in 3 whorls with (from the outside) 2, 4, and 4 tepals, the inner whorl easily deciduous, female flowers >5 mm long ..... **1. *Apodanthes casearia***

1b From the Neotropics, Africa, Iran, or Australia, parasitizing Fabaceae, number of outer tepals usually >2, female flowers <5 mm long **2**

2a Occurring in Australia **3**

2b Not in Australia **5**

3a Tepals in 3 whorls ..... **2. *Pilostyles collina***

3b Tepals in 2 whorls **4**

4a Flower diameter >3 mm ..... **3. *P. hamiltonii***

4b Flower diameter <3 mm ..... **4. *P. coccoidea***

5a Occurring in Africa ..... **5. *P. aethiopica***

5b Not in Africa **6**

6a Occurring in Iran ..... **6. *P. haussknechtii***

6b Occurring in the America 7

7a Tepals in 3 whorls, each with 2 to 7 tepals, anthers in 4 whorls (spiral), number of anther lobes >70, on *Adesmia* ..... **7. *P. berteroi***

7b Tepals in 3 whorls, each with 3 or 4 (rarely more) tepals, anthers in 2 or 3 whorls, number of anther lobes <70, not on *Adesmia* **8**

8a Middle tepals ovoid, on *Calliandra*, *Dalea*, *Parosela* or *Psorothamnus*, anthers in 3 whorls (at least if on *Dalea*, *Parosela* or *Psorothamnus*) **9**

8b Middle tepals mostly diamond-shaped, apparently never on *Calliandra*, *Dalea*, *Parosela* or *Psorothamnus*, anthers in 2 whorls (females of the three species cannot be securely distinguished) ..... **8. *P. blanchetii***

9a On *Calliandra*, tepals in 3 whorls, each with 4 tepals ..... **9. *P. mexicana***

9b On *Dalea*, *Parosela*, or *Psorothamnus*, tepals in 3 whorls, each with 3 or 4 tepals. ..... **10. *P. thurberi***

Allocation of all species names so far described in Apodanthaceae

1. *Apodanthes casearia* Poiteau, Ann. Sci. Nat. (Paris) 3: 422, t. 26. 1824.
http://species-id.net/wiki/Apodanthes_casearia

Apodanthes flacourtiae Karsten, Linnaea 28: 413. 1857. Type: Venezuela, Aragua, Choroni, parasitic on "Flacourtieae" [most like a species that today would be placed in Salicaceae], *H. Karsten s.n.* (W, destroyed in WWII), **syn. nov.**

Apodanthes tribracteata Rusby, Descr. S. Amer. Pl. 15. 1920. Type: Bolivia, near Inglis-Inglis, 8 Aug. 1902, *R. S. Williams 1580* (NY), **syn. nov.**

Nom. inval. *Apodanthes matogrossensis* Vattimo, Vattimo-Gil, Rodriguésia 26(38): 45. 1971, without Latin descr. Type: Brazil, Mato Grosso, parasitic on *Casearia, J. G. Kuhlmann 53076* (R, not seen).

Apodanthes panamensis Vattimo-Gil, Rodriguésia 26(38): 45. 1971, without Latin descr., Latin diagnosis in Rev. Brasil. Biol., 33(1): 140. 1973. Type: Panama, Canal Zone, Aug. 1984, *R. E. Woodson Jr. and R. W. Schery 965* (NY, MO). Already synonymized by A. Gentry (1973).

Apodanthes surinamensis Pulle, Recueil Trav. Bot. Néerl. 6: 259. 1909. Type: Suriname, along the Marowijne River, July-Dec. 1903, parasitic on Flacourtiaceae [most like a species that today would be placed in Salicaceae], *G. M. Versteeg s.n.* (U0007645), **syn. nov.**

Apodanthes roraimae Ida de Vattimo, Rodriguésia 29(44): 48-49. 1978. Type: Brazil, Roraima, 24 Jul. 1974, parasitic on Flacourtiaceae [most likely a species that today would be placed in Salicaceae], *G. T. Prance et al. 21353* (NY), **syn. nov.** Comment: George Yatskievych, a curator at the Missouri Botanical Garden, also studied the NY isotype in 2004 and annotated it as *A. caseariae.*

Apodanthes minarum Vattimo-Gil, Rodriguésia 26 (38): 45. 1971, without Latin descr.; Latin diagnosis in Rev. Brasil. Biol., 33(1): 140. 1973. Type: Brazil, Minas Gerais, Viçosa, 31 Mar. 1930, *Y. Mexia 4540* (L, MO), **syn. nov.**

Type. French Guiana, Karouany, c. 1802, parasitic on *Casearia* spec., *P. A. Poiteau s.n.* (P: P00686413).

Note. Tepals white to yellow, female flowers >5 mm long, tepals in 3 whorls, the outer with 2 tepals, the inner one easily deciduous (Figs 4A–D; 6C, D). Growing in trunk and branches of *Casearia* and occasionally *Xylosma* (Salicaceae, Fig. 3) in Guatemala, Honduras, Costa-Rica, Panama, Colombia, Venezuela, Suriname, French Guiana, Brazil, Peru and Bolivia (Fig. 2).

2. *Pilostyles collina* Dell, Nuytsia 4: 293–294. 1983.
http://species-id.net/wiki/Pilostyles_collina

Type. Australia, Western Australia, Peak Charles, 10 Jan. 1982, parasitic on *Oxylobium*, *B. D. Dell 8216* (G, MO).

Note: Tepals orange to red, in 3 whorls. Growing in young stems of *Gastrolobium* and *Oxylobium* in Western Australia (Figs 2, 3, see Thiele et al. 2008 for pictures of flowers).

3. *Pilostyles hamiltonii* Gardner, J. Roy. Soc. Western Australia 32: 77. 1948.
http://species-id.net/wiki/Pilostyles_hamiltonii

Type. Australia, Western Australia, Darling District, Helena Rover, Mundaring Weir, Mar. 1946, parasitic on *Daviesia pectinata* Lindl., *C. D. Hamilton s.n.* (PERTH, not seen).

Note: Tepals dark burgundy, in 2 whorls, flowers >3 mm in diameter. Growing in young stems of *Daviesia* in Western Australia (Figs 2, 3, see Thiele et al. 2008 for pictures of flowers).

4. *Pilostyles coccoidea* K.R.Thiele, Nuytsia 18: 273–284. 2008.
http://species-id.net/wiki/Pilostyles_coccoidea

Type. Australia, Western Australia, Waddi Road, 30°33'26"S, 115°28'10"E, 7 Mar. 2008, parasitic on *Jacksonia*, *K.R. Thiele 3495* (PERTH 07692447).

Note. Tepals pale orange to brown, in 2 whorls, flowers <3 mm in diameter. Growing in stems of *Jacksonia* in Western Australia (Figs 2 and 3, see Thiele et al. 2008 for pictures of flowers).

5. *Pilostyles aethiopica* Welwitsch, Trans. Linn. Soc. London 27: 66–70. 1871 = *Berlinianche aethiopica* (Welw.) Vattimo-Gil, nom. inval.
http://species-id.net/wiki/Pilostyles_aethiopica

Pilostyles holtzii Engler, Bot. Jahrb. Syst. 46: 293. 1912 = *Berlinianche holtzii* (Engl.) Vattimo-Gil, not validly published. Type: Tanzania, Kilimatinde, July 1907, parasitic on *Berlinia eminii* Taub., *W. Holtz 1422* (B, destroyed during World War II), **syn. nov.** (based on the protologue).

Syntypes. Angola, Huila, 12 May 1860, parasitic on *Berlinia paniculata* Benth. = *Pseudoberlinia paniculata* (Benth.) P.A.Duvign., *F. M. J. Welwitsch 529, 529b* (C, G).

Note. Tepals pink to brown, in 3 to 4 whorls each with 3-6 tepals. Male flowers with 1 or 2 ring(s) of ca. 15 pollen sacs, stamen filaments free from the central column (Fig. 6F), hair cushion at the basis of the inner tepals (Blarer et al. 2004). Growing in branches of *Julbernardia* and *Brachystegia*, maybe also on *Berlinia*, *Westia* and *Pseudoberlinia*, in Zimbabwe, Zambia, Tanzania, Angola and Malawi (Figs 2, 3).

6. *Pilostyles haussknechtii* Boissier, Arch. Sci. Phys. Nat. 25: 255–261. 1866.
http://species-id.net/wiki/Pilostyles_haussknechtii

Type. Middle East, parasitic on *Astragalus*, *J. E. Haussknecht s.n* (G-BOISS, not seen).

Note: Tepals pink to brown in 2 whorls, each with 6 to 10 tepals (Fig. 6G). Found at the basis of young branches of *Astragalus* and occasionally *Onobrychis* and *Halimodendron* in Iran (Figs 2, 3).

7. *Pilostyles berteroi* Guillemin, Ann. Sci. Nat., Bot., sér. 2, 2: 21. 1834 = *Apodanthes berteroi* (Guill.) Gardner, Hooker's Icon. Pl. 7: t. 655. 1844. syn. nov.
http://species-id.net/wiki/Pilostyles_berteroi

Syntypes. Chile, Quillota, parasitic on *Adesmia*, [in Chile 1828-1831] *C. L. G. Bertero s.n.* (P, not seen); Chile [from the collection number this was in "various localities in the Andes", during the period from 27 Oct.-26 Nov. 1841], *T. Bridges 1273* (BM, not seen, K, not seen).

Note: Tepals purple to brown with clearer margins (Fig. 6A), 9–18 in number, stamens in 4 whorls (spirals), with > 70 pollen sacs. Growing in older stems of *Adesmia* shrubs in Chile, Argentina, Peru, and Bolivia (Figs 2, 3). Our morphological (Suppl. material 2) and molecular data (Fig. 1) show that the species is nested among other species of *Pilostyles*, indicating that Gardner's transfer was erroneous.

8. *Pilostyles blanchetii* (Gardner) R.Br., Trans. Linn. Soc. London 19(3): 247. [6 Nov 1844] = *Apodanthes blanchetii* Gardner, Icon. Pl. 7: t. 655 b. 1844 [Jul 1844] = *Frostia blanchetii* (Gardner) H.Karst., Nov. Actorum Acad. Caes. Leop.-Carol. Nat. Cur. 26: 922. 1858.
http://species-id.net/wiki/Pilostyles_blanchetii

Pilostyles calliandrae (Gardner) R.Br., Trans. Linn. Soc. London 19(3): 247. [6 Nov 1844] = *Apodanthes calliandrae* Gardner, Icon. Pl. 7: t. 644. 1844 [Jan 1844] = *Frostia calliandrae* (Gardner) H. Karst., Nov. Actorum Acad. Caes. Leop.-Carol. Nat. Cur. 26: 921. 1858. Type: Brazil, Amazonas, near Maynas [Manaus], Feb. 1840, *G. Gardner 3639* (K000601222), **syn. nov.**

Pilostyles caulotreti (Karsten) Hook.f., Prodr. (DC.) 17: 116. 1873 = *Sarna caulotreti* Karsten, Linnaea 28: 415. Jun 1857 [1856]. Type: Venezuela, *H. Karsten s.n.* (W, destroyed in WWII). Comment: Gentry (1973) considered this name as synonym of *P. blanchetii*, and we agree with this assessment.

Pilostyles ingae (Karsten) Hooker f., Prodr. (DC.) 17: 116. 1873 = *Sarna ingae* H.Karst., Linnaea 28: 415. Jun 1857 [1856]. Type: Colombia, Cauca, Popayán, parasitic on *Inga, H. Karsten s.n.* (W, destroyed in WWII), **syn. nov.** (based on the protologue).

Pilostyles galactiae Ule, Notizbl. Königl. Bot. Gart. Berlin 6: 292. 1915. Type: Brazil, Amazonia, Surumu River, tributary of the Rio Branco, Oct. 1909 and Mar. 1910, parasitic on *Galactia jussiaeana* Kunth., *E. Ule 7895* (B, holotype destroyed in WWII; isotype NY), **syn. nov.**

Pilostyles goyazensis Ule, Ber. Deutsch. Bot. Ges. 33: 475. 1915. Syntypes (all parasitic on *Mimosa*): Brazil, Goias, region near city of Corumba, Sobradinho, Aug. 1892, *E. Ule 3097*; Serra dos Pyreneos, Mun. Corumba, Dec. 1892, *E. Ule 3098*; same location, Dec. 1892, *E. Ule 3099*; in the Corumba region, Aug. 1892, not flowering, *E. Ule s.n.*; Serra dos Pyreneos, Aug. 1892, not reproductive, *E. Ule s.n.* (all in B, material destroyed in WWII), **syn. nov.** (based on the protologue).

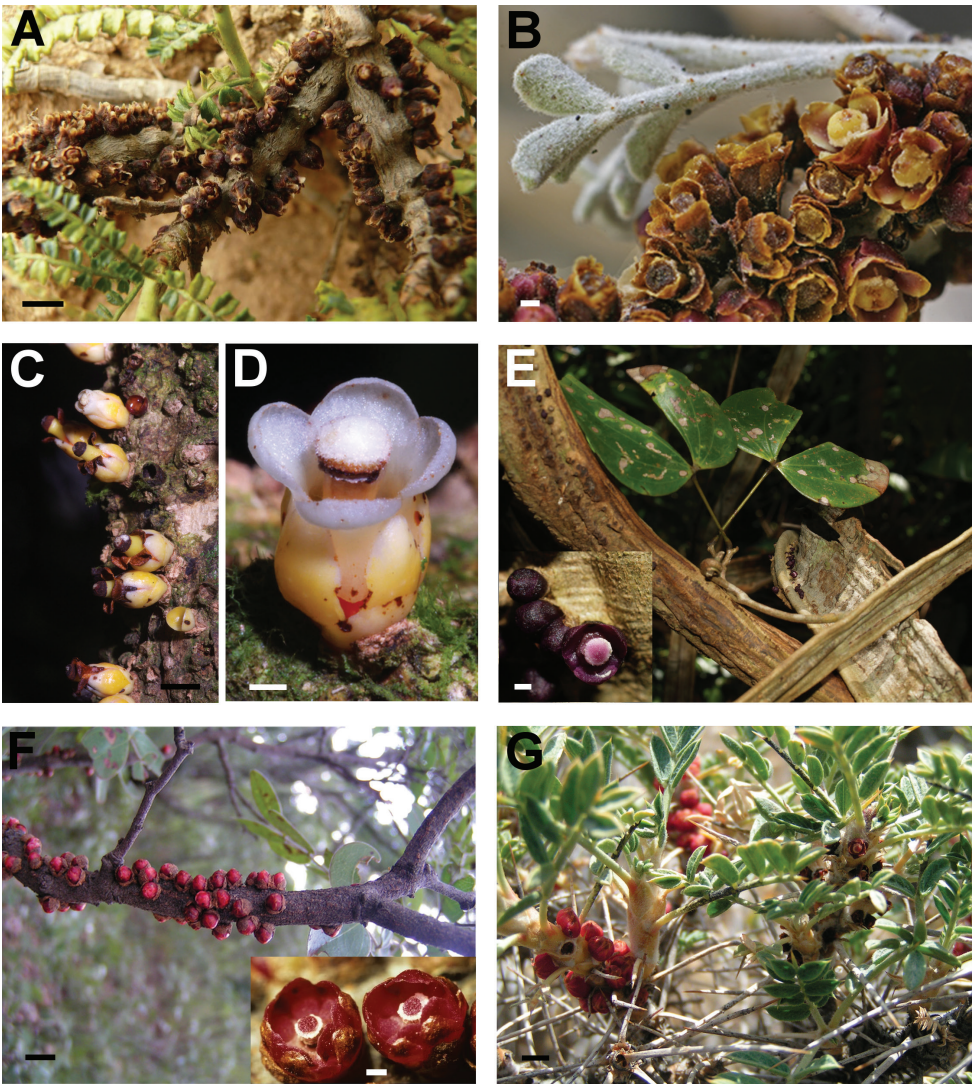

Figure 5. Flowers of Apodanthaceae species **A** *Pilostyles berteroi* on *Adesmia* in Chile (picture by C. Heibl) **B** Male flowers of *P. thurberi* on *Psorothamnus emoryi* in California (picture by L. Hendrickson) **C** Male and female flowers of *Apodanthes casearíae* on *Casearia* (?) in Panama (picture by G. Gerlach) **D** Close-up of a male flower of *A. casearíae* (picture by G. Gerlach) **E** *Pilostyles blanchetii* on *Bauhinia* in Panama; insert: close-up of a flower (pictures by C. Galdames) **F** *Pilostyles aethiopica* on *Julbernardia globiflora* in Zimbabwe (picture by S. Bellot); insert: close-up of male flowers (picture by D. Plowes) **G** Female and male flowers of *P. haussknechtii* on Astragalus in Iran (picture by S. Bellot). The white scale bars correspond to 1 mm and the black ones to 5 mm.

Pilostyles globosa (S.Watson ex Robinson) Hemsl., J. Linn. Soc., Bot. 31: 311. 1896 = *Apodanthes globosa* S.Watson ex Robinson., Bot. Gaz. 16: 83, tab. 9, 1891. Type: Mexico, Northern part, Sierra Madre, parasitic on *Bauhinia lunarioides* A. Gray, *C. G. Pringle 1950* (G), **syn. nov.**

Pilostyles stawiarskii Vattimo-Gil, Revista Brasil. Biol. 10: 196. 1950. Type: Brazil, Paraná, Mun. de Palmas, parasitic on *Mimosa scabrella* Benth. [incl. its synonym *Mimosa bracaatinga* Hoehne], Jan. 1948 and Feb. 1948, *V. Stawiarski* R50.591 and 50.592 (R, photos). There is also a topotypical collection from Dec. 1949, **syn. nov.** (based on the protologue).

Pilostyles ulei Solms-Laub., in Goebel, Organogr. Pfl. 2,1: 434. Figure 292 (1900), descr. in Endriss, Flora, Ergänz.-Bd. 91: 209. 1902. Type: Brazil, Goias, parasitic on Fabaceae, *E. Ule s.n.* (B, destroyed in WWII; R has E. Ule 34, E. Ule 36, E. Ule 38, E. Ule 148, E. Ule 367, E. Ule 482, and E. Ule 483 labeled as this species, not seen). Comment: already Solms-Laubach (1901) and Endriss (1902) considered *P. ulei* as a synonym of *P. ingae.*

Type. Brazil: Bahia, 1839, *J. S. Blanchet 2861* (NY).

Note. Tepals purple to brown sometimes with clearer margins (Fig. 6E), in 3 whorls with usually 4 (rarely 3-6) tepals, the middle tepal diamond-shaped. Stamens in 2 whorls. Found in branches of *Mimosa* and *Bauhinia*, but also *Cassia*, *Dioclea*, *Galactia* and *Schnella*, in Jamaica, Cayman Islands, Costa-Rica, Panama, Colombia, Venezuela, Guyana, Brazil, Argentina and Uruguay (Figs 2, 3).

9. *Pilostyles mexicana* (Brandegee) Rose, Contr. U.S. Natl. Herb. 12(7): 264. 1909 = *Apodanthes mexicana* Brandegee, Zoe 5(11): 245. 1908.
http://species-id.net/wiki/Pilostyles_mexicana

Type. Mexico, near Zacuapan, Tenampa, parasite on *Calliandra grandiflora* Benth., Oct. 1906, *C.A. Purpus 2207* (NY).

Note. Tepals red to brown, in 3 whorls, each with 4 tepals. Growing in branches of *Calliandra* in Guatemala, Honduras and Mexico (possibly further south; Figs 2, 3).

10. *Pilostyles thurberi* Gray, Pl. Nov. Thurb. 326–327. 1854.
http://species-id.net/wiki/Pilostyles_thurberi

Pilostyles covillei Rose, Contr. U.S. Natl. Herb. 12: 263. 1909. Type: USA, Texas, Dickens county, Matador ranch, 14 June 1894, parasitic on *Parosela formosa* (Torr.) Vail, *F. V. Coville 1860* (US, not seen).

Pilostyles glomerata Rose, Contr. U.S. Natl. Herb. 12: 263. 1909. Type: Mexico, Puebla, near Tehuacán, 31 Aug. 1905, parasitic on *Parosela, J. N. Rose and J. H. Painter 9942* (NY, G). The protologue gives the collection number as 8942.

Pilostyles palmeri Rose, Contr. U.S. Natl. Herb. 12: 263. 1909. Type: Mexico, San Luis Potosí, near Alvarez, May 1887, parasitic on *Parosela*, *E. Palmer 584* (US-570088).

Pilostyles pringlei (Watson) Hemsl., J. Linn. Soc., Bot. 31: 311. 1896 = *Apodanthes pringlei* Watson ex B.L.Rob., Bot. Gaz. 16: 83, tab. 9. 1891, no Latin descr.; *Pilo-*

styles pringlei (Watson) Rose, Contr. U.S. Natl. Herb. 12: 264. 1909, superfluous transfer. Type: Mexico, Sierra Madre, near Monterey, 27 June 1888, parasitic on *Dalea frutescens* A. Gray, *C. G. Pringle 1949* (NY, M, G).

Pilostyles sessilis Rose, Contr. U.S. Natl. Herb. 12: 263. 1909. Type: male flowers: Mexico, Hidalgo, Ixmiquilpan, 1905, parasitic on *Parosela*, *J. N. Rose 9041* (NY); female flowers: Mexico, Querétaro, hacienda Ciervo, 20 Aug. 1905, parasitic on *Parosela tuberculata* (Lag.) Rose, *J. N. Rose and J. H. Painter 9636* (NY, US).

Type. USA, probably Arizona, near Gila river, June 1850, parasitic on *Psorothamnus emoryi* (A. Gray) Rydb., *G. Thurber 682* (NY).

Note. Tepals white, red to brown, in 3 whorls, each with 3 or 4 tepals, rarely more (Fig. 6B). Growing in branches of *Dalea*, *Parosela* and *Psorothamnus* in the southern United States of America and Mexico (Figs 2, 3). New York (NY) has a specimen from Mexico of this species annotated as "*Pilostyles mortoni*", a nomen nudum, by Ida de Vattimo in 1952.

Note on an invalid genus name

Harms (1935) tried to place the two African names, *Pilostyles aethiopica* Welw. and *P. holtzii* Engl., in a separate section, *Pilostyles* section *Berlinianche*, named for their legume host species in the genus *Berlinia*, but failed to include a Latin diagnosis for the new section. Later, Vattimo-Gil (1955, 1971) decided to rank this section as a separate genus because of the hair cushions on the inner perianth whorl and strictly tri- and hexamerous flowers compared to the tetramerous flowers of the American species of *Pilostyles*. This assessment, however, could only have been based on specimens of *P. aethiopica*, since the only collection of *P. holtzii* burnt in World War II. Unfortunately, Vattimo-Gil also neglected to provide a Latin diagnosis, and the genus name is therefore not valid. Based on our results (Fig. 1), *P. aethiopica* does not deserve generic status because it is embedded among the other species of *Pilostyles*.

Note on a possible new species of *Pilostyles*

Flavio González and Natalia Pabón-Mora, at the university of Antioquia in Colombia, are studying the ecology and morphology of Apodanthaceae in Colombia (González and Pabón-Mora accepted a) and are describing a new species of *Pilostyles* (González and Pabón-Mora accepted b). This species is the first *Pilostyles* parasitizing the legume genus *Dalea* in South America and occurs in dry valleys of the Colombian Eastern Cordillera at altitudes above 2000 m. Morphologically, the new species is most similar to *P. berteroi*, which grows in the Chilean and Peruvian Andes at up to 3000 m of altitude (Fig. 2) and parasitizes *Adesmia* (closely related to *Dalea*, see Fig. 3).

Conclusion

By combining morphological and molecular information, we show that Apodanthaceae comprise 10 species and that morphological distinctions fit well with geographical disjunctions and specializations on different hosts (Salicaceae vs. Fabaceae). DNA sequences of mitochondrial *matR* and nuclear 18S rDNA, along with morphology and geography permit identifying any collection of Apodanthaceae. A wider sampling of the morphological variation, especially of male *Apodanthes casearіae* and female *Pilostyles blanchetii*, *P. mexicana* and *P. thurberi*, however, is needed to determine whether some unusual specimens might deserve to be ranked as subspecies.

Acknowledgements

We thank C. Galdames, G. Gerlach, C. Heibl, L. Hendrickson, D. Plowes for photos of Apodanthaceae, K. R. Thiele for providing Australian silica-dried material, the reviewers N. Filipowicz and D. Nickrent for corrections, and the herbaria listed in *Materials and Methods* for providing loans. Funding came from the German Research Foundation (DFG RE 603/16-1).

References

Barkman TJ, Lim S-H, Salleh KM, Nais J (2004) Mitochondrial DNA sequences reveal the photosynthetic relatives of *Rafflesia*, the world's largest flower. Proceedings of the National Academy of Sciences of the United States of America 101: 787–792. doi: 10.1073/pnas.0305562101

Bellot S, Renner SS (2013) Pollination and mating systems of Apodanthaceae and the distribution of reproductive traits in parasitic angiosperms. American Journal of Botany 100: 1083–1094. doi: 10.3732/ajb.1200627

Bellot S, Renner SS (in review) Exploring new dating approaches for parasites: the worldwide Apodanthaceae (Cucurbitales) as an example. Molecular Phylogenetics and Evolution.

Blarer A, Nickrent DL, Endress PK (2004) Comparative floral structure and systematics in Apodanthaceae (Rafflesiales). Plant Systematics and Evolution 245: 119–142. doi 10.1007/s00606-003-0090-2

Bouman F, Meijer W (1994) Comparative structure of ovules and seeds in Rafflesiaceae. Plants Systematics and Evolution 193: 187–212. doi: 10.1007/BF00983550

Endriss W (1902) Monographie von *Pilostyles ingae* (Karst.) *Pilostyles ulei* Solms-Laub. Flora, Ergänz.-Bd. 91: 209–236, 1 plate [Diss. Zürich.].

Filipowicz N, Renner SS (2010) The worldwide holoparasitic Apodanthaceae confidently placed in the Cucurbitales by nuclear and mitochondrial gene trees. BMC Evolutionary Biology 10: 219. doi: 10.1186/1471-2148-10-219

Gentry AH (1973) Flora of Panama, Part IV. Family 50A. Rafflesiaceae. Annals of the Missouri Botanical Garden 60: 17–21. doi: 10.2307/2394766

González F, Pabón-Mora N (accepted a) First reports and generic descriptions of the achlorophyllous holoparasites Apodanthaceae (Cucurbitales) of Colombia. Revista Actualidades Biológicas.

González F, Pabón-Mora N (accepted b) *Pilostyles boyacensis*, a new species of Apodanthaceae (Cucurbitales) from Colombia. Phytotaxa.

Harms H (1935) Rafflesiaceae. In: Engler A, Prantl K (Eds) Die natürlichen Pflanzenfamilien (ed. 2). Engelmann, Leipzig, 243–281.

Katoh S (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30: 772–780. doi: 10.1093/molbev/mst010

Sipes S, Huff Hartz KE, Amin H, Anterola A, Nickrent DL (2014) Floral scent and pollinators of the holoparasite *Pilostyles thurberi* (Apodanthaceae). Journal of Pollination Ecology 12: 31–39.

Solms-Laubach H (1901) Rafflesiaceae. In: Engler A (Ed) Das Pflanzenreich IV. Engelmann, Leipzig, 75.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690. doi: 10.1093/bioinformatics/btl446

Thiele KR, Wylie SJ, Maccarone L, Hollick P, McComb JA (2008) *Pilostyles coccoidea* (Apodanthaceae), a new species from Western Australia described from morphological and molecular evidence. Nuytsia 18: 273–284.

Vattimo I de (1950) *Pilostyles stawiarskii*. Revista Brasileira de Biologia 10: 196.

Vattimo I de (1955) Notice sur la tribu Apodantheae R.Br. (Rafflesiaceae). Taxon 4: 211–212. doi: 10.2307/1217760

Vattimo I de (1971) Contribuição ao conhecimento da tribo Apodanthea R. Br. Parte I – Conspecto das espécies (Rafflesiaceae). Rodriguésia 38: 37–62.

Vattimo I de (1973) Notas sobre o genero *Apodanthes* Poit. com descrição de duas novas espécies (Rafflesiaceae). Revista Brasileira de Biologia 33: 140.

Wojciechowski MF, Mahn J, Jones B (2006) Fabaceae legumes. Version 14 June 2006. http://tolweb.org/Fabaceae/21093/2006.06.14 *in* The Tree of Life Web Project, http://tolweb.org/ [accessed 02.2014]

Supplementary material 1

List of accessions used in this study with author names for each species, herbarium vouchers, and GenBank accession numbers

Authors: Sidonie Bellot, Susanne S. Renner

Data type: species list

Link: doi: 10.3897/phytokeys.36.7385.app1

Supplementary material 2

Results of the dissections of 123 flowers from 82 specimens of Apodanthaceae

Authors: Sidonie Bellot, Susanne S. Renner

Data type: measurements

Link: doi: 10.3897/phytokeys.36.7385.app2

Table S1. List of accessions used in this study with author names for each species, herbarium vouchers, and GenBank accession numbers

Species name	Voucher (Herbarium) 18S / matR	Source 18S / matR	GenBank accession 18S	GenBank accession *matR*
Apodanthes casearіae Poiteau	R. Callejas 8062 (NY)	This paper	KJ634137	KJ634133
Apodanthes casearіae Poiteau	C. Galdames 7260 (M)	Bellot and Renner, in review	KJ634128	KJ634100
Apodanthes casearіae Poiteau	Y. Mexia 4540 (MO)	This paper	KJ634138	KJ634134
Apodanthes casearіae Poiteau	D. N. Smith et al. 14057 (MO)	Bellot and Renner, in review	KJ634127	KJ634101
Probably *Casearia sp.*	Host bark of R. Callejas 8062 (NY)	This paper	-	KJ634135
Corynocarpus laevigatus J.R.Forst & G.Forst	[1]M. W. Chase 236 (NCU); [2]M. W. Chase s.n. (NCU)	[1]Soltis et al., 2000; [2]Zhu et al., 2007	AF206892[1]	AY121499[2]
Pilostyles aethiopica Welwitsch	S. Bellot 29 (M)	Bellot and Renner, in review	KJ634129	-
Pilostyles berteroi Guill.	J. N. Rose and L. B. Rose 18849 (NY)	This paper	-	KJ634136
Pilostyles berteroi Guill.	C. Heibl 05-017 (M)	Bellot and Renner, in review	KJ634126	KJ634102
Pilostyles blanchetii (Gardner) Brown	H. S. Irwin et al. 20350 (NY)	This paper	-	KJ634132
Pilostyles blanchetii (Gardner) Brown	E. P. Heringer et al. 5768 (MO)	Bellot and Renner, in review	KJ634125	KJ634103
Pilostyles blanchetii (Gardner) Brown	H. S. Irwin et al. 31560 (MO)	This paper	KJ634139	-
Pilostyles blanchetii (Gardner) Brown	H. S. Irwin et al. 14880 (MO)	Bellot and Renner, in review	KJ634122	KJ634106
Pilostyles blanchetii (Gardner) Brown	A. Glaziou 22029 (G)	Bellot and Renner, in review	KJ634119	KJ634108
Pilostyles blanchetii (Gardner) Brown	G. T. Ceccantini et al. 3593 (SPF)	Bellot and Renner, in review	KJ634114	KJ634113
Pilostyles blanchetii (Gardner) Brown	M. Nee et al. 48695 (NY)	Bellot and Renner, in review	KJ634115	KJ634112
Pilostyles coccoidea K.R.Thiele	K. R. Thiele 3495 (PERTH)	Bellot and Renner, in review	KJ634124	KJ634104
Pilostyles collina B.Dell	K. R. Thiele 4501 (PERTH)	Bellot and Renner, in review	KJ634123	KJ634105
Pilostyles hamiltonii C.A.Gardner	D. Dixon 1039 (PERTH)	Bellot and Renner, in review	KJ634121	KJ634107
Pilostyles haussknechtii Boissier	A. Chehregani and S. Zarre 17834 (M)	Bellot and Renner, in review	KJ634120	-
Pilostyles mexicana (Brandegee) Rose	D. E. Breedlove 27233 (NY)	Bellot and Renner, in review	KJ634118	KJ634109
Pilostyles thurberi Gray	F. Lyle Wynd and C. H. Mueller 256 (NY)	Bellot and Renner, in review	KJ634117	KJ634110
Pilostyles thurberi Gray	F. Chiang 9034 (NY)	This paper	-	KJ634131
Pilostyles thurberi Gray	M. C. Johnston et al. 11286 (MO)	Bellot and Renner, in review	KJ634116	KJ634111
Pilostyles thurberi Gray	D. L. Nickrent 4342 (SIU)	Nickrent et al., 2004	-	AY739003
Pilostyles thurberi Gray	J. Rzedowski 11303 (A)	This paper	-	KJ634130

Nickrent, D.L., Blarer, A., Qiu, Y.L., Vidal-Russell, R., Anderson, F.E., 2004. Phylogenetic inference in Rafflesiales: the influence of rate heterogeneity and horizontal gene transfer. BMC Evol. Biol. 4, 40.

Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, W.J., Nixon, K.C., Farris, J.S., 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. Botanical Journal of the Linnean Society 133, 381–461.

Zhu, X.-Y., Chase, M.W., Qiu, Y.-L., Kong, H.-Z., Dilcher, D.L., Li, J.-H., Chen, Z.-D., 2007. Mitochondrial matR sequences help to resolve deep phylogenetic relationships in rosids. BMC Evolutionary Biology 7, 217.

Table S2. Results of the dissections of 123 flowers from 82 specimens of Apodanthaceae. Only the headers and the first dissected flowers are present here, the rest is available online at http://phytokeys.pensoft.net/articles.php?id=1521&element_type=5&display_type=element&element_id=32

Number	Observations	Genus	Species	Collector	Collection-number	Herbarium	Host	Country / Area	Sex	Flower shape; G: globose, O: ovoid, hO: horizontally ovoid	Flower length (from base to apex in mm)	Flower width (from side to side in mm)	Number of tepals whorls	Number of segments in outer whorl (oW)	Shape of segments of oW; L: losange, O: ovoid, hO: horizontally ovoid	Length of segments of oW (mm)	Width of segments of oW (mm)	Relative colour of segments of oW compared to mW	Number of segments in middle whorl (mW)
pile 1																			
1	Tepal edges ciliated and clearer. Fruit (seeds)	Pilostyles	blanchetii	Anderson	37104	NY	Bauhinia	Br	F	G	3	3	3	4	L	2	2	same	4
2	Tepal edges ciliated and clearer	Pilostyles	ulei	Irwin	31560	NY	Dioclea	Br	M	G	2	2	3	4	L	1-1.5	1.5	same	4 (1 lost)
2b	Tepal edges ciliated and clearer	Pilostyles	ulei	Irwin	31560	NY	Dioclea	Br	M	G	2	2	3	4	L	1.5	1.5	same	4
3	No cils. fruit (seeds)	Pilostyles	ulei	Nee	48695	NY	Mimosa	Bol	F	G	2.5	2.5	3	4	L	1.5	1.5	same	4
3b	No cils. Less mature than 3	Pilostyles	ulei	Nee	48695	NY	Mimosa	Bol	F	G	2.5	2.5	3	4	L	1	1	same	3, 4?
4	tepals ciliated on the edge	Pilostyles	goyazensis	Irwin	14880	MO	Mimosa	Br	F	G	3	3	3	4	L	1.5-2	1.5-2	same	4
4b	tepals ciliated on the edge	Pilostyles	goyazensis	Irwin	14880	MO	Mimosa	Br	F	G	3	3	3 (4)	4 (2+2)	L	1.5-2	1.5-2	same	4
6	No cils	Pilostyles	ulei	Delprete	10494	NY	Mimosa	Br	F	G	2	2	3	4	L	1-1.5	1-1.5	same	4
7	tepals ciliated on the edge	Pilostyles	blanchetii	del Puerto	3171b	B	Bauhinia	Uru	F	O	2.5	2	3 (4)	4 (2+2)	L	1.5	1.5	same	4
5	tepals ciliated on the edge, especially Ti	Pilostyles	blanchetii	Liesner	10950	MO	Bauhinia	Ven	F	G	3	3	3 (4)	4 (2+2)	L	2	2-2.5	same	4

Shape of segments of mW; L: losange, O: ovoid, hO: horizontally ovoid	Length of segments of mW (mm)	Width of segments of mW (mm)	Number of segments in inner whorl (iW)	Shape of segments of iW; L: losange, O: ovoid, S: spatulate	Length of segments of iW (mm)	Width of segments of iW (mm)	Relative colour of segments of iW	Nectary position; C: around column, T: between tepals, bTi: base internal tepals, bTm: base middle tepals	Male column shape: #c: number of cylinders, e: exact, t: trapezoidal, i: inner cylinder, o: outer cylinder	Male column hairs position; Ab: apex basis, aA: around apex, A: all apex	Male column length (incl. apex in mm)	Male column largest diameter (mm)	Male column apex diameter (mm)	Male column apex shape; d: dome, fd: flat dome, f: flat	Male column apex ornament; b: similar to taste buds	Male column apex hole shape; c: circular hole, s:slit	Number of stamens whorls	Stamens fusion; L: free, F: fused (where they are fused)	Number of anther lobes
L	1.5-2	1.5-2	4	O	1.5	1-1.5	same	C, T, bTi, bTm, bTo											
L	1.5	1.5	4	L?	1	1	clearer	C, T, bTi	1ce	Ab	1	0.5 (almost invisible under anthers)	0.7	d	b	no hole	2?	F filaments	40?
O	1.5	1	4	L	0.5-1.5	0.5-1	clearer	C, T, bTi	1ce	Ab	1	0.5 (almost invisible under anthers)	0.8	d	b	no hole	2	F filaments	36-40
L	1-1.5	1-1.5	4	L	1.5	1.5	same	C, T, bTi, bTm, bTo											
L	1.5	1.5	4	L	1	1	clearer	C, T, bTi, bTm, bTo											
L	1.5	1.5	4	L	1.5-2	1.5-2	same	C, T, bTi, bTm, bTo											
O, L	1.5-2	1.5	4 (1 lost)	L	1.5	1.5	same	C, T, bTi, bTm, bTo											
L	1	1	4	L	1-1.5	1-1.5	same	C, T, bTi, bTm, bTo											
L	1.5	1.5	4	L	1.5	1.5	clearer	C, T, bTi, bTm, bTo											
L	2	2-2.5	4	L	1.5	1.5	same	C, T, bTi, bTm, bTo											

Carpels number	Gynoecium type; S: syncarp, A: apocarp	Locule shape; C: circular, O: ovoid, e: elliptic, hO: horizontally ovoid	Locule length (mm)	Locule width (mm)	Ovary position; S: supere, I: infere, HI: hemi-infere	Placenta number	Placenta position; P: parietal, C: central	Number of ovules (approximative)	Style length (incl. Stigma, in mm)	Style width (mm)	Style shape; C: cylinder, T: trapezoidal	Stigma shape; d: dome, fd: flat dome, c: cylinder, s: sphere, p:pyramid, td: depression on the top	Stigma diameter (mm)	Stigma ornament; b: similar to taste buds, h: hairs around the stigma, hb: hairs at the around the lower part of the stigma	Stigma hole shape; c: circular hole, s:slit
NA	S	C	2	2	HI	NA	P	100	1.5	1	C	d	1	b+h	no hole
NA	S	C	2	2	HI	NA	P	250 (counted precisely)	0.7	0.7	T	d	0.3	b+h	S
NA	S	C	1.5	1.5	HI	NA	P	150?	0.5	1	T	d	0.5	b+h	S
NA	S	C	2	2	HI	NA	P	150?	1	1	C	fd	0.7	b+h	C
NA	S	C	1.5	1.5	HI	NA	P	150	1.5	1	C	d	0.7	b+h	C
NA	S	C	1.5	1.5	HI	NA	P	100	1	0.7	C	fd	0.7	b+h	no hole
NA	S	C	1.5	1.5	HI	4?	P	150 (counted precisely)	1	0.8	C	d	0.8	b+h	no hole
NA	S	C	2	2	HI	NA	P	150?	1	1	C	d	0.8	b+h	S

Chapter 3

POLLINATION AND MATING SYSTEMS OF APODANTHACEAE AND THE DISTRIBUTION OF REPRODUCTIVE TRAITS IN PARASITIC ANGIOSPERMS

Bellot, S. and S. S. Renner

American Journal of Botany 100: 1083–109

American Journal of Botany

American Journal of Botany 100(6): 1083–1094. 2013.

SPECIAL INVITED PAPER—EVOLUTION OF PLANT MATING SYSTEMS

Pollination and mating systems of Apodanthaceae and the distribution of reproductive traits in parasitic angiosperms[1]

Sidonie Bellot[2] and Susanne S. Renner[2]

Systematic Botany and Mycology, University of Munich (LMU), Menzinger Str. 67 80638 Munich, Germany

- *Premise of the study:* The most recent reviews of the reproductive biology and sexual systems of parasitic angiosperms were published 17 yr ago and reported that dioecy might be associated with parasitism. We use current knowledge on parasitic lineages and their sister groups, and data on the reproductive biology and sexual systems of Apodanthaceae, to readdress the question of possible trends in the reproductive biology of parasitic angiosperms.
- *Methods:* Fieldwork in Zimbabwe and Iran produced data on the pollinators and sexual morph frequencies in two species of Apodanthaceae. Data on pollinators, dispersers, and sexual systems in parasites and their sister groups were compiled from the literature.
- *Key results:* With the possible exception of some Viscaceae, most of the ca. 4500 parasitic angiosperms are animal-pollinated, and ca. 10% of parasites are dioecious, but the gain and loss of dioecy across angiosperms is too poorly known to infer a statistical correlation. The studied Apodanthaceae are dioecious and pollinated by nectar- or pollen-foraging Calliphoridae and other flies.
- *Conclusions:* Sister group comparisons so far do not reveal any reproductive traits that evolved (or were lost) concomitant with a parasitic life style, but the lack of wind pollination suggests that this pollen vector may be maladaptive in parasites, perhaps because of host foliage or flowers borne close to the ground.

Key words: Apodanthaceae; dioecy; field observations; fly pollination; phylogeny; sexual system.

A ground-breaking study published 5 years ago revealed that among flowering plants, parasitism evolved at least 13 times (Barkman et al., 2007), namely in the Apodanthaceae, Balanophoraceae, *Cassytha* (Lauraceae), *Cuscuta* (Convolvulaceae), Cynomoriaceae, *Cytinus* (Cytinaceae), Hydnoraceae, *Krameria* (Krameriaceae), Lennoaceae, *Mitrastemon* (Mitrastemonaceae), Rafflesiaceae, higher Santalales, and Orobanchaceae (excluding the first-branching genus *Lindenbergia*; Fig. 1). Together, these clades may comprise some 4500 species, 2040 of them in the Orobanchaceae and another 2100 in the Santalales (Table 1 provides species numbers and references). The phylogenetic placements of 12 of the 13 parasite lineages found by Barkman et al. (2007) have been upheld in studies with denser gene or taxon sampling (Nickrent, 2007; Vidal-Russell and Nickrent, 2008a, b; Filipowicz and Renner, 2010). Only the placement of Cynomoriaceae, which was unsupported in the preferred topology of Barkman et al., remains unclear. The single genus of this family has been placed in Myrtales, Rosales, Santalales, Saxifragales, and Sapindales (Nickrent et al., 2005; Jian et al., 2008; Qiu et al., 2010; Zhang et al., 2009), but the situation is not yet settled (Zhang et al., 2011).

With the independent parasitic lineages of flowering plants now clear, it is timely to return to the question of traits that may be associated with the evolution of parasitism. This question was last addressed in 1995 in a review of the reproductive biology of parasitic angiosperms that found no discernible trends unique to parasitic plants (Molau, 1995). The review followed the classification most accepted at the time (Cronquist, 1988), which lumped several parasite groups since revealed to be independent, such as Apodanthaceae and Rafflesiaceae, and Balanophoraceae and Cynomoriaceae, while separating others since revealed to have a single ancestor, such as Eremolepidaceae and Santalaceae. The focus of the review also was not evolutionary since relevant sister group relationships were not yet known.

An obvious parallel evolutionary trend in parasitic angiosperms is the reduction of the capacity to photosynthesize. The loss of photosynthetic capacity is complex because genes coding for photosynthesis are partly located in the chloroplast genome, partly in the nucleus, and plastids are involved in other metabolic pathways than those directly related to photosynthesis (Neuhaus and Emes, 2000; Krause, 2012). All holoparasites investigated so far probably therefore still have a plastid genome (Table 1: column 4). Other life history or morphological traits that could have undergone parallel evolution in parasites concern the timing of reproduction (especially in seasonal

[1]Manuscript received 6 December 2012; Revision accepted 25 March 2013.

The authors thank J. Kuijt for discussion of parasite biology; D. Plowes, M. Coates Palgrave, and M. Hyde for help with fieldwork in Zimbabwe; D. Plowes, B. Schlumpberger, and G. Gerlach for photos; S. Zarre and Y. Salmaki for help with fieldwork in Iran; M. Correa and C. Galdames for information on species from Panama; K. Thiele for information from Australia; G. Ceccantini for information from Brazil, B. Muller, T. Dikow, S. Marshall, T. Pape, K. Rognes, E. Kameneva, D. Doczkal, and M. Kotrba for identifications of flies; M. Sharaf and B. Taylor for ant identifications; and D. Plowes and an anonymous reviewer for pointing us to important literature.

[2]Authors for correspondence (e-mails: sido.bellot@neuf.fr, renner@lmu.de)

doi:10.3732/ajb.1200627

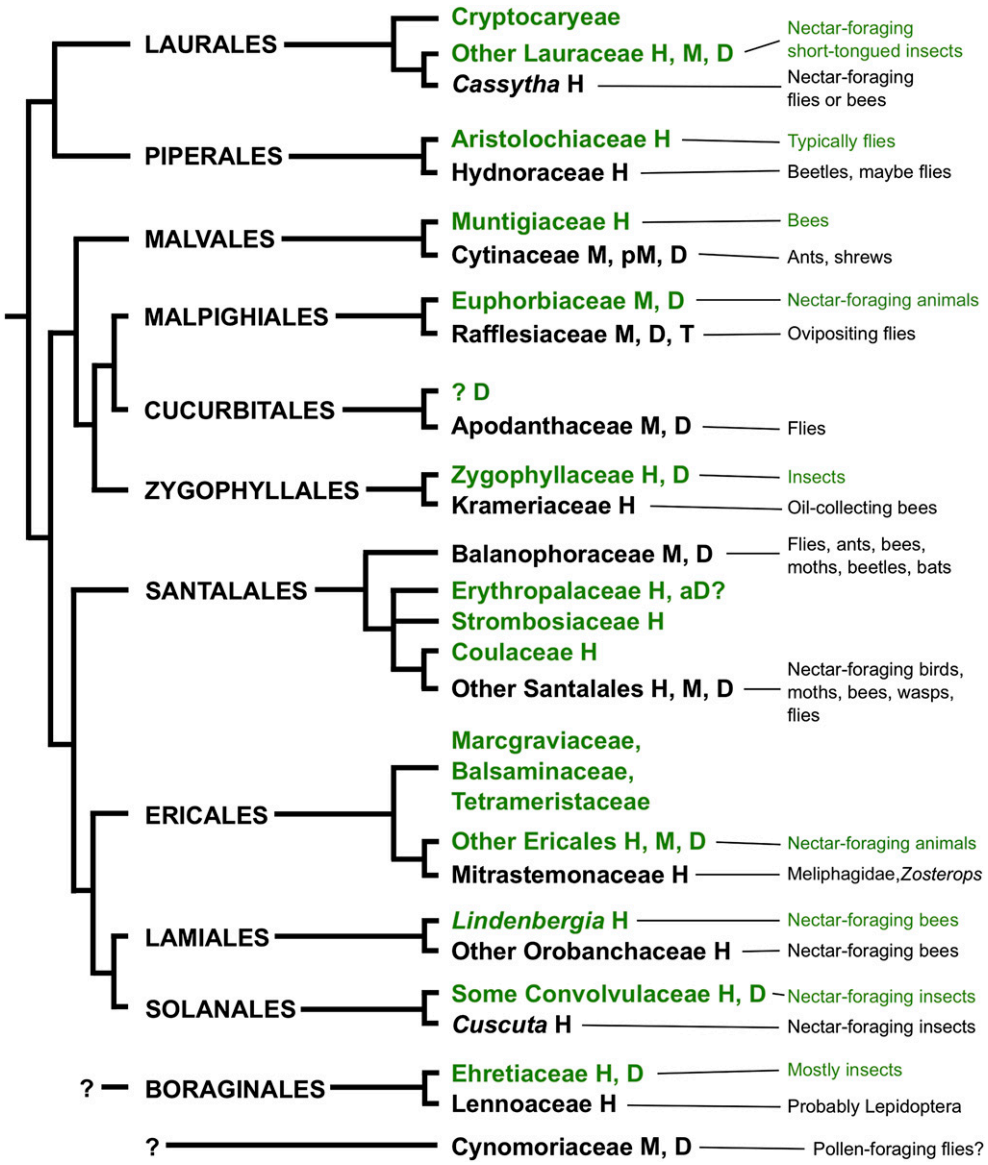

Fig. 1. Phylogenetic relationships, sexual systems, and main pollinators of the 13 parasitic angiosperm lineages and their sister groups. Autotrophic lineages are shown in green, parasitic ones in black. aD = androdioecy, D = dioecy, H = hermaphrodism, M = monoecy, pM = polygamomonoecy. Question marks indicate uncertainty. See Table 1 for references.

climates), the mode of seed dispersal (Kuijt, 1969), or conceivably also the mode of pollination. For example, animal pollination might be more reliable than wind pollination if parasite flowers are often located near the ground or under host foliage.

Because most parasitic lineages can only be studied in natural conditions, not in the laboratory or botanical gardens, knowledge of their reproductive biology is patchy. Pollinators of Cynomoriaceae and Lennoaceae have never been observed, and the first data on the pollination of Apodanthaceae are reported in the present study. Parallel or convergent evolution in the reproductive modes of parasitic angiosperms is thus an open question. In terms of sexual systems, a study published 17 yr ago suggested an overrepresentation of dioecy among parasitic angiosperm genera. This was based on the finding that of 959 angiosperm genera with dioecious species, at least 43 (4%) also contained parasitic species. Of the angiosperm genera lacking dioecious species, only 1% contained parasitic species (Renner and Ricklefs, 1995). A seeming correlation between dioecy and a parasitic life style was also noted in the flora of the southeastern United States: ca. 17% of Carolina heterotrophs and mycoheterotrophs are dioecious (Conn et al., 1980). More recently, an analysis of diversification rates in parasitic plants (Hardy and Cook, 2012) revealed that a parasitic life style may contribute to high extinction risks (for hosts and parasites alike), and this risk might be compounded by a dioecious sexual system in which species may suffer a competitive disadvantage because only half of the individuals in a dioecious population are seed bearing (Heilbuth et al., 2001; Vamosi and Vamosi, 2004). It would therefore be surprising if dioecy were indeed overrepresented among parasitic angiosperms.

Here we use today's data on the placement of the 13 parasite families and their reproductive biology to readdress the question

TABLE 1. Biological traits of parasitic clades and their sister clades analyzed in this review.

Parasitic clades: No. of genera/ No. of species	Flower phenotype	Population sexual system, Minimum no. of dioecious species	Plastid genome investigated	Pollination	Fruit or seed dispersal	Longevity	Sister group and no. of genera/ no. of species	Flower phenotype in sister group	Population sexual system in sister group	Pollination in sister group
Apodanthaceae: 2/18	Unisexual	Monoecious or dioecious 5	Ongoing: *Pilostyles aethiopica* and *P. hamiltonii* (this study)	Flies (this study)	Birds or other animals eating small fruits	Perennial	Cucurbitales (all or some) up to 2600 species in 109 genera	Mostly unisexual[2]	Mostly monoecious[2]	Biotic or abiotic[3]
Balanophoraceae: 16/42	Unisexual[4]	Monoecious or dioecious[4], 21	None	Flies, wasps, ants, moths, beetles[6, 65]; bees[66]; bats[65]	Rodents, rain wash[6]	Unknown	Other Santalales 106–138/1552–3304 parasites + 11–13/44–83 autotrophs	Bisexual or unisexual[8]	Hermaphrodite, monoecious or dioecious (or other sexual systems)[8]	Nectar-foraging birds, moths, bees, wasps, flies[58]
Cassytha (Lauraceae): 1/21	Bisexual[12]	Hermaphrodite[12]	None	Nectar-foraging flies or bees, extrapolating from flower morphology	Fruit-eating animals[58]	Perennial[14]	Lauraceae without Cryptocaryeae ca. 40/1730[57]	Unisexual or bisexual[15]	Hermaphrodite, monoecious or dioecious[15]	Nectar-foraging short-tongued insects[44]
Cuscuta (Convolvulaceae): 1/200[43]	Bisexual[42]	Hermaphrodite[42]	Published: Cuscuta *reflexa*, *C. exaltata, C. obtusiflora*, *C. gronovii*[61, 54]	Shallow tubular flowers, small amounts of nectar visited by nectar foragers[42]	Birds[51, 58]; water[50]	Mostly annual[42, 43]	Somewhere in Convolvulaceae[52]	Mostly bisexual but also unisexual[7]	Mostly hermaphrodite, some dioecious[7]	Nectar-foraging animals
Cynomorium (= Cynomoriaceae): 1/1	Unisexual and bisexual[16]	Monoecious or dioecious[16]	None	Flies feeding on pollen?	Possibly rain[58]	Perennial[17]	Some Saxifragales lineage?	?	?	?
Cytinaceae: 2/10	Unisexual or bisexual[18]	Monoecious, polygamo-monoecious[55] or dioecious[18], 6	None	Ants[18]; shrews[11]	Rodents and ants[19]; beetles[67]; mechanical[58]	Perennial[18]	Muntingiaceae[20] 3/3[57]	Bisexual[20]	Hermaphrodite [20]	Bees[21]
Hydnoraceae: 2/9	Bisexual[7]	Hermaphrodite[7]	None	Based on morphology and scent, prob. flies[22] or beetles[22, 64]	Mammals[22]	Perennial[23]	Aristolochiaceae[24] 5–8/480[57]	Bisexual[25, 7]	Hermaphrodite,[25, 7]	Typically flies[26]
Krameria (= Krameriaceae): 1/18	Bisexual[27]	Hermaphrodite[27]	None	Oil-collecting bees[27]	Nuts barbed with spines[27]	Perennial[27]	Zygophyllaceae[29] 22/285[57]	Mostly bisexual[7]	Hermaphrodite or dioecious[7]	Insects[7]
Lennoaceae 2/4	Bisexual[27]	Hermaphrodite[27]	None	Probably Lepidoptera[63]	Wind, rain[62, 63]	Perennial and annual[27, 53]	Ehretiaceae[30] 8/170[57]	Bisexual or unisexual[31]	Hermaphrodite or dioecious[31]	Insects, maybe sometimes abiotic[31]
Mitrastemon (= Mitrastemonaceae): 1/2	Bisexual[12]	Hermaphrodite[12]	None	Meliphagidae[1] *Zosterops*[58]	Birds, rain[32]	Perennial[58]	Ericales excluding Balsaminaceae, Marcgraviaceae, and Tetrameristaceae[33] 10275 species[33]	Bisexual or unisexual[34]	Hermaphrodite, monoecious or dioecious[34]	Mostly animal-pollinated[35]
Orobanchaceae (excluding *Lindenbergia*): 96/2040[57]	Bisexual[12]	Hermaphrodite[12]	Published: *Epifagus virginiana*[60]; ongoing: 7 other species[49]	Nectar-foraging bees[58]	Grazing mammal[48]; rain[12]	Perennial and annual[12]	*Lindenbergia*[46] (Orobanchaceae) 3-30	Bisexual	Hermaphrodite	Nectar-foraging bees

Table 1. Continued.

Parasitic clades: No. of genera/ No. of species	Flower phenotype	Population sexual system, Minimum no. of dioecious species	Plastid genome investigated	Pollination	Fruit or seed dispersal	Longevity	Sister group and no. of genera/ no. of species	Flower phenotype in sister group	Population sexual system in sister group	Pollination in sister group
Rafflesiaceae: 3/34	Bisexual or unisexual[36, 37]	Monoecious, dioecious or trioecious[36, 37], 5–30	None	Ovipositing flies[59]	Mammals[32]	Perennial?	Euphorbiaceae[38] 218/5735[57]	Unisexual[45]	Monoecious or dioecious[45]	Mostly biotic
Santalales (Amphorogynaceae, Aptandraceae, Cervantesiaceae, Comandraceae, Loranthaceae, Misodendraceae, Nanodeaceae, Octoknemaceae, Olacaceae, Opiliaceae, Santalaceae, Schoepfiaceae, Thesiaceae, Viscaceae and Ximeniaceae, excluding Balanophoraceae and the three basal-branching lineages Coulaceae, Strombosiaceae, Erythropalaceae): 148/2111[8]	Bisexual or unisexual[8]	Hermaphrodite, monoecious or dioecious (or other sexual systems)[8], 419	None	Nectar-foraging birds, moths, bees, wasps, flies[58, 68]; two fly-pollinated species of *Arceuthobium* also may receive wind-carried conspecific pollen[68]	Birds, wind, explosion[10,56]	Perennial, some annual in *Thesium*[4]	Coulaceae 2–3/2–4[8]	Bisexual[8]	Hermaphrodite[8]	?

Notes: for sources on species and genus numbers see *Materials and Methods*. [1] Beehler, 1994. [2] Schaefer and Renner, 2011. [3] Encyclopædia Britannica, 2012. [4] Shumei and Murata, 2003. [5] Goto et al., 2012. [6] Borchsenius and Olesen, 1990. [7] Watson and Dallwitz, 1992. [8] Nickrent et al., 2010. [9] Allen and Hiscock, 2008. [10] Nickrent, 2011. [11] Johnson et al., 2011. [12] Nickrent, 1997 onward. [13] Nelson, 2008. [14] Wikipedia, 2013. [15] Rohwer, 1993. [16] Nickrent et al., 2005. [17] García et al., 2004. [18] de Vega et al., 2009. [19] Garcia-Franco and Rico-Gray, 1997. [20] Nickrent, 2007. [21] de Figueiredo et al., 2008. [22] Maass and Musselman, 2004. [23] Tennakoon et al., 2007. [24] Nickrent et al., 2002. [25] Nair and Narayanan, 1962. [26] Sakai, 2002. [27] Jepson Flora Projects, 2012. [28] Hannon Williams and Finke, 2011. [29] Wang et al., 2009. [30] APG, 2003. [31] Gottschling, M., et al., University of Munich, personal communication. [32] Bouman and Meijer, 1994. [33] Hardy and Cook, 2012. [34] Encyclopædia Britannica, 2012. [35] Smets and Pyck, 2003. [36] Barcelona et al., 2011. [37] Banziger and Hansen, 2000. [38] Davis et al., 2007. [39] Pélabon et al., 2012. [40] Narbona et al., 2005. [41] Pahlevani and Akhani, 2011. [42] Wright, 2011. [43] Costea, 2007. [44] Kubitzki and Kurz, 1984. [45] Radcliffe-Smith, 2001. [46] Bennett and Mathews, 2006. [47] Yoder, 1998. [48] Jacobsohn et al., 1987. [49] Wicke, 2012. [50] Frost et al., 2003. [51] Sakai et al., 1995. [52] Stefanović and Olmstead, 2004. [53] Yatskievych and Mason, 1986. [54] Funk et al., 2007. [55] Alvarado-Cárdenas, 2009. [56] Kuijt, J., University of Victoria, British Columbia, Canada, personal communication. [57] Stevens, 2001 onward. [58] Heide-Jorgensen, 2008. [59] Beaman et al., 1988. [60] Wolfe et al., 1992. [61] McNeal et al., 2007. [62] Yatskievych, 1985. [63] Yatskievych, G., Missouri Botanical Garden, personal communication. [64] Schlumpberger, B., Herrenhäuser Garden, Hannover, Germany, personal communication. [65] Ecroyd, 1996. [66] Govindappa and Shivamurthy, 1975. [67] de Vega et al., 2011. [68] Player, 1979.

whether parasitic angiosperm lineages exhibit shared traits in their pollination biology, dispersal modes, or sexual systems. We also describe the pollination and mating system of two species of Apodanthaceae. Apodanthaceae comprise 18 species in two genera, *Apodanthes* and *Pilostyles* (a third genus name, *Berlinianche*, has never been validly published). They occur in North and South America, Africa, Iran, and Australia. Like most other Cucurbitales, Apodanthaceae have unisexual flowers. The flowers and fruits measure 4–10 mm in diameter (Blarer et al., 2004; present paper) and appear on the hosts' stems and branches once a year (Fig. 2). Insect pollination had been inferred from flower morphology and color, and occasional observations of visitors in Africa, Australia, and Brazil (Dell and Burbidge, 1981; Blarer et al., 2004; Brasil, 2010). Blarer et al. (2004) detected no flower odor in the African *Pilostyles* (*Berlinianche*) *aethiopica* Welwitsch, while Blassingame (1968) perceived an indoloid odor in the Californian *P. thurberi* Gray.

Questions about reproductive trait correlations across angiosperms are difficult to answer because of insufficient data; however, any strong trend, such as lack of wind pollination or preponderance of bisexual flowers, can be picked up even from incomplete data.

MATERIALS AND METHODS

Fieldwork on* Pilostyles aethiopica *and* Pilostyles haussknechtii—Pilostyles aethiopica* Welwitsch was studied between 16 and 29 February 2012 in the Mukuvisi woodlands, at 17°50′44.16″S and 31°05'24.12″E in Harare, Zimbabwe. These woodlands are protected. The parasite flowers were borne on the branches of 1.5 to 4 m tall trees of *Julbernardia globiflora* (Benth.) Troupin (Fabaceae). *Pilostyles haussknechtii* Boissier was studied between 28 May and 11 June 2012 in two populations in Iran, one north of Teheran in the province of Alborz (35°55′14.09″N and 51°03′12.78″E), the other in the province of Isfahan (32°55′11.3″N and 50°41′23.1″E). The local vegetation type was steppe (Fig. 2), and the hosts were 0.20 to 1 m tall shrubs of *Astragalus floccosus* Boissier near Teheran and of *A. verus* Olivier in Isfahan. Herbarium vouchers of the parasites and their hosts have been deposited in the herbarium of Munich, M (*S. Bellot 5* to *S. Bellot 56*).

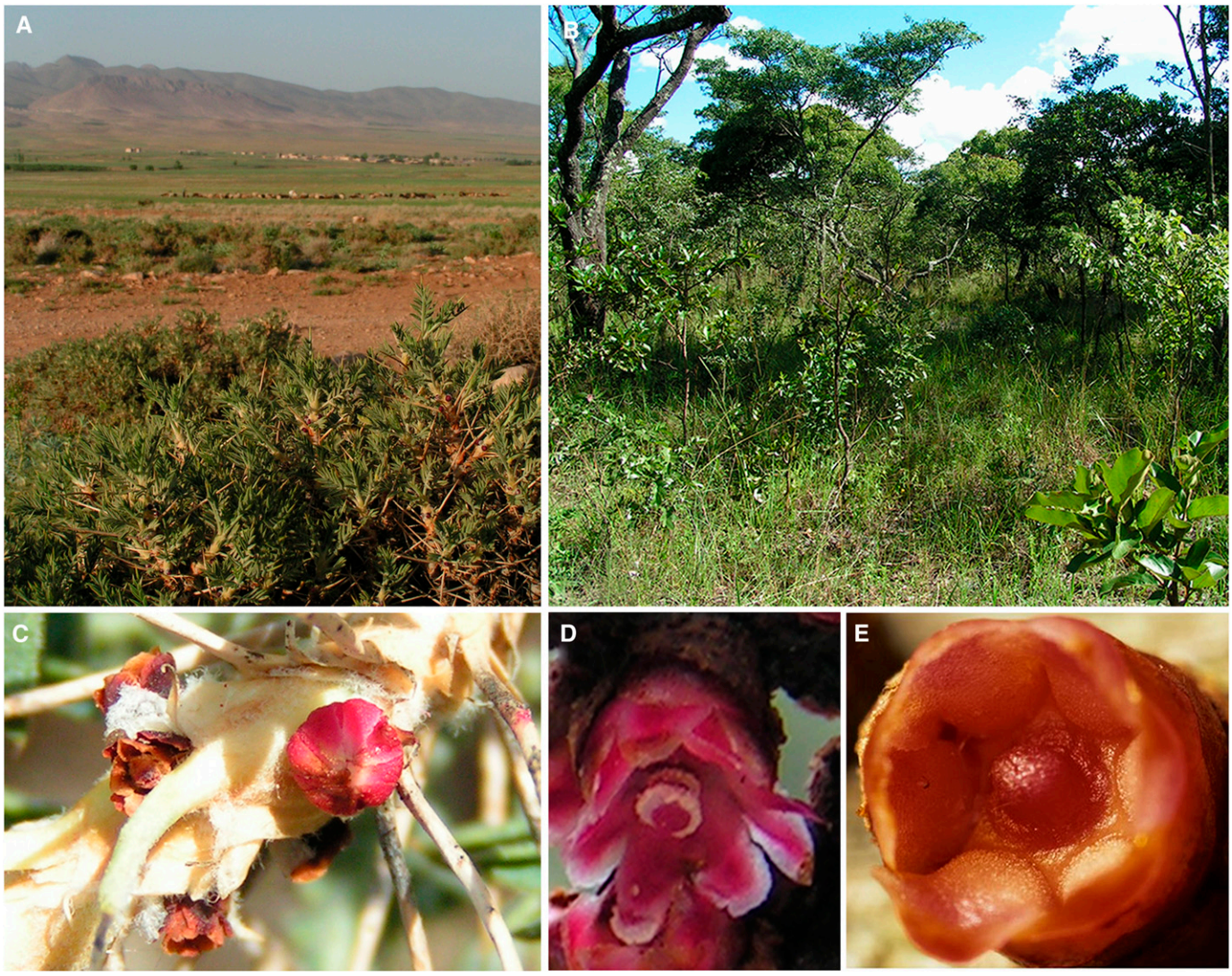

Fig. 2. Habits and floral morphology of two species of *Pilostyles*. (A) *Pilostyles haussknechtii* on *Astragalus floccosus* near Isfahan, Iran. (B) *Pilostyles aethiopica* on *Julbernardia globiflora* in the Mukuvisi woodlands, Harare, Zimbabwe. (C) Possible monoecy, with male (left) and female (right) flowers of *P. haussknechtii* on a stem of *A. floccosus*. (D) Male flower of *P. aethiopica*. (E) Female flower of *P. aethiopica* (courtesy of D. Plowes).

At the Harare site, 20 host trees carrying numerous parasite flowers were selected. Male *Pilostyles aethiopica* flowers have a shiny, white hair crown at the top of their central column, whereas female flowers do not (Fig. 2), making it possible to sex flowers by eye. At the Isfahan site, 27 infected shrubs were studied, and at the Teheran site, 26 infected and 123 uninfected shrubs. To sex the flowers of *P. haussknechtii* required close observation or prying open flowers, and it was therefore only possible to sex a subset of the flowers on each host. At the Teheran site, we also studied whether host flowering was more or less abundant in infected vs. noninfected hosts.

In Harare, parasite flowers in five host tree patches were observed during 10-min periods between 0900 and 1700 hours on several days for a total of 36 observation periods. Notes were taken on flower sex, scent, number and kind of insect visitor, and visitor behavior. Visitors were also photographed, and captured visitors have been deposited in the zoological collections of Munich (Zoologische Staatsammlung München; vouchers *S. Bellot 1* to *6*). Photographs and specimens were studied and identified by the fly and ant specialists listed in the acknowledgments. For the six most common visitors to both flower sexes, visitation frequency was quantified. The scent of male and female flowers was assayed by letting it accumulate in closed small glass vials. Nectar presence was tested with diabetes testing strips that were inserted into flowers.

Assessment of reproductive traits in parasitic angiosperms and their sister groups—All angiosperm genera recorded as parasitic were databased; following previous reviews, we included Santalales lineages thought to be parasitic based on their phylogenetic position but lacking information on haustoria formation. Most important sources were D. L. Nickrent's Parasitic Plant Connection website (http://www.parasiticplants.siu.edu/UsingPPC.html, last accessed on 28 October 2012) and R. Walker's Parasitic plants database (http://www.omnisterra. com/bot/pp_home.cgi, last accessed 28 October 2012). For each genus, species names were downloaded from Tropicos (http://www.tropicos.org), and synonyms or illegitimate names were deleted. The Plant List (http://www.theplantlist.org/) and the Angiosperm Phylogeny Website (Stevens, 2001) were also consulted. This resulted in a table of 5058 accepted species names (Appendix S1; see Supplemental Data with the online version of this article); for the largest groups, *Cuscuta*, Orobanchaceae, and Santalales, we accepted species numbers of 200 (Digital Atlas of *Cuscuta*, Costea [2007]), 2040 (Stevens, 2001), and 2111 (Nickrent et al., 2010), slightly higher (*Cuscuta*) or lower than those obtained from other sources.

Traits scored for parasitic species (Appendix S1) were flower morphology (whether unisexual or bisexual), population sexual system(s), longevity (annual, perennial), pollinator(s), and seed and fruit dispersal (partly assessed only from fruit types). Data sources (Appendix S1) included online floras, websites, articles in journals, and books of parasitic plants, most importantly Kuijt (1969) and Heide-Jorgensen (2008). The scored data were used to construct a summary table at the parasitic family level or other relevant taxonomic levels (Results).

RESULTS

Sexual systems and flowering of Apodanthaceae—Table 2 summarizes the results of the distribution of flower morphs in the Zimbabwean and Iranian populations. In Zimbabwe, 12 hosts carried only male parasite flowers, seven only female flowers, and one flowers of both sexes. In the two Iranian populations, 19 or 17 host shrubs carried only male parasite flowers, while six or 10 carried only female parasite flowers. One shrub carried parasite flowers of both sexes (Fig. 2C). At the Teheran site, of 123 uninfected hosts almost all flowered, while of the 26 infected hosts, only seven flowered. When the parasite was flowering on a flowering host, flowers of host and parasite never occurred on the same host stem.

Flowers visitors and floral scent—At the Harare site, four fly species (Fig. 3A–D) and two ant species (Fig. 3E, F) were observed on both flower sexes and during more than one observation period. The ant species *Camponotus flavomarginatus* Mayr, 1862 and *Lepisiota capensis* Mayr, 1862 (both Formicinae; vouchers *S. Bellot 4*, *5*, *6*) are unlikely pollinators because they return to the nest between forays to different food sources. In Harare, the most common flies were *Chrysomya chloropyga* Wiedemann, 1818 (voucher *S. Bellot 1*; Fig. 3B) and *Stomorhina* sp. (not captured; Fig. 3A), both belonging to the Calliphoridae. In Iran, the fly *Timia* (*Empyelocera*) spec. aff. *abstersa* (Ulidiidae; voucher *S. Bellot 7*; Fig. 3G) was common on flowers of both sexes. At both sites, flies alighted on branches near flowers (which were too small for them to land on), approached the flowers, and then probed the stamens or the base of the floral column with their proboscis, apparently looking for pollen grains or liquid. Their entire heads, including the upper part of the proboscis, came in contact with the anther ring or stigma. In Harare, floral scent could be perceived throughout the day, with no apparent change in intensity; female flowers produced scent before they opened. In Iran, flower scent was not always perceptible and was weak, even after accumulation in glass vials (done at ca. 1000 and 1500 hours). No difference in scent between male and female flowers could be perceived in either species. Nectar presence could not be verified with the diabetes testing strips because nectar quantities were too minute. No insect eggs or larvae were found in any flower.

Reproductive traits of parasitic angiosperms and their sister groups—Of a total of 4510 parasitic species, about 10% are dioecious (Table 1 and Appendix S1). Dioecy is found in Apodanthaceae, Balanophoraceae, higher Santalales, Cytinaceae, and Rafflesiaceae, and in the sister groups of three of them (Table 1). Of the eight parasitic lineages lacking dioecious species, five have sister groups that include dioecious species (Table 1). Table 1 also shows that all parasitic lineages contain perennial species, including trees (Fig. 4E), while five contain perennial as well as annual species. Only *Cuscuta* is predominantly annual. All dioecious parasite species appear to be perennial.

With the exception of the Viscaceae genus *Arceuthobium*, in which nectar-offering and fly-pollinated species may sometimes receive wind-carried conspecific pollen (Player, 1979), parasitic species are animal-pollinated, with all pollinator categories from vertebrates to insects represented (Table 1). Figure 4 shows some of the pollination syndromes represented (Fig. 4C: bee/bumblebee pollination, Fig. 4D: animal pollination with trapping). Among parasites' sister groups, several contain wind-pollinated species (Table 1). Fruit or seed dispersal by animals predominates in the parasite lineages (Fig. 4A, B), except for Cynomoriaceae, Lennoaceae, and some Orobanchaceae, the

TABLE 2. Host distribution of male and female flowers of *Pilostyles aethiopica* and *P. haussknechtii*

Species	Population	Infected hosts	Hosts with only male parasite flowers	Hosts with only female parasite flowers	Hosts with male and female parasite flowers
P. aethiopica	Mukuvisi woodlands	20	12	7	1
P. haussknechtii	Teheran	26	19	6	1
	Isfahan	27	17	10	0

Fig. 3. Flower visitors of (A–F) *Pilostyles aethiopica* in Zimbabwe and (G) *P. haussknechtii* in Iran. (A) *Stomorhina* sp. (Calliphoridae). (B) *Chrysomya chloropyga* (Calliphoridae). (C, D) Fly flower visitors. (E) *Lepisiota capensis* (Formicidae). (F) *Camponotus flavomarginatus* (Formicidae). (G) *Timia* (*Empyelocera*) sp. aff. *abstersa* (Ulidiidae). The flower diameter is 4–6 mm.

seeds of which are dispersed by rain splash or wind (cf. references in Table 1).

DISCUSSION

Sexual systems and reproductive biology of Apodanthaceae— Flowers of *Pilostyles aethiopica* and *P. haussknechtii* were strictly unisexual, without traces of aborted organs of the other sex. Unisexual flowers are the ancestral condition in the Cucurbitales, apparently combined with a dioecious mating system (Zhang et al., 2006). A study of the Brazilian species *P. ingae* (Karsten) Hooker recently reported that 1.3% of the male flowers had small pistils, but were functionally male (Brasil, 2010: The numbers of dissected flowers were not given.). In the species studied here, *P. aethiopica* and *P. haussknechtii*, dioecy appears to be the normal sexual system (Table 1). Among the remaining Apodanthaceae, three or four are monoecious, namely the Australian *P. collina* B. Dell and *P. coccoidea* K.R.Thiele (Thiele et al., 2008), the Brazilian *Apodanthes minarum* Vattimo (de Vattimo, 1973), and possibly the Brazilian *P. stawiarskii* de Vatt. (de Vattimo, 1950). Most remaining species

Fig. 4. Examples of reproductive traits of parasitic angiosperms. (A) Seedlings of the hemiparasite *Krameria* sp. (Krameriaceae) at the foot of a host (Cactaceae) with barbed seeds (inset) adapted to exozoochory. (B) Bisexual flowers and fruits of the bird-pollinated and dispersed *Tristerix longibracteatus* (Loranthaceae). (C) Herb habit of the root parasite *Orobanche lavandulacea* (Orobanchaceae) and close-up of its bee-pollinated inflorescence (inset). (D) Bisexual flower of the root holoparasite *Prosopanche americana* (Hydnoraceae) in male phase (courtesy of B. Schlumpberger). (E) Tree habit and inflorescence (inset) of the hemiparasite *Gaiadendron* (Loranthaceae). Images A–C, E: courtesy of G. Gerlach.

are dioecious [*P. ingae* and *P. thurberi* (Brasil, 2010); the Australian *P. hamiltonii* Gardner (Thiele et al., 2008), the American *P. mexicana* (Brandegee) Rose, *P. blanchetii* (Gardner) Brown, *P. calliandrae* (Gardner) Brown (de Vattimo, 1971), *A. tribracteata* Rusby (Rusby, 1920), and *A. caseariae* Poiteau (Poiteau, 1824)]. The sexual system of the American species *P. galactiae* Ule, *P. goyazensis* Ule, and *A. berteroi* (Guill.) Gardner, and of the African *P. holtzii* Engler is unknown, the later species being known only from a type destroyed during World War II.

The endoparasitic lifestyle of Apodanthaceae makes it difficult to decide whether cases of monoecy in fact represent multiple infections with two or more parasite individuals, each representing one sex. Next-generation sequencing data that will facilitate the design of highly variable genetic markers have now been generated for *Pilostyles hamiltonii* and *P. aethiopica* and may resolve this question (S. Bellot and S. S. Renner, unpublished data). Nothing is known about germination and initial haustorium formation, but animal dispersal is certain because the fruits are small red, orange, or yellow berries with numerous tiny (0.5 mm) seeds, embedded in a fleshy sticky pulp (Bouman and Meijer, 1994; S. Bellot, personal observation). Birds have been observed eating the fruits of *Pilostyles ingae* in Brazil (Brasil, 2010).

Apodanthaceae flowers are sessile, minute, and dark-red or yellow (depending on the species), with nectar offered in minute quantities (this study) on a shallow nectary cushion at the base of the flowers. Pollination by short-tongued nectar-foragers was therefore expected, and as demonstrated here, at least *Pilostyles aethiopica* is indeed pollinated by short-tongued flies. An earlier study of this species also reported flies as flower visitors, namely *Drosophila*, Sciaridae, Psychodidae, and Cecidomyiidae (Blarer et al., 2004), but provided neither details on their foraging behavior on the flowers nor whether they visited both sexes.

The flower scent likely consists of ethanol and related compounds in addition to fruity compounds as known from other fly-pollinated flowers, including Araceae (Kite and Hetterscheid, 1997), Schisandraceae (Yuan et al., 2008), *Trollius* (Ibanez et al., 2010), but also wasp-pollinated flowers (Johnson et al., 2009: *Drypetes natalensis*). Females of Calliphoridae (blow-flies), that also visit carrion, are attracted by such scents (e.g., Jürgens et al., 2006). No insect eggs or larvae have been reported from Apodanthaceae flowers, nor were any found in our study, and the flies therefore must come for pollen or nectar. In an Australian species of *Pilostyles*, a Brazilian species of *Pilostyles*, and the Panamanian *Apodanthes caseariae*, wasps have been observed on the flowers (Dell and Burbidge, 1981; Brasil, 2010), with the Panamanian wasp being identified as *Pepsis menechma* Lepeletier, 1845 (C. Galdames, Smithsonian Tropical Research Institute, Panama; personal communication).

Is dioecy overrepresented in parasitic angiosperms?—Dioecy has evolved in five of the 13 parasitic angiosperm clades, amounting to 456 species or 10% of 4510 parasitic species (Table 1; Appendix S1). Among nonparasitic angiosperms, 14 164, or 6%, are dioecious, assuming a total of 240 000 angiosperm species (Renner and Ricklefs, 1995). These estimates have unknown error margins. A phylogenetically informed analysis of a possible correlation between dioecy and parasitism would require a phylogeny for the ca. 14 000 genera of angiosperms with presence/absence of dioecy and parasitism coded for each genus. Neither the phylogeny nor the sexual system data are so far available. If dioecy is indeed an evolutionary handicap that reduces species persistence as suggested by Heilbuth et al. (2001; also Vamosi and Vamosi, 2004), it might contribute to the low species proliferation seen in parasitic angiosperms (Hardy and Cook, 2012). However, one would then expect dioecious parasite lineages to be more species poor than the strictly bisexual ones, which is not the case (Table 1). The situation may be complicated by additional correlations, for example, between host specificity and dioecy, presumably both requiring highly efficient dispersal mechanisms to generate the dense seed shadow required to maintain viable populations (Wilson and Harder, 2003).

Dioecy is less likely in families of herbs and more likely in families containing shrubs and trees (Renner and Ricklefs, 1995), and we therefore looked at the longevity of parasitic plants with or without dioecious species. All 13 parasitic lineages contain perennial species, while four or five also contain annual species (Table 1). None of the annual species is dioecious (the longevity of Balanophoraceae is still unknown), matching the finding of Renner and Ricklefs (1995). A perennial life cycle may be advantageous for parasites, especially when hosts are difficult to find (i.e., when a huge seed shadow is necessary for successful establishment). Testing this would require phylogenetic information at the genus level for parasites and other angiosperms.

Animal pollination is overrepresented in parasitic angiosperms—As shown in Table 1, parasitic angiosperms are animal-pollinated, while among the remaining angiosperms at least 24 000 species are wind-pollinated, with wind pollination having evolved at least 65 times (Linder, 1998). Ecological reasons for the absence of wind pollination among parasites may be their occurrence at ground level (root parasites), on the stems of trees (e.g., Apodanthaceae), or within leafy tree crowns (e.g., mistletoes), where pollen export and import by wind might be relatively difficult. All animal pollinators are represented among parasites, ranging from ants, via birds, to bees, butterflies, beetles, and flies, to shrews (Figs. 3, 4; Table 1 and its references). Among seed dispersal modes, dispersal by animals predominates, although not as pervasively as animal pollination.

Conclusions—In conclusion, the absence of wind pollination is clearly a feature uniting parasitic angiosperms (not noted by Molau, 1995). Dioecy appears to be more common among parasites than it is among other flowering plants, but phylogenetic resolution so far is insufficient to infer shifts to and from dioecy. The African and Iranian species of Apodanthaceae studied here are dioecious and pollinated by flies. Resolving whether a combination of animal pollination, perennial life history, and dioecy facilitated the evolution of parasitism or evolved subsequent to the evolution of parasitism will require ancestral trait reconstructions in several of the 13 lineages. At the moment, we are far from understanding the ecological and evolutionary interactions among hosts, parasites, their pollinators and dispersers even in the best-studied systems, *Cuscuta*, *Viscum*, or the economically important pests in Orobanchaceae, and compared to earlier reviews (Kuijt, 1969; Molau, 1995), ecological data have been slower to accumulate than molecular ones.

LITERATURE CITED

ALLEN, A. M., AND S. J. HISCOCK. 2008. Evolution and phylogeny of self-incompatibility systems in angiosperms. *In* V. E. Franklin-Tong [ed.], Self-incompatibility in flowering plants—Evolution, diversity, and mechanisms, 73–101. Springer, Heidelberg, Germany.

ALVARADO-CÁRDENAS, L. O. 2009. Systematics of the genus *Bdallophytum* (Cytinaceae). *Acta Botanica Mexicana* 87: 1–21.

ANGIOSPERM PHYLOGENY GROUP. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141: 399–436.

BANZIGER, H., AND B. HANSEN. 2000. A new taxonomic revision of a deceptive flower, *Rhizanthes* Dumortier (Rafflesiaceae). *Natural History Bulletin of the Siam Society* 48: 117–143.

BARCELONA, J. F., E. S. FERNANDO, D. L. NICKRENT, D. S. BALETE, AND P. B. PELSER. 2011. An amended description of *Rafflesia leonardi* and a revised key to Philippine *Rafflesia* (Rafflesiaceae). *Phytotaxa* 24: 11–18.

BARKMAN, T. J., J. R. MCNEAL, S.-H. LIM, G. COAT, H. B. CROOM, N. D. YOUNG, AND C. W. DEPAMPHILIS. 2007. Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. *BMC Evolutionary Biology* 7: 248.

BEAMAN, S., P. J. DECKER, AND J. H. BEAMAN. 1988. Pollination of *Rafflesia. American Journal of Botany* 75: 1148–1162.

BEEHLER, B. M. 1994. Canopy-dwelling honeyeater aggressively defends terrestrial nectar source. *Biotropica* 26: 459–461.

BENNETT, J. R., AND S. MATHEWS. 2006. Phylogeny of the parasitic plant family Orobanchaceae inferred from phytochrome A1. *American Journal of Botany* 93: 1039–1051.

BLARER, A., D. L. NICKRENT, AND P. K. ENDRESS. 2004. Comparative floral structure and systematics in Apodanthaceae (Rafflesiales). *Plant Systematics and Evolution* 245: 119–142.

BLASSINGAME, J. L. 1968. Revision of the genus *Pilostyles* in the southwestern United States. Ph.D. dissertation, Texas Technological College, Texas, USA.

BORCHSENIUS, F., AND J. M. OLESEN. 1990. The Amazonian root holoparasite *Lophophytum mirabile* (Balanophoraceae) and its pollinators and herbivores. *Journal of Tropical Ecology* 6: 501–505.

BOUMAN, F., AND W. MEIJER. 1994. Comparative structure of ovules and seeds in Rafflesiaceae. *Plant Systematics and Evolution* 193: 187–212.

BRASIL, B. DE AZEVEDO. 2010. Life cycle, phenology and floral anatomy of *Pilostyles* (Apodanthaceae-Rafflesiaceae s.l.): Phylogenetic position allowances for Apodanthaceae family. Abstract of M.Sc. thesis, Instituto de Biociências, Sao Paulo, Brazil.

CONN, J. S., T. R. WENTWORTH, AND U. BLUM. 1980. Patterns of dioecism in the flora of the Carolinas. *American Midland Naturalist* 103: 310–315.

COSTEA, M. 2007 [onward]. Digital atlas of *Cuscuta* (Convolvulaceae). Wilfrid Laurier University Herbarium, Ontario, Canada. Website http://www.wlu.ca/page.php?grp_id=2147&p=8968 [accessed 31 October 2012].

CRONQUIST, A. 1988. The evolution and classification of flowering plants. New York Botanical Garden Press, Bronx, New York, USA.

DAVIS, C. C., M. LATVIS, D. L. NICKRENT, K. J. WURDACK, AND D. A. BAUM. 2007. Floral gigantism in Rafflesiaceae. *Science* 315: 1812.

DE FIGUEIREDO, R. A., A. A. DE OLIVEIRA, M. A. ZACHARIAS, S. M. BARBOSA, F. F. PEREIRA, G. N. CAZELA, J. P. VIANA, AND R. A. DE CAMARGO. 2008. Reproductive ecology of the exotic tree *Muntingia calabura* L. (Muntigiaceae) in Southeastern Brazil. *Revista Árvore* 32: 993–999.

DE VATTIMO, I. 1950. Pilostyles stawiarskii. *Revista Brasileira de Biologia* 10: 196.

DE VATTIMO, I. 1971. Contribuçoa ao conhecimento da tribo Apodanthea R. Br. Parte I—Conspecto das espécies (Rafflesiaceae). *Rodriguésia* 38: 37–62.

DE VATTIMO, I. 1973. Notas sobre o genero *Apodanthes* Poit. com descriçao de duas novas especies (Rafflesiaceae). *Revista Brasileira de Biologia* 33: 140.

DE VEGA, C., M. ARISTA, P. L. ORTIZ, C. M. HERRERA, AND S. TALAVERA. 2009. The ant-pollination system of *Cytinus hypocist*is (Cytinaceae), a Mediterranean root holoparasite. *Annals of Botany* 103: 1065–1075.

DE VEGA, C., M. ARISTA, P. L. ORTIZ, C. M. HERRERA, AND S. TALAVERA. 2011. Endozoochory by beetles: A novel seed dispersal mechanism. *Annals of Botany* 107: 629–637.

DELL, B., AND A. H. BURBIDGE. 1981. Notes on the biology of *Pilostyles* (Rafflesiaceae) in Western Australia. *Western Australian Herbarium Research Notes* 5: 71–79.

ECROYD, C. E. 1996. The ecology of *Dactylanthus taylorii* and threats to its survival. *New Zealand Journal of Ecology* 20: 81–100.

ENCYCLOPÆDIA BRITANNICA, INC. 2012. Cucurbitales. Encyclopædia Britannica Online. Website http://www.britannica.com/EBchecked/topic/1432268/Cucurbitales [accessed 1 October 2012].

FILIPOWICZ, N., AND S. S. RENNER. 2010. The worldwide holoparasitic Apodanthaceae confidently placed in the Cucurbitales by nuclear and mitochondrial gene trees. *BMC Evolutionary Biology* 10: 219.

FROST, A., J. C. LOPEZ-GUTIERREZ, AND C. B. PURRINGTON. 2003. Fitness of *Cuscuta salina* (Convolvulaceae) parasitizing *Beta vulgaris* (Chenopodiaceae) grown under different salinity regimes. *American Journal of Botany* 90: 1032–1037.

FUNK, H. T., S. BERG, K. KRUPINSKA, U. G. MAIER, AND K. KRAUSE. 2007. Complete DNA sequences of the plastid genomes of two parasitic flowering plant species, *Cuscuta reflexa* and *Cuscuta gronovii. BMC Plant Biology* 7: 45.

GARCÍA, M. A., E. H. NICHOLSON, AND D. L. NICKRENT. 2004. Extensive intraindividual variation in plastid rDNA sequences from the holoparasite *Cynomorium coccineum* (Cynomoriaceae). *Journal of Molecular Evolution* 58: 322–332.

GARCIA-FRANCO, J. G., AND V. RICO-GRAY. 1997. Reproductive biology of the holoparasitic endophyte *Bdallophyton bambusarum* (Rafflesiaceae). *Botanical Journal of the Linnean Society* 123: 237–247.

GOTO, R., G. YAMAKOSHI, AND T. MATSUZAWA. 2012. A novel brood-site pollination mutualism?: The root holoparasite *Thonningia sanguinea* (Balanophoraceae) and an inflorescence-feeding fly in the tropical rainforests of West Africa. *Plant Species Biology* 27: 164–169.

GOVINDAPPA, D. A., AND G. R. SHIVAMURTHY. 1975. The pollination mechanism in *Balanophora abbreviata* Blume. *Annals of Botany* 39: 977–978.

HANNON WILLIAMS, L. E., AND D. L. FINKE. 2011. Breeding system and potential pollinators of the desert shrub *Krameria erecta* (Krameriaceae). *In* Proceedings of the 96th Ecological Society of America Annual Meeting, 2011, Austin, Texas, USA [abstract]. Website http://eco.confex.com/eco/2011/preliminaryprogram/abstract_30622.htm [accessed 1 October 2012].

HARDY, N. B., AND L. G. COOK. 2012. Testing for ecological limitation of diversification: A case study using parasitic plants. *American Naturalist* 180: 438–449.

HEIDE-JORGENSEN, H. S. 2008. Parasitic flowering plants. Brill, Leiden, Netherlands.

HEILBUTH, J. C., K. L. ILVES, AND S. P. OTTO. 2001. The consequences of dioecy for seed dispersal: Modeling the seed-shadow handicap. *Evolution; International Journal of Organic Evolution* 55: 880–888.

IBANEZ, S., S. DÖTTERL, M. C. ANSTETT, S. BAUDINO, J. C. CAISSARD, C. GALLET, AND L. DESPRÉS. 2010. The role of volatile organic compounds, morphology, and pigments of globeflowers in the attraction of their specific pollinating flies. *The New Phytologist* 188: 451–463.

JACOBSOHN, R., D. BEN-GHEDALIA, AND K. MARTON. 1987. Effect of the animal's digestive system on the infectivity of *Orobanche* seeds. *Weed Research* 27: 87–90.

JEPSON FLORA PROJECTS [eds.]. 2012. Jepson eFlora, Index to treatments and keys from the Jepson Manual. Website http://ucjeps.berkeley.edu/interchange/I_treat_indexes.html [last accessed on 1 October 2012].

JIAN, S., P. S. SOLTIS, M. A. GITZENDANNER, M. J. MOORE, R. LI, T. A. HENDRY, Y. L. QIU, A. DHINGRA, C. D. BELL, AND D. E. SOLTIS. 2008. Resolving an ancient, rapid radiation in Saxifragales. *Systematic Biology* 57: 38–57.

JOHNSON, S. D., P. M. BURGOYNE, L. D. HARDER, AND S. DÖTTERL. 2011. Mammal pollinators lured by the scent of a parasitic plant. *Proceedings. Biological Sciences* 278: 2303–2310.

JOHNSON, S. D., M. E. GRIFFITHS, C. I. PETER, AND M. J. LAWES. 2009. Pollinators, "mustard oil" volatiles, and fruit production in flowers of the dioecious tree *Drypetes natalensis* (Putranjivaceae). *American Journal of Botany* 96: 2080–2086.

JÜRGENS, A., S. DÖTTERL, AND U. MEVE. 2006. The chemical nature of fetid floral scents in stapeliads (Apocynaceae-Asclepoadoideae-Ceropegieae). *The New Phytologist* 172: 452–468.

KITE, G. C., AND W. L. A. HETTERSCHEID. 1997. Inflorescence odours of *Amorphophallus* and *Pseudodracontium* (Araceae). *Phytochemistry* 46: 71–75.

KRAUSE, K. 2012. Plastid genomes of parasitic plants: A trail of reductions and losses. *In* C. E. Bullerwell [ed.], Organelle genetics, 79–103. Springer, Heidelberg, Germany.

KUBITZKI, K., AND H. KURZ. 1984. Synchronized dichogamy and dioecy in neotropical Lauraceae. *Plant Systematics and Evolution* 147: 253–266.

KUIJT, J. 1969. The biology of parasitic flowering plants. University of California Press, Berkeley, California, USA.

LINDER, H. P. 1998. Morphology and the evolution of wind pollination. *In* S. J. Owens and P. J. Rudall [eds.], Reproductive biology in systematics, conservation and economic botany, 123–135. Royal Botanic Gardens, Kew, UK.

MAASS, E., AND L. J. MUSSELMAN. 2004. *Hydnora triceps* (Hydnoraceae)—First record in Namibia and first description of fruits. *Dinteria* 29: 1–10.

MCNEAL, J. R., J. V. KUEHL, J. L. BOORE, AND C. W. DEPAMPHILIS. 2007. Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. *BMC Plant Biology* 7: 57.

MOLAU, U. 1995. Reproductive ecology and biology. *In* M. C. Press and J. D. Graves [eds.], Parasitic plants, 141–176. Chapman and Hall, London, UK.

NAIR, N. C., AND K. R. NARAYANAN. 1962. Studies on the Aristolochiaceae: I. Nodal and floral anatomy. *Proceedings of the National Institute of Sciences of India* 28B: 211–227.

NARBONA, E., M. ARISTA, AND P. L. ORTIZ. 2005. Explosive seed dispersal in two perennial Mediterranean *Euphorbia* species (Euphorbiaceae). *American Journal of Botany* 92: 510–516.

NELSON, S. C. 2008. *Cassytha filiformis*. Plant Disease, no. 42. University of Hawaii at Manoa, College of Tropical Agriculture and Human Resources, Cooperative Extension Service, Honolulu, Hawaii, USA.

NEUHAUS, H. E., AND M. J. EMES. 2000. Nonphotosynthetic metabolism in plastids. *Annual Review of Plant Physiology and Plant Molecular Biology* 51: 111–140.

NICKRENT, D. L. 1997 [onward]. The parasitic plant connection. Website http://www.parasiticplants.siu.edu/UsingPPC.html [last accessed on 1 October 2012].

NICKRENT, D. L. 2007. Cytinaceae are sister to Muntingiaceae (Malvales). *Taxon* 56: 1129–1135.

NICKRENT, D. L. 2011. Santalales (Including Mistletoes). *In* Encyclopedia of life sciences. John Wiley, Chichester, UK.

NICKRENT, D. L., A. BLARER, Y.-L. QIU, D. E. SOLTIS, P. S. SOLTIS, AND M. ZANIS. 2002. Molecular data place Hydnoraceae with Aristolochiaceae. *American Journal of Botany* 89: 1809–1817.

NICKRENT, D. L., J. P. DER, AND F. E. ANDERSON. 2005. Discovery of the photosynthetic relatives of the "Maltese mushroom" *Cynomorium*. *BMC Evolutionary Biology* 5: 38.

NICKRENT, D. L., V. MALÉCOT, R. VIDAL-RUSSELL, AND J. P. DER. 2010. A revised classification of Santalales. *Taxon* 59: 538–558.

PAHLEVANI, A. H., AND H. AKHANI. 2011. Seed morphology of Iranian annual species of *Euphorbia* (Euphorbiaceae). *Botanical Journal of the Linnean Society* 167: 212–234.

PÉLABON, C., P. THÖNE, T. F. HANSEN, AND W. S. ARMBRUSTER. 2012. Signal honesty and cost of pollinator rewards in *Dalechampia scandens* (Euphorbiaceae). *Annals of Botany* 109: 1331–1339.

PLAYER, G. 1979. Pollination and wind dispersal of pollen in *Arceuthobium*. *Ecological Monographs* 49: 73–87.

POITEAU, M. A. 1824. Description de *l'Apodanthes*, nouveau genre de plantes phanérogames parasite. *Annales des Sciences Naturelles* 3: 422, t. 26.

QIU, Y. L., L. LI, B. WANG, J.-Y. XUE, T. A. HENDRY, R. Q. LI, J. W. BROWN, ET AL. 2010. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. 48: 391–425.

RADCLIFFE-SMITH, A. 2001. Genera Euphorbiacearum. Royal Botanical Gardens, Kew, UK.

RENNER, S. S., AND R. E. RICKLEFS. 1995. Dioecy and its correlates in the flowering plants. *American Journal of Botany* 82: 596–606.

ROHWER, J. G. 1993. Lauraceae. *In* K. Kubitzki, J. G. Rohwer, and V. Bittrich [eds.], The families and genera of vascular plants, vol. 2, 366–391. Springer, Heidelberg, Germany.

RUSBY, H. H. [ed.]. 1920. Descriptions of three hundred new species of South American plants. Published by the author, New York, New York, USA.

SAKAI, A. K., W. L. WAGNER, D. M. FERGUSON, AND D. R. HERBST. 1995. Origins of dioecy in the Hawaiian flora. *Ecology* 76: 2517–2529.

SAKAI, S. 2002. *Aristolochia* spp. (Aristolochiaceae) pollinated by flies breeding on decomposing flowers in Panama. *American Journal of Botany* 89: 527–534.

SCHAEFER, H., AND S. S. RENNER. 2011. Phylogenetic relationships in the order Cucurbitales and a new classification of the gourd family (Cucurbitaceae). *Taxon* 60: 122–138.

SHUMEI, H., AND J. MURATA. 2003. Balanophoraceae. Pp. 272-276 in Z. Y. Wu and P. Raven [eds.], Flora of China 5, Science Press, Beijing & Missouri Botanical Garden Press, St. Louis, MO, USA.

SMETS, E., AND N. PYCK. 2003. Ericales (*Rhododendron*) [online]. *In* Encyclopedia of life sciences [eLS]. John Wiley, Chichester, UK. Website http://onlinelibrary.wiley.com/doi/10.1038/npg.els.0003730/full, doi:10.1038/npg.els.0003730.

STEFANOVIĆ, S., AND R. G. OLMSTEAD. 2004. Testing the phylogenetic position of a parasitic plant (*Cuscuta*, Convolvulaceae, Asteridae): Bayesian inference and the parametric bootstrap on data drawn from three genomes. *Systematic Biology* 53: 384–399.

STEVENS, P. F. 2001 [onward]. Angiosperm Phylogeny Website, version 12, July 2012 [and more or less continuously updated since]. Website http://www.mobot.org/MOBOT/research [accessed 1 October 2012].

TENNAKOON, K. U., J. F. BOLIN, L. J. MUSSELMAN, AND E. MAASS. 2007. Structural attributes of the hypogeous holoparasite *Hydnora triceps* Drege & Meyer. *American Journal of Botany* 94: 1439–1449.

THIELE, K. R., S. J. WYLIE, L. MACCARONE, P. HOLLICK, AND J. A. MCCOMB. 2008. *Pilostyles coccoidea* (Apodanthaceae), a new species from Western Australia described from morphological and molecular evidence. *Nuytsia* 18: 273–284.

VAMOSI, J. C., AND S. M. VAMOSI. 2004. The role of diversification in causing the correlates of dioecy. *Evolution; International Journal of Organic Evolution* 58: 723–731.

VIDAL-RUSSELL, R., AND D. L. NICKRENT. 2008a. The first mistletoes: Origins of aerial parasitism in Santalales. *Molecular Phylogenetics and Evolution* 47: 523–537.

VIDAL-RUSSELL, R., AND D. L. NICKRENT. 2008b. Evolutionary relationships in the showy mistletoe family (Loranthaceae). *American Journal of Botany* 95: 1015–1029.

WANG, H., M. J. MOORE, P. S. SOLTIS, C. D. BELL, S. F. BROCKINGTON, R. ALEXANDRE, C. C. DAVIS, ET AL. 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences, USA* 106: 3853–3858.

WATSON, L., AND M. J. DALLWITZ. 1992 [onward]. The families of flowering plants: Descriptions, illustrations, identification, and information retrieval, version 19 December 2012. Website http://delta-intkey.com/angio [accessed 1 October 2012].

WICKE, S. 2012. Assessing the evolutionary patterns of plastid genome reduction in a group of non-photosynthetic parasitic angiosperms (Orobanchaceae). Ph.D. dissertation, University of Vienna, Vienna, Austria.

WIKIPEDIA. 2013. *Cassytha* [online]. Website http://en.wikipedia.org/wiki/Cassytha [accessed 25 April 2013].

WILSON, W. G., AND L. D. HARDER. 2003. Reproductive uncertainty and the relative competitiveness of simultaneous hermaphroditism versus dioecy. *American Naturalist* 162: 220–241.

WOLFE, K. H., C. W. MORDEN, AND J. D. PALMER. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proceedings of the National Academy of Sciences, USA* 89: 10648–10652.

WRIGHT, M. 2011. The evolution of sexual reproduction in *Cuscuta* (Convolvulaceae). M.Sc. thesis, Wilfrid Laurier University, Ontario, Canada.

YATSKIEVYCH, G. 1985. Notes on the biology of the Lennoaceae. *Cactus & Succulent Journal* 57: 73–79.

YATSKIEVYCH, G., AND C. T. MASON JR. 1986. A revision of the Lennoaceae. *Systematic Botany* 11: 531–548.

YODER, J. I. 1998. Self and cross-compatibility in three species of the hemiparasite *Triphysaria* (Scrophulariaceae). *Environmental and Experimental Botany* 39: 77–83.

YUAN, L. C., Y. B. LUO, L. B. THIEN, J. H. FAN, H. L. XU, J. YUKAWA, AND Z. D. CHEN. 2008. Pollination of *Kadsura longipedunculata* (Schisandraceae), a monoecious basal angiosperm, by female, pollen-eating *Megommata* sp. (Cecidomyiidae: Diptera) in China. *Biological Journal of the Linnean Society. Linnean Society of London* 93: 523–536.

ZHANG, L.-B., M. P. SIMMONS, A. KOCYAN, AND S. S. RENNER. 2006. Phylogeny of the Cucurbitales based on DNA sequences of nine loci from three genomes: Implications for morphological and sexual system evolution. *Molecular Phylogenetics and Evolution* 39: 305–322.

ZHANG, S.-D., D. E. SOLTIS, Y. YANG, D.-Z. LI, AND T.-S. YI. 2011. Multi-gene analysis provides a well-supported phylogeny of Rosales. *Molecular Phylogenetics and Evolution* 60: 21–28.

ZHANG, Z.-H., C.-Q. LI, AND J. LI. 2009. Phylogenetic placement of *Cynomorium* in Rosales inferred from sequences of the inverted repeat region of the chloroplast genome. *Journal of Systematics and Evolution* 47: 297–304.

Appendix S1. Parasitic angiosperm species scored for the following traits: flower morphology (whether unisexual or bisexual), population sexual system, longevity (annual, perennial), pollinator, seed and fruit dispersal (partly assessed from fruit morphology), host specificity, type of parasitism (exo- or endo-parasitism, root or stem parasitism), growth form (herb, shrub, tree). Appendix S1 is a excel workbook with three sheets. Sheet 1 of Appendix 1 consists in ca. 300 pages, it is available online at http://www.amjbot.org/content/100/6/1083/suppl/DC1, only the first pages are displayed here as an example. Sheet 2 summarizes the counts to estimate proportion of dioecious species in Angiosperms. Sheet 3 summarizes the references used to fill sheet 1.

Appendix S1, sheet 1 (only headers and first 33 rows)

Family	Genus	Species	Species author	Sum of accepted species listed in the Parasitic Plants Database (http://www.omnisterra.com/bot/pp_home.cgi)	Sum of accepted species listed in Tropicos (http://www.tropicos.org/)	Sum of accepted species in other sources listed in column H	References for column G	Number used to calculate presentages in sheet 2, columns F and G, with priority given to # in G, then E, then F	Sexual system at the population level. H: each plant hermaphrodite, M: monoecious, D: dioecious, O: others (see column V)	Flower sexuality (monclinous or diclinous). B: bisexual, U: unisexual, M: male, F: female	Photosynthesis capacity. He: hemiparasite, Ho: holoparasite (some He are photosynthetic only at seedling stage)	Vegetative development. En: endoparasite, Ex: exoparasite
Lauraceae	**Cassytha**			**21**	**21**			**21**	**H**	**B**	**He**	**Ex**
Lauraceae	Cassytha	aurea	Weber						H	B	He	Ex
Lauraceae	Cassytha	candida	Weber						H	B	He	Ex
Lauraceae	Cassytha	capensis	Meissn.						H	B	He	Ex
Lauraceae	Cassytha	capillaris	Meissn.						H	B	He	Ex
Lauraceae	Cassytha	ciliolata	Nees						H	B	He	Ex
Lauraceae	Cassytha	filiformis	L.						H	B	He	Ex
Lauraceae	Cassytha	flava	Nees						H	B	He	Ex
Lauraceae	Cassytha	flindersii	Weber						H	B	He	Ex
Lauraceae	Cassytha	glabella	R.Br.						H	B	He	Ex
Lauraceae	Cassytha	larsenii	Kosterm.						H	B	He	Ex
Lauraceae	Cassytha	melantha	R.Br.						H	B	He	Ex
Lauraceae	Cassytha	micrantha	Meissn.						H	B	He	Ex
Lauraceae	Cassytha	nodiflora	Meissn.						H	B	He	Ex
Lauraceae	Cassytha	paradoxae	Proctor						H	B	He	Ex
Lauraceae	Cassytha	pedicellosa	Weber						H	B	He	Ex
Lauraceae	Cassytha	peninsularis	Weber						H	B	He	Ex
Lauraceae	Cassytha	pomiformis	Nees						H	B	He	Ex
Lauraceae	Cassytha	pondoensis	Engl.						H	B	He	Ex
Lauraceae	Cassytha	pubescens	R.Br.						H	B	He	Ex
Lauraceae	Cassytha	racemosa	Nees						H	B	He	Ex
Lauraceae	Cassytha	rufa	Weber						H	B	He	Ex
Hydnoraceae	**Hydnora**			**6**	**7**			**6**	**H**	**B**	**Ho**	**Ex**
Hydnoraceae	Hydnora	abyssinica	A.Braun						H	B	Ho	Ex
Hydnoraceae	Hydnora	africana	Thunb.						H	B	Ho	Ex
Hydnoraceae	Hydnora	esculenta	Jumelle&Perrier						H	B	Ho	Ex
Hydnoraceae	Hydnora	sinandevu	Beentje & Luke						H	B	Ho	Ex
Hydnoraceae	Hydnora	triceps	Drege & E.Mey.						H	B	Ho	Ex
Hydnoraceae	Hydnora	visseri	Bolin, E.Maass & Musselman						H	B	Ho	Ex
Hydnoraceae	**Prosopanche**			**3**	**3**			**3**	**H**	**B**	**Ho**	**Ex**
Hydnoraceae	Prosopanche	americana	(R.Br.) Baill.						H	B	Ho	Ex
Hydnoraceae	Prosopanche	bonacinae	Speg.						H	B	Ho	Ex
Hydnoraceae	Prosopanche	costaricensis	Gomez & Gomez						H	B	Ho	Ex

Attachment to host. R: on the root, S: on the stem	Host specificity. S: one species, G: many species of one genus, F: many genera of one family, F+: many families, V: various (for genus level)	Growth form. H: herb, S: shrub, T: tree, C: climber or liana	Longevity. P: perennial, A: annual	Pollinator	Diaspore dispersal	Fruit morphology	References. D: http://delta-intkey.com/angio/index.htm, W: Wikipedia, Nickrent: http://www.parasiticplants.siu.edu/index.html, Wjstor: http://plants.jstor.org/, Wflorabase: http://florabase.dec.wa.gov.au/, Wville-ge: http://www.ville-ge.ch/musinfo/bd/cjb/africa/recherche.php?langue=a), numbers correspond to references in the legend of table 1 of the manuscript, a list of other references is given in sheet 3	Notes	Sexual systems to count by using the function "conditional sum": genus level and autotrophs are removed	Number 1 attributed to count sexual systems (results of the count are on sheet 2
S	**F+**	**C**	**P**		birds, mammals	drupe	**Nickrent, D, W, 58**			**1**
S	F+	C	P				Nickrent, D, W		H	1
S	F+	C	P				Nickrent, D, W		H	**1**
S	F+	C	P				Nickrent, D, W		H	1
S	F+	C	P				Nickrent, D, W		H	**1**
S	F+	C	P				Nickrent, D, W		H	1
S	F+	C	P				Nickrent, D, W		H	**1**
S	F+	C	P				Nickrent, D, W		H	1
S	F+	C	P				Nickrent, D, W		H	**1**
S	F+	C	P				Nickrent, D, W		H	1
S	F+	C	P				Nickrent, D, W		H	**1**
S	F+	C	P				Nickrent, D, W		H	1
S	F+	C	P				Nickrent, D, W		H	**1**
S	F+	C	P				Nickrent, D, W		H	1
S	F+	C	P				Nickrent, D, W		H	**1**
S	F+	C	P				Nickrent, D, W		H	1
S	F+	C	P				Nickrent, D, W		H	**1**
S	F+	C	P				Nickrent, D, W		H	1
S	F+	C	P				Nickrent, D, W		H	**1**
S	F+	C	P				Nickrent, D, W		H	1
S	F+	C	P				Nickrent, D, W		H	**1**
S	F+	C	P				Nickrent, D, W		H	1
R	**V**	**H**	**P**	beetle, flies	mammals	edible, up to 2-years to mature	**Nickrent, 22**			**1**
R	Uk	H	P						H	1
R	F	H	P				24		H	**1**
R	G	H	P				24		H	1
R	Uk	H	P						H	**1**
R	F	H	P	blowflies	mammals		24, D, 22		H	1
R	Uk	H	P						H	**1**
R	**V**	**H**	P	beetle		edible, up to 2-years to mature	**Nickrent, D**			1
R	F	H	P				24		H	**1**
R	F+	H	P				24		H	1
R		H	P						H	**1**

Appendix S1, sheet 2

Sexual system as in sheet 1: H: hermaphrodite, M: monoecious, D: dioecious, O: others, ?: maybe	Totals resulting from sheet 1	Numbers for *Dendrophthora* from J. Kuijt, pers. comm. 2012	Numbers for *Viscum* from Heide-Jorgensen 2008	Totals with *Dendrophthora* and *Viscum*	**Percentages using 5059**	**Percentages using 4510**
H?	137			138	2.7	3.1
H	3903			3979	78.7	88.2
M?	25	3		28	0.6	0.6
M	140	34	62	236	4.7	5.2
D?	**109**	**8**		**117**	**2.3**	**2.6**
D	**371**	**40**	**45**	**456**	**9.0**	**10.1**
O	21			30	0.6	0.7
Uk/UK/?	121	48		169	3.3	3.7
Totals	4827	133	107		101.9	114.3
Total -8 because 8 species counted twice (when M+D)	4819					
Total number of species here (B12+C11+D11)	5059		**Using 5059 or 4510, the percentage of dioecious species is 9 or 10%**			
Total number of species modified according to M. Costea, M. Chase and D. Nickrent (see Sheet 1)	4510		**Whatever we take as total number the list of the sheet before ("Number Here"=5058) or if we report the number of dioecious species from this list to a slightly lower estimate total number (4510) as we chose in the manuscript, minimum percentage of dioecious species is ca. 10%**			

Appendix S1, sheet 3

Amuchastegui, A, L Petryna, JJ Cantero and C Nunez. 2003. Plantas parasitas del centro de Argentina. Acta Botanica Malacitana 28: 37-46.
Baillon, HE. 1880. Sur lÕHachettea, nouveau genre de Balanophorees. Bulletin Mensuel de la Societe Linneenne de Paris 1: 229.
Balete, DS, PB Pelser, DL Nickrent and JF Barcelona. 2010. Rafflesia verrucosa (Rafflesiaceae), a new species of small-flowered Rafflesia from eastern Mindanao, Philippines. Phytotaxa 10: 49-57.
Balle, S. 1964. Bakerella. Adansonia n.s., 4: 108.
Bentham, G and JD Hooker. 1862. Genera Plantarum 1: 344-349.
Blume, CL von. 1826. Bijdragen tot de flora van Nederlandsch Indie 1148.
Blume, CL von. 1830. Systema Vegetabilium 7: 1611-1730.
Brasil, B de Azevedo. 2010. Life cycle, phenology and floral anatomy of Pilostyles (Apodanthaceae - Rafflesiaceae s.l.): phylogenetic position allowances for Apodanthaceae family. Abstract of M.Sc. thesis. Instituto de Bioci • ncias, Sao Paulo, Brazil.
Breteler, FJ. 2001. Novitates Gabonenses 40. A new species of Strombosiopsis (Olacaceae) from Gabon. Adansonia, ser. 3. 23(2): 303-306.
Britton, N L and Millspaugh. 1920a. The Bahama Flora 109.
Britton, N L and Millspaugh. 1920b. The Bahama Flora 108.
Buen, L Lopez de and JF Ornelas. 2002. Host compatibility of the cloud forest mistletoe Psittacanthus schiedeanus (Loranthaceae) in central Veracruz, Mexico. American Journal of Botany 89(1): 95Ð102.
Burgoyne, PM. 2006. A new species of Cytinus (Cytinaceae) from South Africa and Swaziland, with a key to the Southern African species. Novon 16(3): 316Ð319.
Candolle, ALPP de. 1857. Prodromus Systematis Naturalis Regni Vegetabilis 14: 635.
Carvell, WN and WH Eshbaugh. 1982. A systematic study of the genus Buckleya (Santalaceae). Castanea 47(1): 17-37.
Dell, B, and AH Burbidge. 1981. Notes on the biology of Pilostyles (Rafflesiaceae) in Western Australia. Western Australian Herbarium Research Notes 5: 71-79.
Der, JP and DL Nickrent. 2008. A Molecular Phylogeny of Santalaceae (Santalales). Systematic Botany 33(1): 107Ð116.
Don, GA. 1834. General History of the Dichlamydeous Plants 3: 432.
Dostalek, T and Z Munzbergova. 2010. Habitat requirements and host selectivity of Thesium species (Santalaceae). Botanical Journal of the Linnean Society164: 394Ð408.
Ducke, WA. 1922. Plantes nouvelles ou peu connues de la region Amazonienne (IIeme partie). Archivos do Jardim Botanico do Rio de Janeiro 3: 41.
Eberwein, R, DL Nickrent and A Weber. 2009. Development and morphology of flowers and inflorescences in Balanophora papuana and B. elongata (Balanophoraceae). American Journal of Botany 96: 1055-1067.

Eberwein, RK and A Weber. 2004. Exorhopala ruficeps (Balanophoraceae): morphology and transfer to Helosis. Botanical Journal of the Linnean Society 146: 513Ð517.
Eichler, AW. 1868. Lathrophytum, ein neues Balanophoreengeschlecht aus Brasilien. Botanische Zeitung (Berlin) 26: 550.
Engler, HGA. 1889. Die Naturlichen Pflanzenfamilien 3(1): 189.
Fernald, ML. 1928. Contributions from the Gray herbarium of Harvard University Ð No. LXXIX. Rhodora 30(350): 23Ð24.
Gonzalez, LA and JF Morales. 2004. Una nueva especie de Struthanthus (Loranthaceae) para Costa Rica. SIDA 21(1): 97-102.
Gosline, G and V Malecot. 2011. A monograph of Octoknema (Octoknemaceae Ñ Olacaceae s.l.). Kew Bulletin 66: 367-404.
Grisebach, AHR. 1879. Abhandlungen der Kšniglichen Gesellschaft der Wissenschaften zu Gšttingen 24: 151.
Halle, N. 1978. Illustration de deux rares Balanophoracees d'Afrique equatoriale appartenant aux genres Chlamydophytum Mildbr. et Balanophora Forst. Adansonia, ser. 2, 17(3): 249-261.
Hawkeswood, TJ. 2005. Record of the mistletoe Muellerina celastroides (Sieb. ex Schulles) Tieghem (Loranthaceae) parasitizing a pear tree (Pyrus malus L., Rosaceae) at Faulconbridge, New South Wales, Australia. Calodema 4.
Hiepko, P. 1982. A revision of Opiliaceae II. Opilia Roxb. Willdenowia 12(2): 161-182.
Hooker, JD. 1859. On a new genus of Balanophoreae from New Zealand, and two new species of Balanophora. Transactions of the Linnean Society of London 22: 426.
Hurlimann, H and HU Stauffer. 1957. Daenikera, a New Santalaceae Genus in Santalales-Studien II. Mitteilungen aus dem Botanischen Museum der Universitat Zurich CCX (210).
Johri, BM and JS Agrawal. 1954. Morphology and embryology of Helicanthes elastica (Desr.) Dans. Current Science 23: 96-98.
Kawakita, A and M Kato. 2002. Floral biology and unique pollination system of root holoparasites Balanophora kuroiwai and B. tobiracola (Balanophoraceae). American Journal of Botany 89(7): 1164Ð1170.
Kuijt, J. 1988. Revision of Tristerix (Loranthaceae). Systematic Botany Monographs 19: 1-61.
Kuijt, J. 1990. A second species of Ligaria (Loranthaceae). Brittonia 42(1): 66-69.
Kuijt, J. 1991. Two new species of Ixocactus (Loranthaceae) and a reformulation of the genus. Systematic Botany 16(2): 292-298.
Kuijt, J. 1992. Two New Species of Cladocolea (Loranthaceae) from Mexico and Surinam. Novon 2(4): 351-354.
Kuijt, J. 1996. Miscellaneous Mistletoe notes 20-36. Novon 6(1): 33.
Kuijt, J. 2000. Two new Brazilian species of Oryctina (Loranthaceae) with a revised key to the genus. Novon 10(4): 391-397.
Kuijt, J. 2003. Miscellaneous Mistletoe notes 37-47. Novon 13(1): 82.
Kuijt, J. 2011. Pulling the skeleton out of the closet: resurrection of Phthirusa sensu Martius and consequent revival of Passovia (Loranthaceae). Plant Div. Evol. 129(2): 159Ð211.
Kuijt, J. 2011. Thirteen new species of Neotropical Viscaceae (Dendrophthora and Phoradendron). Novon: A Journal for Botanical Nomenclature, 21(4): 444-462.

Kuijt, J. 2011. Two new species of Oryctanthus (Loranthaceae) from Colombia and French Guiana. Novon: A Journal for Botanical Nomenclature 21(4): 463-467.
Lauterbach, CAG. 1912. Opiliaceae. Nova Guinea 8: 817.
Lehn, CR, SM Salis, P P—voa de Mattos and W Arruda. 2008. Especies arboreas parasitadas por Langsdorffia hypogaea Mart. (Balanophoraceae) no Pantanal da Nhecolandia, Corumba-MS, Brasil. II Simposio internacional Savana Tropicas 12-17 October 2008, Brasilia, DF, Brazil.
Linnaeus, C von. 1753. Species Plantarum 1: 110.
Malecot, V, GE Schatz, and J Bosser. 2003. Synoptic Revision of the genus Phanerodiscus Cavaco (Olacaceae) in Madagascar. Adansonia, ser. 3. 25(1): 119-128.
Martius CFP von. 1830. Einige Bemerkungen uber Loranthus. Flora 13(1): 106Ð107.
Michaux, A. 1803. Flora Boreali-Americana 2: 231Ð233.
Michaux, M. 1966a. Contribution a l'etude des Olacacees d'Afrique tropicale. pp 262-264 (total 157-290) in R Schnell, F Grout de Beaufort, F Bernhard, M Michaud, eds. Memoires de l'institut Fondamental d'Afrique Noire, Melanges Botaniques, vol. 75. Dakar, Senegal: Musee de Institut Fondamental d'Afrique (IFAN).
Michaux, M. 1966b. Contribution a l'etude des Olacacees d'Afrique tropicale. pp 262-264 (total 157-290) in R Schnell, F Grout de Beaufort, F Bernhard, M Michaud, eds. Memoires de l'institut Fondamental d'Afrique Noire, Melanges Botaniques, vol. 75. Dakar, Senegal: Musee de Institut Fondamental d'Afrique (IFAN).
Miquel, FAW. 1856. Flora van Nederlandsch Indie 1(1): 810, 834.
Nickrent, DL, MA Garcia, MP Martin and RL Mathiasen. 2004. A phylogeny of all species of Arceuthobium (Viscaceae) using nuclear and chloroplast DNA sequences. American Journal of Botany 91(1): 125Ð138.
Patschovsky, M. 1911. Loranthaceae peruvianae in Plantae novae andinae imprimis Weberbauerianae. V. in Botanische Jahrbucher fur Systematik, Pflanzengeschichte und Pflanzengeographie 45(4): 437.
Poiteau, MA. 1824. Description de l'Apodanthes, nouveau genre de plantes phanŽrogames parasite, Annales des Sciences Naturelles (Paris) 3: 422, t. 26
Polhill, RM and D Wiens. 1997. New Loranthaceae and Viscaceae. Novon 7(3): 277.
Rasingam, L, P Lakshminarashimhan and PG Diwakar. 2010. Rediscovery of Ginalloa andamanica Kurz (Angiosperms: Viscaceae) Ð an endemic and threatened species from Little Andaman Island, India. Journal of Threatened Taxa 2(9): 1158-1159.
Reif, C. 2007. Contribution to the taxonomy of Eremolepidaceae, Loranthaceae and Viscaceae from Rio de Janeiro State, Brazil. Universidade Santa Ursula.
Richard, LCM. 1822. Memoire sur une nouvelle famille de plantes, les Balanophorees. Memoires du Museum d'Histoire Naturelle 8: 416-432.
Rogers, ZS, DL Nickrent and V Malecot. 2008. Staufferia and Pilgerina: two new endemic monotypic arborescent genera of Santalaceae from Madagascar. Annals of the Missouri Botanical Garden 95(2): 398Ð399.

Rosales, D Alvarado and L de L Saavedra Romero. 2005. El genero Cladocolea (Loranthaceae) en Mexico: muerdago verdadero o injerto. Revista Chapingo. Serie ciencias forestales y del ambiente 11(001): 5-9.
Rotta, E, YM Malheiros de Oliveira, AJ de Araœjo and MT Inoue. 2005. Reconhecimento pratico de cinco especies de Erva-de-Passarinho na arboriza • ao de Curitiba, PR. Colombo : Embrapa Florestas, 2005.
Ruiz, H Lopez and JA Pavon. 1798. Systema Vegetabilium Florae Peruvianae et Chilensis 73.
Rusby, HH. [ed.], 1920. Descriptions of three hundred new species of South American plants. New York, New York, USA.
Rzedowski G Calderon de. 1995. Olacaceae. Flora del bajio y de regions adyacentes 34.
Schlechter, FRR and RKF Pilger. 1906. Beitrage zur Kenntnis der Flora von Neu-Kaledonien. Botanische Jahrbucher fur Systematik, Pflanzengeschichte und Pflanzengeographie 39: 103.
Smart, C. 1952. The life history of Tupeia Cham. et Schl. (Loranthaceae). Transactions of the Royal Society of New Zealand 79(3-4): 459-466.
Sobero, JR, MA Sgariglia, DA Sampietro, EN Quiroga, MG Sierra and MA Vattuone. 2010. Purification and identification of antibacterial phenolics from Tripodanthus acutifolius leaves. Journal of Applied Microbiology 108: 1757Ð1768.
Sonder, OW. 1857. Enumeratio Santalacearum in Africa australi extratropica crescentium quas Dr. Ecklon et C. Zeyher collegerunt. Flora 40: 364.
Spitzer, I, I Marx, JGS Maia, and K Pfeilsticker. 1990. Curupira tefeensis (Olacaceae) A Rich Source of Very Long Chain Fatty Acids. Fat Sci. Technol. W. 4.
Standley, PC. 1920. The North-American species of Agonandra. Journal of the Washington Academy of Sciences 10: 506.
Stapf, O. 1905. Contributions to the Flora of Liberia. Journal of the Linnean Society, Botany 37: 89.
Stauffer, HU and H Hurlimann. 1957. Amphorogyne, another Santalaceae Genus from New Caledonia in Santalales-Studien III. Mitteilungen aus dem Botanischen Museum der Universitat Zurich CCXI(211).
Stearn, WT. 1972. Kunkeliella, a new genus of Santalaceae in the Canary Island. Cuad. Bat. Canar. XVP: 11-26.
Thiele, KR, SJ Wylie, L Maccarone, P Hollick, and JA McComb. 2008. Pilostyles coccoidea (Apodanthaceae), a new species from Western Australia described from morphological and molecular evidence. Nuytsia 18: 273-284.
Tieghem, PEL van. 1894. Sur le groupement des esp • ces en genres dans la tribu des PsittacanthŽes de la famille des LoranthacŽes. Bulletin de la SociŽtŽ Botanique de France 41: 601.
Tieghem, PEL van. 1895. Sur le groupement des esp • ces en genres dans la tribu des PsittacanthŽes de la famille des LoranthacŽes. Bulletin de la SociŽtŽ Botanique de France 42: 440.
Ulloa, C Ulloa and P Moller Jorgensen. 1998. Acanthosyris annonagustata (Santalaceae), a new species from Eastern Ecuador. Novon 8(1): 84Ð86.
Urban, I. 1897. Additamenta ad cognitionem florae Indiae occidentalis. Botanische Jahrbucher fur Systematik, Pflanzengeschichte und Pflanzengeographie 24: 26.

Vattimo, I de. 1950. Pilostyles stawiarskii. Revista Brasileira de Biologia 10: 196. Vattimo, I de. 1971. Contribucoa ao conhecimento da tribo Apodanthea R. Br. Parte I Ð Conspecto das espŽcies (Rafflesiaceae). RodriguŽsia 38: 37-62.
Vattimo, I de. 1973. Notas sobre o genero Apodanthes Poit. com descricao de duas novas especies (Rafflesiaceae). Revista Brasileira de Biologia 33(1): 140.
Veenendaal, EM, IK Abebrese, ME Walsh and MD Swaine. 1996. Root hemiparasitism in a West African rainforest tree Okoubaka aubrevillei (Santalaceae). New Phytol. 134: 487-493.
Werthand, CR, WV Baird, and LJ Musselman. Root Parasitism in Schoepfia Schreb. (Olacaceae). Biotropica 11(2): 140-143.
Zebec, M and M Idzojtic. 2006. Hosts and distribution of yellow mistletoe, Loranthus europaeus Jacq. in Croatia. Hladnikia 19: 41-46.

Chapter 4

EXPLORING NEW DATING APPROACHES FOR PARASITES: THE WORLDWIDE APODANTHACEAE (CUCURBITALES) AS AN EXAMPLE.

Bellot, S. and S. S. Renner

Molecular Phylogenetics and Evolution 80: 1–1

Molecular Phylogenetics and Evolution 80 (2014) 1–10

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Exploring new dating approaches for parasites: The worldwide Apodanthaceae (Cucurbitales) as an example

Sidonie Bellot *, Susanne S. Renner

Systematic Botany and Mycology, University of Munich (LMU), Menzinger Str. 67, 80638 Munich, Germany

ARTICLE INFO

Article history:
Received 7 January 2014
Revised 4 July 2014
Accepted 7 July 2014
Available online 22 July 2014

Keywords:
Local clocks
Parasitic plants
Random local clocks
Substitution rates
Uncorrelated lognormal clock
Bayes factors

ABSTRACT

Gene trees of holoparasitic plants usually show distinctly longer branch lengths than seen in photosynthetic closest relatives. Such substitution rate jumps have made it difficult to infer the absolute divergence times of parasites. An additional problem is that parasite clades often lack a fossil record. Using nuclear and mitochondrial DNA sequences of Apodanthaceae, a worldwide family of endoparasites living inside Fabaceae and Salicaceae, we compared several dating approaches: (i) an uncorrelated lognormal (UCLN) model calibrated with outgroup fossils, (ii) ages of host lineages as a maximal age in an UCLN model, (iii) user-assigned local clocks, and (iv) outgroup-fossil-calibrated random local clocks (RLC) with varying prior probabilities on the number of permitted rate changes (RLCu and RLCp models), a variable that has never been explored. The resulting dated phylogenies include all 10 species of the family, three in Australia, one in Iran, one in Africa, and the remainder in the Americas. All clock models infer a drastic rate jump between nonparasitic outgroups and Apodanthaceae, but since they distribute the rate heterogeneity differently, they result in much-different age estimates. Bayes factors using path and stepping-stone sampling indicated that the RLCp model fit poorly, while for *matR*, topologically unconstrained RLCu and UCLN models did not differ significantly and for 18S, the UCLN model was preferred. Under the equally well fitting models, the Apodanthaceae appear to be a relatively old clade, with a stem age falling between 65 and 81 my, the divergence of *Apodanthes* from *Pilostyles* between 36 and 57 my ago, and the crown age of the Australian clade 8–18 my ago. In our study system, host-age calibrations did not yield well-constrained results, but they may work better in other parasite clades. For small data sets where statistical convergence can be reached even with complex models, random local clocks should be explored as an alternative to the exclusive reliance on UCLN clocks.

1. Introduction

Over the past 15 years, relaxed molecular clock approaches (Sanderson, 1997, 2002; Thorne et al., 1998; Huelsenbeck et al., 2000; Drummond et al., 2006; Rannala and Yang, 2007) have become the dominant strategy for inferring absolute times from molecular data. This is because fossil-calibrated studies have often revealed punctuated rates of substitutions, which prevent the use of global clocks (Li and Tanimura, 1987; Yoder and Yang, 2000; Bell et al., 2010). To accommodate such situations, workers have assumed either that substitution rates of ancestors and descendants are correlated (Sanderson, 1997, 2002; Thorne et al., 1998; Rannala and Yang, 2007) or that rate variation is unrelated to species relationships (Drummond et al., 2006; Rannala and Yang, 2007). In the latter case, the rate for each branch is drawn from a single underlying distribution, such as a lognormal or exponential distribution of which the parameters are estimated in the analysis (Drummond et al., 2006). A problem with the uncorrelated lognormal (UCLN) model of Drummond et al. is that it may bias results in favor of rate variation, making UCLN clocks poorly suited for data sets in which a strict clock best explains the branch lengths. Rannala and Yang (2007) suggested that this bias should decrease in trees with many branches.

Another approach for dealing with abrupt rate variation is to use local clocks in which regions of a phylogeny are calibrated using paleontological time points, with each region being given its own substitution rate (Li and Tanimura, 1987; Yoder and Yang, 2000). Local molecular clocks have been used in relatively few studies (e.g., Bailey et al., 1991; Yoder and Yang, 2000; Baum et al., 2004; Aguileta et al., 2006; Nunome et al., 2007; Aris-Brosou, 2007), perhaps because they require an *ad hoc* choice of nodes in the phylogeny where a rate change is assumed to have occurred. This permits neither the modeling of uncertainty in tree topology nor in the clade most likely to have undergone the abrupt rate change. To overcome this problem, Drummond and Suchard

* Corresponding author.
E-mail address: sido.bellot@gmail.com (S. Bellot).

http://dx.doi.org/10.1016/j.ympev.2014.07.005
1055-7903/

(2010) developed the random local clock (RLC) approach, which employs a Monte Carlo Markov chain to investigate nested local clock configurations. Their method co-estimates the phylogenetic tree and the number, magnitude and location of rate changes along it. Importantly, the RLC model includes the possibility of zero rate changes, meaning it can also serve to test whether a single global rate fits a particular data set. To compare the fit of RLC and UCLN models, authors have used Bayes factors because these models are not nested (Avaria-Llautureo et al., 2012). So far, nobody seems to have studied the impact of changing the *a priori* permitted number of rate changes in an RLC model.

Abrupt shifts in molecular rates characterize the holoparasitic lineages of flowering plants, judging from branch lengths in trees from nuclear 18S and 26S ribosomal RNA sequences, mitochondrial *matR*, *atp1*, *cox1*, and *nad1*, and the plastid 16S ribosomal RNA (Nickrent and Starr, 1994; Barkman et al., 2007; Bromham et al., 2013). Parasitism has evolved at least 13 times in the angiosperms (Barkman et al., 2007) and is associated with drastic changes in photosynthesis-related genes of the plastome, probably also in the mitochondrial and nuclear genomes (Krause, 2012; Wicke et al., 2013). In addition, many physiological, ecological, and other life history changes occur with the onset of a parasitic life style, perhaps including changes in effective population size or mating system; most parasitic lineages are long-lived and many are dioecious (Bellot and Renner, 2013). These changes likely modified selection on mutation rates or DNA repair. Regardless of their causes, the extremely long stem lineages seen in holoparasitic angiosperm clades have made it difficult to fit clock models for these groups. Nevertheless, Naumann et al. (2013) recently calculated the stem ages of all 13 parasitic lineages using an UCLN clock model and (mostly) single representatives of each parasite lineage.

Here we use the parasite family Apodanthaceae to compare local clocks, random local clocks, and clock models relying on the UCLN prior rate distribution, calibrated with either outgroup fossils or host ages. If a parasite species or clade occurs only on one host species or genus, the divergence time of the host provides a maximal age for the parasite's age, because a host-specific parasite is unlikely to be older than its host. This type of calibration has not been tried in plants, but may be useful for specialized endoparasites, such as Apodanthaceae, at least as a cross-validation of other calibration approaches. All species in this family lack any green parts and live permanently inside trees or shrubs of Fabaceae or Salicaceae from the stems of which they emerge to flower and fruit (Bellot and Renner, 2014). The family is distributed in North and South America, Africa, Iran, and Australia, and its 10 species belong to two genera, *Pilostyles* and *Apodanthes* (Bellot and Renner, 2014).

From the worldwide distribution of Apodanthaceae, it is almost certain that transoceanic dispersal must have played a role in the family reaching the different continents. Compared to the three other endoparasitic lineages of flowering plants, Cytinaceae with 10 species, Mitrastemonaceae with two, and Rafflesiaceae with 34, Apodanthaceae have by far the largest geographic and host range, which suggests that they may be older than these other lineages. With a temporal framework it would be possible to infer how long it took these parasites to evolve host specificity to either Fabaceae or Salicaceae and to acquire their extreme adaptations. Absolute ages also are needed to test hypotheses about trait correlations (Hardy and Cook, 2012) and genomic changes, for example in the plastome (Wicke et al., 2013). To achieve these ultimate goals, we here focus on three questions, (i) what is the most plausible model of substitution rate change in the Apodanthaceae; (ii) how different are the divergence times estimated with user-specified local clocks, random local clocks, or an uncorrelated lognormal clock model; and (iii) can plausible ages be inferred by using host ages as maximal constraints on parasite ages (as an alternative calibration instead of an outgroup fossil).

2. Material and methods

2.1. Plant material, DNA isolation, PCR amplifications and sequencing

DNA was extracted from 21 specimens of Apodanthaceae representing most of the named species (de Vattimo, 1971). Table S1 shows herbarium vouchers and Genbank accession numbers. Total genomic DNA was extracted from herbarium specimens or silica-dried flowers using the commercial plant DNA extraction kit NucleoSpin (Macherey–Nagel, Düren, Germany) or alternatively the NucleoSpin® Food kit of the same company. To overcome problems with fragmented DNA, we designed custom primers for *matR* and 18S using Primer3Plus v. 2.3.6 (Untergasser et al., 2012); Table S2 lists all primers, which were used for both polymerase chain reactions (PCRs) and sequencing. PCR products were purified with the ExoSAP or FastAP clean-up kits (Fermentas Life sciences, St. Leon-Rot, Germany), and sequencing relied on the Big Dye Terminator cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) and an ABI 3130-4 automated capillary sequencer. In total, 35 *matR* and 18S sequences were newly obtained for this study. The 18S sequence of *P. aethiopica* was obtained by a BLAST search against Illumina data of its genomic DNA (Bellot and Renner, unpublished). Additional 18S and *matR* sequences of Malpighiales, Fagales and Cucurbitales were retrieved from GenBank, and their accession numbers are provided in Table S1.

2.2. Sequence alignment and phylogenetic analyses

All sequences were blasted against GenBank to rule out contamination. Chromatograms were checked and sequences were edited using BioEdit v. 7.2.0 (Hall, 1999) and Geneious R7 (Biomatters, available from http://www.geneious.com). Two alternative alignments were performed for each marker, using the program MAFFT v. 7 (Katoh, 2013). The 18S alignments were obtained using either the "E-INS-i" strategy with manual editing or the "Q-INS-i" strategy, which takes into account the secondary structure of the RNA, without further manual editing. The *matR* alignments were obtained using the E-INS-i strategy with some manual editing, with and without translation into amino acids. Since the alignment methods did not affect the maximum likelihood (ML) topologies obtained, we chose the Q-INS-i-based alignment of 18S and the alignment performed without considering amino acids for *matR* in final analyses. Alignments and ML trees have been deposited in TreeBASE (accession number 15658). To select the best fitting model of nucleotide substitutions, JModelTest v. 2.1 (Darriba et al., 2012) was run on the final matrices with the following parameters: 11 substitution schemes, +F, +G 4 categories, ML optimized, and Best tree search, and the best model was then chosen using the corrected Akaike Information Criterion with model averaging allowed. For both markers, GTR + G was the preferred model. Phylogenetic searches using the separate or combined DNA matrices were performed under maximum likelihood optimization as implemented in RAxML-7.2.8-ALPHA (Stamatakis, 2006), with 100 bootstrap replicates, and rooting on *Clusia rosea* (Malpighiales).

2.3. Analysis of substitution rates and molecular dating

For clock dating, we constructed DNA alignments of nuclear 18S and mitochondrial *matR* with and without outgroups of Apodanthaceae. Matrices included 8 ingroup *matR* sequences, 10 ingroup 18S sequences, and 11, 12 or no outgroup sequences (the monospecific genus *Apodanthes* was then used as the outgroup). We used the Bayesian approach implemented in the software BEAST 1.8.0 (Drummond et al., 2006, 2012). All analyses used the GTR + G (4

categories) model of nucleotide substitutions, with empirical base frequencies, and a Yule tree prior. Each analysis was performed at least twice, with Markov chain lengths of 60, 200, 300 or 500 million generations, sampling every 10,000th generation. Some analyses were performed using the CIPRES Science Gateway (Miller et al., 2010). A burn-in fraction of 10% was removed in Tree Annotator (part of the BEAST package) before exporting a maximum clade credibility tree. Convergence was checked in Tracer v. 1.5 (Rambaut and Drummond, 2007), and the final tree was visualized in FigTree v. 1.3.1 or with the R modules PHYLOCH and SDMTools (Heibl, 2008; VanDerWal et al., 2012; R Core Team, 2013). Effective sample sizes (ESS) for all parameters were >200 after 60 million generations for most analyses and 200, 300 or 500 million generations for random local clocks analyses. To obtain Bayes factors, we performed marginal likelihood estimations for all analyses using the path and stepping-stone sampling methods implemented in BEAST v. 1.8.0, with the default settings.

We constrained the monophyly of the ingroup (Apodanthaceae), the monophyly of Fagales and Cucurbitales (in three 18S analyses), the relationships between African and Australian species (in one 18S analysis), the monophyly of the Australian species (in one *matR* and one 18S analyses), and a clade of *P. berteroi*, *P. blanchetii*, *P. mexicana* and *P. thurberi*, the relationship of which was resolved by *matR* and the combined data, but not by 18S alone. In addition, we constrained three family relationships outside Apodanthaceae (each in one analysis) to ensure identical topologies as necessary for some model comparisons. Analyses were also performed on unconstrained topologies, and the mean ages obtained are identical or near-identical, at least with Apodanthaceae. We provide Bayes factors for constrained and unconstrained analyses (Section 3).

We applied four dating approaches, the first of which used host divergence time as a maximal age cut-off for the divergence time of the parasites, while the other three relied on an outgroup fossil for calibration. The host calibration approach focused on the Australian *P. hamiltonii*, which only parasitizes species of *Daviesia*, a legume genus with 119 species in the region of Australia where *P. hamiltonii* occurs. Its sister species *P. collina* and *P. coccoidea* parasitize species of *Oxylobium*, *Gastrolobium* and *Jacksonia* (Thiele et al., 2008), legume genera with ca. 30, 48, and 67 species in the relevant region of Australia. These legume genera are closely related (Crisp and Cook, 2003), and the divergence between *Daviesia* and the other three dates to 48.4 million years (my; standard deviation 1.3 my; Lavin et al., 2005), so we calibrated the divergence of *P. hamiltonii* and *P. collina* with a normal prior distribution with a mean of 48 my and standard deviation (SD) of 15 my, truncated at 48 my, letting the majority of ages fall between 10 and 48 my ago (Fig. 1A). This "host calibration" was applied to the 18S and *matR* Apodanthaceae-only matrices using an UCLN clock.

The fossil used to calibrate the other three relaxed clocks was the oldest Fagales cupule dated to 72.1–83.6 my (Upper Cretaceous, Campanian), which provides a minimal age for the divergence between Fagales and Cucurbitales (Herendeen et al., 1995). Based on Wang et al.'s (2009) dating of the Fagales/Cucurbitales split to 85–109 my ago, we assigned the fossil a normal prior distribution with a mean of 95 my, such that 95% of the ages would fall between 85 and 105 my (SD 5 my). For the random local clock (RLC) we chose as a prior distribution for the clock rate a normal distribution with a mean of 0 and standard deviation of 1. We set the prior number of rate changes to the default Poisson distribution of log2 (hereafter called RLCp), meaning that the prior number of rate changes had 50% probability to be null (global clock). Alternatively, we used a uniform prior that gave equal probability to any number of rate changes between 0 and 20 (hereafter RLCu), a range that included the number of rate changes inferred under the UCLN model (see Section 3.3).

For the local clock (LC; Fig. 1B), branch lengths in Cucurbitales (Fig. S1) were calibrated with the Fagales fossil to obtain a substitution rate with which the stem age of Apodanthaceae (node T_2 in Fig. 1B) was then calculated. This stem age in turn was used to calibrate branch lengths in the Apodanthaceae and thus obtain Apodanthaceae-specific *matR* and 18S rates.

3. Results

3.1. Phylogenetic reconstructions

The maximum likelihood phylogenies from the nuclear and mitochondrial DNA matrices are shown in Fig. 2A and B. They use the traditional species names before the synonymizations carried out by Bellot and Renner (2014, including morphological measurements for 123 dissected flowers) who recognize 10 biological species. Figs. 2C and 3 show the accepted species names. *Pilostyles* is monophyletic, and the African species *P. aethiopica* (only 18S available) is nested among *Pilostyles* species from Iran and Australia. Both 18S and *matR* reveal an American and an Australian clade of specimens, and 18S moreover indicates that the African species (*P. aethiopica*) and the Iranian species (*P. haussknechtii*) are closer to the Australian species than to the American ones. Both genetic markers placed several individuals previously bearing different names in morphologically and geographically homogenous clusters; examples are *P. thurberi* (one asterisk on Fig. 2A and B), *P. blanchetii* (two asterisks) and *A. casearia*e (three asterisks). Since none of the topological differences between the nuclear and mitochondrial trees was statistically supported, the two data matrices were combined (Fig. 2C). In the tree from the combined matrix, the USA-Mexican *P. thurberi* is sister to the remaining American *Pilostyles*.

3.2. Inferred substitution rates and rate jumps

For both *matR* and 18S, branch lengths (genetic distances) in the crown group of Apodanthaceae are not longer than branches in Cucurbitales. However, the apodanth stem is comparatively long (Fig. S1 and Table S3); in the *matR* gene it has on average 5× more substitutions per site than in the remaining cucurbits. The mean number of substitutions per site in 18S is 11× higher than in the other Cucurbitales. Using the Fagales fossil (Section 2 and Fig. S1), we calculated a substitution rate for the non-parasitic Cucurbitales (Table S3), which allowed us to calculate a stem age of 79 my from the Apodanthaceae *matR* data (node T_2 in Fig. 1B) or 73 my from 18S data. We then applied 79 or 73 my as secondary calibrations (Table S3) and with the former obtained an Apodanthaceae-specific rate of 0.0022 substitutions per site per my (subst./site/my) for *matR* and of 0.0018 subst./site/my for 18S, with the latter 0.0024 and 0.0016 subst./site/my. Since the rates derived from the 79 or 73 my stem constraints were close to each other, we used their means, i.e., 0.0023 subst./site/my for *matR* and 0.0017 subst/site/my for 18S to calibrate our local Apodanthaceae clocks.

Fig. 3 shows the *matR* and 18S chronograms obtained with the UCLN and the two RLC approaches (Section 2), with slow and fast rates color-coded; mean rates in subst./site/my are reported in Table 1. For both loci, the UCLN model inferred a high rate in the stem of Apodanthaceae compared to the rate of their cucurbit outgroups (22-fold for *matR* and 46-fold for 18S), with a subsequent decrease of the rate in crown group Apodanthaceae to a value only 2× (*matR*) or 4× (18S) higher than that of other Cucurbitales (Table 1). Within Apodanthaceae, the UCLN model showed 8 rate changes on the *matR* topology and 7 on the 18S tree (Fig. 3A and D).

The RLCp clock model inferred a 2× lower difference between the Apodanthaceae stem rate and that of sister cucurbits than did the UCLN clock model, and no subsequent rate decrease in

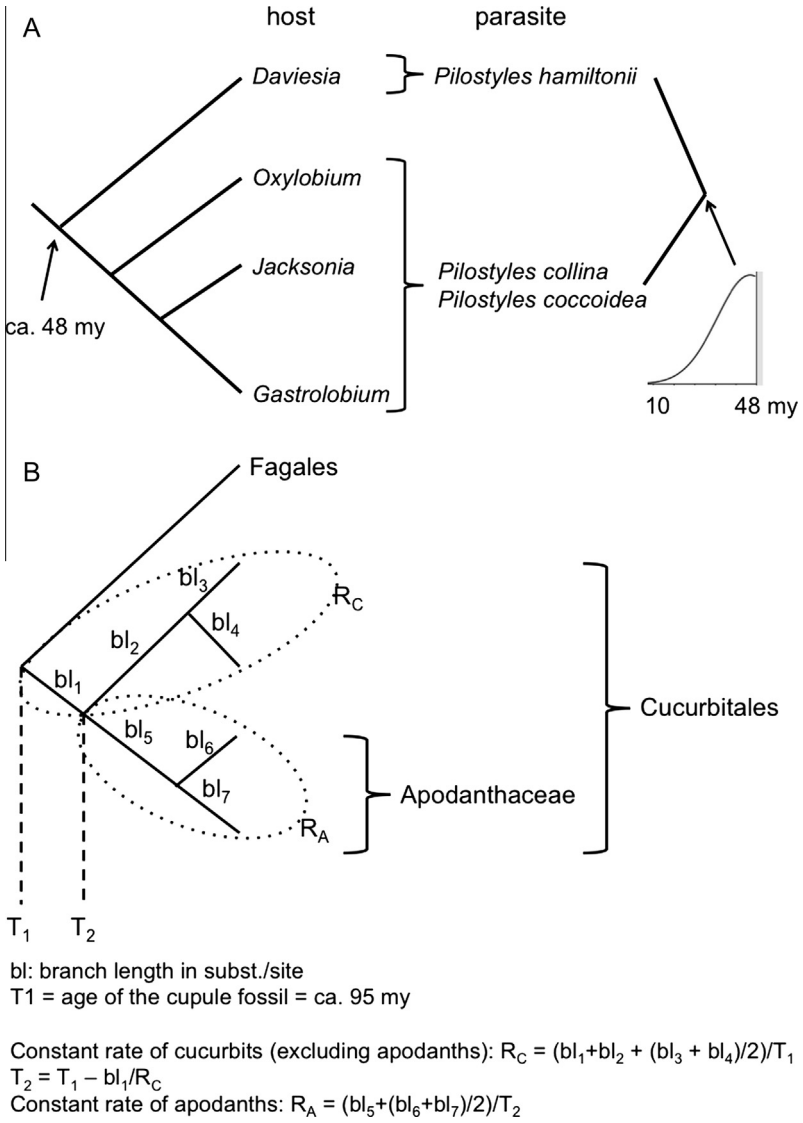

Fig. 1. (A) The "host calibration" approach, which consisted of using the divergence between specific hosts as a maximal age for specialized parasites (Section 2). Host phylogeny adapted from Crisp and Cook (2003) and host divergence times from Lavin et al. (2005). (B) The local clock approach, which consisted in assuming a divergence time T_1 between Fagales and Cucurbitales of ca. 95 my (Herendeen et al., 1995) and different, but constant, substitution rates within Apodanthaceae (R_A) and within the remaining Cucurbitales (R_C).

the crown group of Apodanthaceae. Instead, the apodanth rate stabilized around a value 6× (*matR*) or 27× (18S) higher than the rate of the remaining Cucurbitales (Fig. 3B and E; Table 1). The numbers of inferred rate changes within Apodanthaceae under the RLCp model were 1 for *matR* and 3 for 18S (Table 1).

The RLCu model produced a pattern of rate distribution intermediate to the ones obtained using the two other models. It inferred a stem rate for apodanths 14× (*matR*) or 35× (18S) higher than the stem rates of other cucurbits, followed by a decrease of the rate in the apodanth crown to a value 4× (*matR*) or 10× (18S) higher than the mean rates of nonparasitic cucurbits. The number of rate changes inferred by the RLCu clock model inside Apodanthaceae was 5 for matR and 10 for 18S (Fig. 3C and F; Table 1).

3.3. Results from different calibration approaches, and model comparisons using Bayes factors

Divergence times and 95% highest posterior density (HPD) intervals obtained with the four calibration approaches vary considerably (Table 2), and the mitochondrial and nuclear markers yielded different ages, albeit with almost completely overlapping 95% HPDs. In general, 18S gave older ages than *matR* (see Table 2). The host calibration approach resulted in extremely large confidence intervals, making it difficult to compare its result to those from the other three approaches. As expected from the stochastic nature of nucleotide substitutions, short branches, such as those between the Australian *P. hamiltonii* and *P. collina*, and among the American species of *Pilostyles*, showed the largest differences among the different models.

The UCLN clock model yielded older ages than the RLCu and RLCp models (Table 2), while ages from the RLCp and the LC were similar, except for the split between *P. aethiopica* and *P. haussknechtii* (18S: RLCp: 5 my, LC: 15 my) and that between *P. aethiopica* and *P. hamiltonii* (18S: RLCp: 12 my, LC: 20 my). Ages from the RLCu model were always older than ages from the RLCp model but younger than the UCLN ages. Ages obtained with the UCLN and RLCu models had larger confidence intervals than those obtained with the RLCp and LC models.

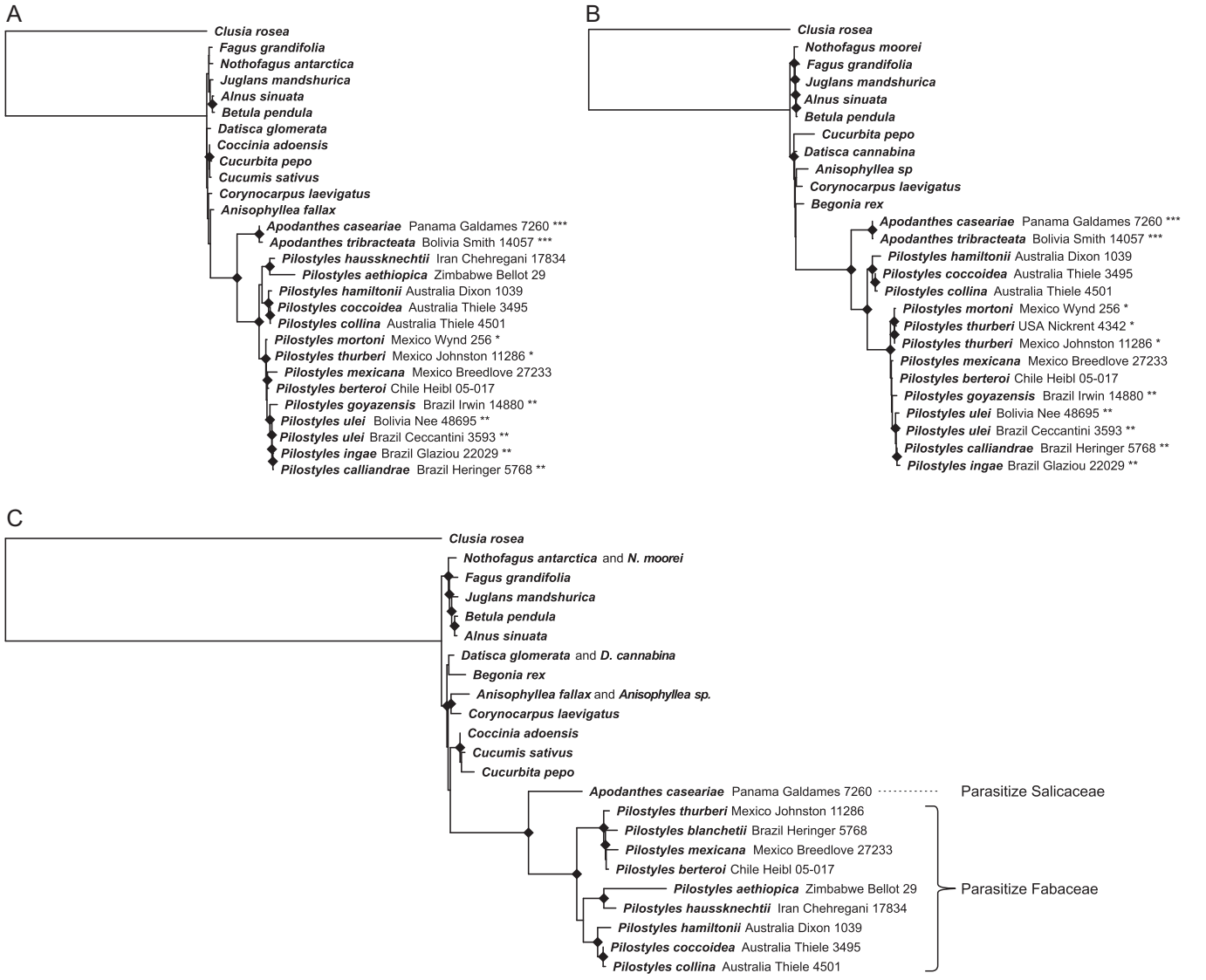

Fig. 2. Phylogenetic relationships of Apodanthaceae and outgroups obtained from 18S (A), *matR* (B) and the combined data (C) with different taxon sampling. Black diamonds at nodes indicate bootstrap support ⩾70%. Numbers after species names are GenBank accessions or collection numbers (see Table S1). Names marked with one, two, or three asterisks are synonyms of, respectively, *P. thurberi*, *P. blanchetii*, or *Apodanthes casearіae*. Traditional species names are used on (A) and (B) while (C) uses the names accepted by Bellot and Renner (2014).

The natural logarithms (ln) of the marginal likelihoods of the UCLN, RLCp and RLCu models are reported in Table 3 for *matR* and in Table 4 for 18S, which also includes the ln(Bayes factors), that is, the ratios of the ln(marginal likelihoods) of the different models. Path sampling and stepping-stone sampling gave similar results. We followed Kass and Raftery (1995) in interpreting Bayes factors in terms of decision-making. For both markers, the models involving constrained topologies were significantly better than the alternatives, with ln(Bayes factor) > 12. The RLCp model always performed worst. For 18S, the UCLN model was better than the RLCu model, with a ln(Bayes factor) > 3. For *matR*, the fit of the topologically unconstrained RLCu and UCLN models did not differ significantly.

Based on these Bayes factors, the models that best explain our data are the UCLN and RLCu models. Under these models, applied either to *matR* or 18S, Apodanthaceae originated 65–81 my ago (Table 2) and *Apodanthes* and *Pilostyles* diverged 36–57 my ago. The common ancestor of the American *Pilostyles* is 6–25 my old, and the most recent American split, involving the Mexican *P. mexicana* and the Chilean *P. berteroi* dates to less than 3–13 my.

4. Discussion

Our comparison of the random local clock model (Drummond and Suchard, 2010) with the more widely used local clock and uncorrelated lognormal clock models illustrates the great effects that model choice is having on Bayesian age estimation. The magnitude of these effects may currently be underappreciated because few studies can afford the time required to explore them, at least if they focus on larger data sets. That the family Apodanthaceae comprises just 10 species worldwide means that data matrices were sufficiently small for even parameter-rich models to reach stationarity, a problem in other explorations of the random local clock approach (Dornburg et al., 2012; Rothfels and Schuettpelz, 2014). Despite their small number of taxa (and hence nodes), however, our 18S and *matR* datasets contained distinctly uneven substitution rates and an abrupt rate increase in the Apodanthaceae stem lineage (Fig. 3). For this reason, the four compared clock models (UCLN, LC, RLCp, RLCu) resulted in highly different distributions of ancestor/descendent rate heterogeneity. The next sections discuss the different dating approaches and the

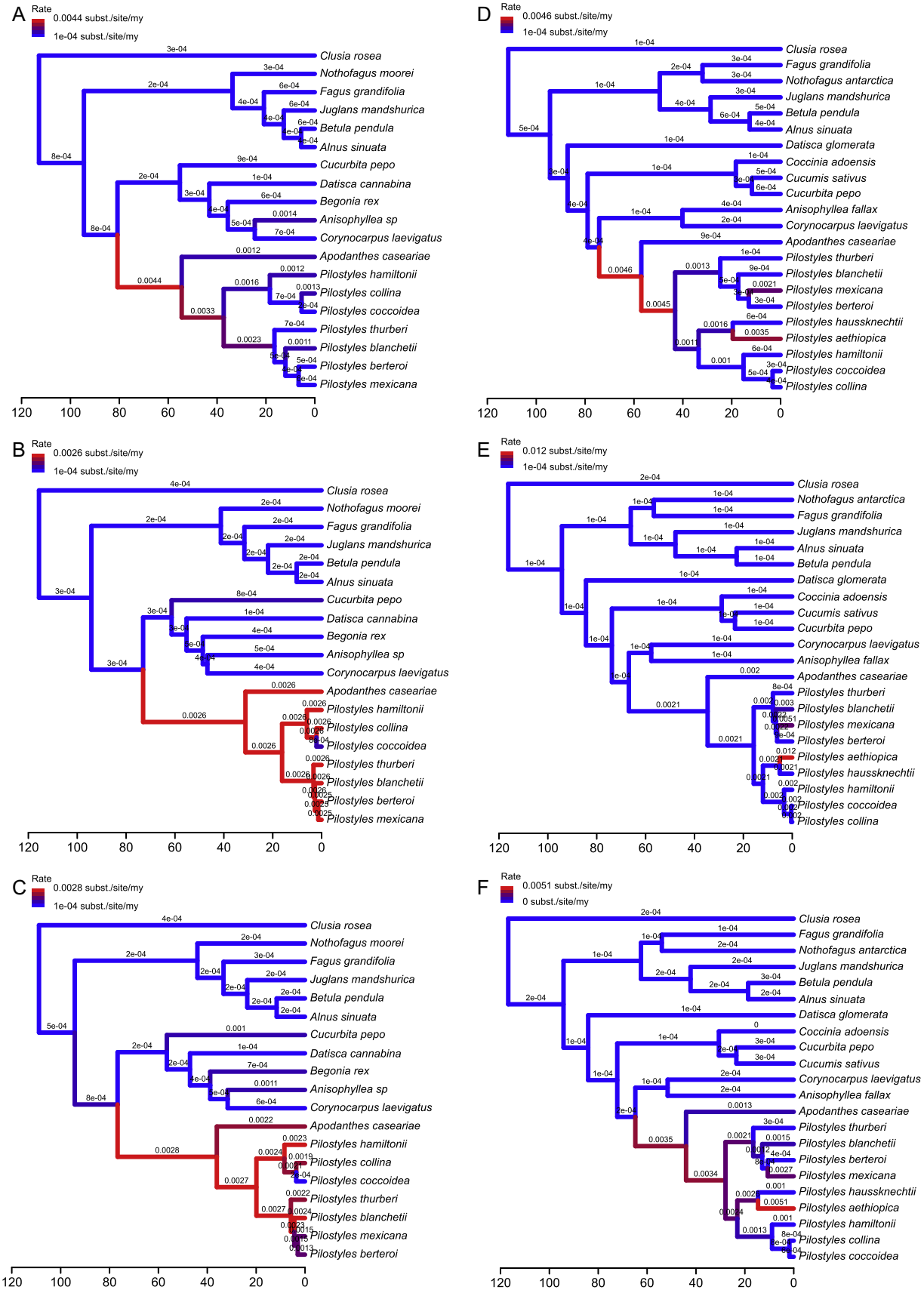

Fig. 3. Chronograms illustrating substitution rates in the mitochondrial *matR* tree (A–C) or the nuclear 18S tree (D–F). Numbers above branches are substitutions per site per million years (subst./site/my). (A and D) Rates obtained under the assumption of an uncorrelated lognormal relaxed clock (UCLN). (B and E) Rates obtained from the random local clock (RLC) with a Poisson prior distribution on the number of rate changes (RLCp). (C and F) Rates obtained from the RLC with a uniform prior distribution on the number of rate changes (RLCu). The scale bars represent the time in my before the present.

Table 1
Mean rates of substitution in subst./site/my of *matR* and 18S in Apodanthaceae and other Cucurbitales calculated as described in Section 2.

	matR UCLN	*matR* RLCp	*matR* RLCu	18S UCLN	18S RLCp	18S RLCu
Mean rate of non-parasitic cucurbits	0.00057	0.00041	0.00053	0.00029	0.0001	0.00016
Stem rate of the apodanths sister group	0.0002	0.0003	0.0002	0.0001	0.0001	0.0001
Stem rate of apodanths	0.0044	0.0026	0.0028	0.0046	0.0021	0.0035
Crown rate of apodanths	0.0011	0.0025	0.0020	0.0011	0.0027	0.0016
Mean rate of apodanths	0.0013	0.0025	0.0020	0.0013	0.0027	0.0017
Ratio apodanths/sister group stem rates	22	9	14	46	21	35
Ratio apodanths stem rate/other cucurbits mean rate	8	6	5	16	21	22
Ratio apodanths crown rate/other cucurbits mean rate	2	6	4	4	27	10
Number of rate changes inside Apodanthaceae as visible on Fig. 3	8	1	5	7	3	10

implications of our preferred models for the biogeography of the family.

4.1. Using host ages as maximal constraints

Our exploration of host ages as a possible maximal age constraint for parasites, an approach never before tried for plants, yielded disappointing results because the inferred parasite ages all had very large confidence intervals. The single species of *Apodanthes* is sister to a monophyletic *Pilostyles* (Fig. 2C), and the two genera parasitize either Salicaceae or Fabaceae. Our results based on the outgroup-fossil-calibrated-RLCu approach show that the divergence between *Pilostyles* and *Apodanthes* is probably younger than one of the two host genera of *Apodanthes*, *Casearia*, a large genus of 180 species that diverged from the remaining Salicaceae 63 (55–71) my ago (Bell et al., 2010). The age of the other host genus of *Apodanthes*, *Xylosma* (also Salicaceae), has not been inferred. In the same way, the RLCu and UCLN models show that the Australian species of *Pilostyles* diverged 8–18 my ago, much later than the 48 my-old divergence of their hosts (Lavin et al., 2005). Apodanthaceae genera and species thus are younger than their hosts, justifying our use of the host divergence as a maximal age for the parasites. In cases where a specialized parasite and its host species both are embedded in the crown groups of their respective phylogenies (making them *a priori* relatively young), host ages as maximum calibrations could be useful as a complementary approach to other calibration approaches.

4.2. Choosing the best-fit clock model

Following the introduction of the BEAST software (Drummond et al., 2006), an ever-increasing number of studies are using the uncorrelated lognormal (UCLN) relaxed clock model, usually without reporting whether alternative models, such as local clocks or a global strict clock were tested. The two other studies that have compared random local clocks to other models (Dornburg et al., 2012; Rothfels and Schuettpelz, 2014) also found strong evidence for punctuated shifts in nucleotide substitution rates as did we. Using simulated and empirical data, Dornburg et al. (2012) concluded that the RLC model better handled abrupt rate shifts than did the UCLN model, although they did not use Bayes factors to compare model fit. Their dataset (32 nodes) contained several clades with distinct rates, locally homogenous but very different from each other, resulting in a distribution of separate rates around multiple modes. This situation does not fit the prior assumption of the UCLN model, namely a unimodal distribution of rate variation (Drummond et al., 2006). The UCLN model best fits situations where only a few branches have a drastically different rate (these are outliers in the unimodal prior distribution). By contrast, the RLC model has an autocorrelation parameter and implements independent rate multipliers (Drummond and Suchard, 2010), fitting data with drastically different rates among clades, but homogeneous rates within clades. Rothfels and Schuettpelz (2014; dataset with 24 nodes) could not choose between the UCLN and RLC models.

We are unaware of any other study exploring the effects of the number of *a priori* permitted rate changes in RLC models. In our data set, the prior on the number of rate changes proved to have strong effects, and workers using RLC models should bear this possibility in mind. The default Poisson prior distribution of the number of rate changes in BEAST assigns 50% probability to there not being any rate jumps. The alternative tested here is to use a prior distribution that assigns equal probability to, for example, 0–20 jumps as in our RLCu model. The default RLCp model led to the inference of only two local clocks, one for outgroups and one for Apodanthaceae, while the RLCu model, which we prefer (below), inferred a similar (*matR*) or a higher (18S) number of rate jumps than the UCLN, albeit smaller ones. These different ways of distributing rate heterogeneity across the tree resulted in younger ages in the RLCp model compared to the UCLN model, and intermediate ages in the RLCu model.

The large difference of ages obtained with the RLCu and the UCLN models (up to 18 my) is worrisome and appears due to the bias of the UCLN model against autocorrelation in sister branches (Rannala and Yang, 2007), which can be seen in our chronograms in which the UCLN almost never inferred the exact same rate in two sister branches (never in *matR* and only once in 18S), whereas the RLCu inferred identical rates in sisters three times in *matR* and six times in18S. The UCLN may indeed be inferring rate changes that never happened in reality, with consequences for the obtained ages. In addition, the RLCu may be biased in favor of the autocorrelation of rates between mother and daughter branches (implemented via the autocorrelation parameter $\rho_{pa(k)}$; see Drummond and Suchard, 2010). This favors many small changes over few drastic changes, explaining the higher number of rate jumps inferred by the RLCu compared to the UCLN. The difference in the number of inferred jumps for the 18S compared to the *matR* data (Fig. 3A and C vs Fig. 3D and F) could indicate that the 18S rates are less autocorrelated between mother and daughter branches than the *matR* rates.

Statistical model choice between the non-nested RLC, UCLN, and LC models is "a very challenging task" (Drummond and Suchard, 2010: p. 10). However, Bayes factors can perform this task, especially since the introduction of path sampling and stepping stone sampling as ways to estimate the marginal likelihood of the models (Baele et al., 2012). Using Bayes factors, the RLCp model, which assigns a 50% prior probability to 0 rate changes, performed the worst, which suggests that no single local clock fits our data, and this is the reason why we decided to reject both the RLCp and the LC models. For 18S, the UCLN model was better than the RLCu model, but for *matR*, the fit of the topologically unconstrained RLCu and UCLN models did not differ significantly. Both models have biases (against or in favor of rate autocorrelation; above), and in our data set, statistical comparison did not provide a clear

Table 2
Ages in millions of years (my) for Apodanthaceae divergences of biogeographic interest, with 95% HPDs given in brackets. UCLN: Uncorrelated lognormal relaxed clock model implemented in BEAST (Drummond et al., 2006). Host calibration refers to the approach in which the host genus age was used as a maximal constraint for the parasite (Fig. 1). Fossil calibration refers to a Fagales cupule dated to 84 my and for which we assigned a prior age distribution of 95 my, with a SD 5 my. The random local clock model was first implemented in BEAST v. 1.5.4 (Drummond and Suchard, 2010). The local clock approach is described in Section 2 and Fig. 1B. Calibrated nodes are in bold. -C indicates that the topology was constrained. NA indicates that the respective species was not present in the matrix.

Nodes of interest	Geographic separation involved	Host calibration + UCLN		Fossil calibration + UCLN		Fossil calibration + random local clock with Poisson prior		Fossil calibration + random local clock with uniform prior		Local clock	
		matR-C	18S-C	*matR*-C	18S-C	*matR*	18S-C	*matR*	18S-C	*matR*	18S-C
Cucurbitales/Fagales	NA	NA	NA	**94 (85, 104)**	**94 (85, 105)**	**94 (85, 104)**	**94 (84, 104)**	**94 (84, 104)**	**94 (85, 104)**	NA	NA
Stem of Apodanthaceae	NA	NA	NA	81 (62, 98)	74 (54, 94)	73 (55, 90)	67 (50, 84)	77 (50, 98)	65 (45, 84)	NA	NA
Pilostyles/*Apodanthes*	NA	60 (0, 187)	88 (0, 341)	54 (35, 74)	57 (38, 77)	31 (21, 42)	35 (21, 48)	36 (19, 55)	44 (25, 64)	33 (27, 40)	36 (29, 43)
P. thurberi/*P. hamiltonii*	America/Australia	37 (0, 114)	62 (0, 239)	37 (20, 56)	43 (24, 60)	16 (11, 22)	16 (9, 22)	20 (9, 35)	28 (11, 48)	18 (15, 22)	22 (18, 26)
P. collina/*P. hamiltonii*	NA	**17 (0, 44)**	**14 (0, 43)**	18 (5, 34)	15 (2, 30)	6 (4, 9)	3 (1, 6)	8 (3, 17)	9 (1, 20)	6 (4, 9)	4 (2, 6)
P. thurberi/*P. berteroi*	Mexico + USA/Brazil + Central America	10 (0, 32)	24 (0, 92)	17 (5, 30)	25 (9, 41)	3 (2, 5)	8 (2, 14)	6 (1, 18)	17 (4, 32)	3 (2, 5)	7 (5, 9)
P. blanchetii/*P. berteroi*	Chile/Brazil + Central America	8 (0, 25)	17 (0, 67)	12 (3, 23)	17 (4, 30)	3 (1, 4)	7 (1, 12)	4 (1, 11)	13 (2, 26)	3 (2, 4)	6 (4, 8)
P. mexicana/*P. berteroi*	Mexico/Brazil + Central America	4 (0, 13)	14 (0, 56)	7 (0, 15)	13 (3, 26)	1 (0, 2)	6 (1, 11)	3 (0, 9)	11 (1, 22)	1 (1, 2)	6 (4, 7)
P. aethiopica/*P. hamiltonii*	Africa/Australia	NA	47 (0, 184)	NA	33 (17, 50)	NA	12 (7, 17)	NA	23 (8, 38)	NA	19 (16, 23)
P. aethiopica/*P. haussknechtii*	Africa/Middle East	NA	28 (0, 114)	NA	19 (5, 35)	NA	5 (2, 8)	NA	15 (3, 26)	NA	14 (11, 18)
Number of tips in the tree		8	10	19	22	19	22	19	22	8	10

Table 3
Marginal likelihood estimations performed with BEAST using the stepping-stone sampling method and ln(Bayes factors) of the UCLN, RLCp and RLCu models applied to *matR*. (c) Indicates that the topology was constrained, (nc) that it was not constrained.

	Model 1	UCLN (c)	UCLN (nc)	RLCp (nc)	RLCu (nc)
Model 2	Ln(ml)	−6196.38	−6201.22	−6211.54	−6202.52
UCLN (c)	−6196.38	–	−4.84	−15.16	−6.14
UCLN (nc)	−6201.22	4.84	–	−10.32	−1.31
RLCp (nc)	−6211.54	15.16	10.32	–	9.02
RLCu (nc)	−6202.52	6.14	1.31	−9.02	–

model preference so we discuss the biogeographic implications of the entire age range obtained with the *matR* and 18S data.

4.3. Biogeography of Apodanthaceae

Under the RLCu and the UCLN models applied to *matR* or 18S, the mean stem age of Apodanthaceae is 65–81 my (Table 2). A dating study of holoparasitic angiosperms that used mitochondrial *cox1* and *matR* sequences, and included a single representative of Apodanthaceae with a single species of *Begonia* as cucurbit outgroup, obtained an age of 75 my (95% HPD 59–92 my) for the stem of Apodanthaceae (Naumann et al., 2013), which agrees well in spite of the completely different taxon sampling and clock model employed here. Unfortunately, the precise sister group of Apodanthaceae in Cucurbitales is not yet known (Schaefer and Renner, 2011).

Based on the inferred ages, the present distribution of Apodanthaceae in Australia, Africa, Iran, and the Americas must have involved at least some transoceanic dispersal. The family's relatively old stem age means that it may have already existed when Africa and South-America had just begun to break apart (100 my ago), but the other divergence times are much too young to allow for overland dispersal to the various parts of the family's range. Whether the range expansion from Africa to West Australia and across the Arabian Peninsula to the Middle East involved a floating infected host (tree trunk) is a matter of speculation. The RLCu and UCLN mean estimates of 6–25 my for the divergence between Mexican and South American species would fit the shoaling of the Panamanian Isthmus. Full closure of the Isthmus, ending communication between Caribbean and Pacific waters, is dated to 3.5 my (Coates et al., 2004; Gutiérrez-García and Vázquez-Domínguez, 2013), but geologic evidence indicates that some mountain ranges of the Darien, such as the San Blas range, were above sea level in the Eocene and until the Miocene when they acted as a peninsula extending from South America northwards (Montes et al., 2012), perhaps facilitating dispersal. *Pilostyles blanchetii*, commonly parasitizing *Bauhinia*, occurs on Jamaica and on the Cayman Islands, both ancient continental fragments, which may indicate that Apodanthaceae can establish on suitable hosts following trans-oceanic dispersal. Nevertheless, their low species number worldwide (just 10) suggests that it is difficult for Apodanthaceae to disperse to new regions and/or that extinction may outweigh speciation, matching findings of unusually high rates of extinction in non-photosynthetic lineages in general (Hardy and Cook, 2012).

Acknowledgments

We thank Martina Silber for help in the lab, Natalie Cusimano for help with the figures, and the curators of the herbaria who provided type images and loans of the material under their care. Funding came from the German Research Foundation (DFG RE 603/16-1).

Table 4
Marginal likelihood estimations performed with BEAST using the stepping-stone sampling method and ln(Bayes factors) of the UCLN, RLCp and RLCu models applied to 18S. (c) Indicates that the topology was constrained, (nc) that it was not constrained.

	Model 1	UCLN (c)	UCLN (nc)	RLCp (c)	RLCp (nc)	RLCu (c)	RLCu (nc)
Model 2	Ln(ml)	−5826.09	−5847.31	−5845.6	−5858.15	−5834.76	−5850.46
UCLN (c)	−5826.09	–	−21.22	−19.50	−32.05	−8.67	−24.37
UCLN (nc)	−5847.31	21.22	–	1.71	−10.84	12.55	−3.15
RLCp (c)	−5845.6	19.5	−1.71	–	−12.55	10.84	−4.86
RLCp (nc)	−5858.15	32.05	10.84	12.55	–	23.39	7.68
RLCu (c)	−5834.76	8.67	−12.55	−10.84	−23.39	–	−15.7
RLCu (nc)	−5850.46	24.37	3.15	4.86	−7.68	15.7	–

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2014.07.005.

References

Aguileta, G., Bielawski, J.P., Yang, Z., 2006. Evolutionary rate variation among vertebrate beta globin genes: implications for dating gene family duplication events. Gene 380, 21–29.

Aris-Brosou, S., 2007. Dating phylogenies with hybrid local molecular clocks. PLoS One 2, e879.

Avaria-Llautureo, J., Hernández, C.E., Boric-Bargetto, D., Canales-Aguirre, C.B., Morales-Pallero, B., Rodríguez-Serrano, E., 2012. Body size evolution in extant Oryzomyini rodents: Cope's rule or miniaturization? PLoS One 7, e34654.

Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.S., Alekseyenko, A.V., 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol. Biol. Evol. 29, 2157–2167.

Barkman, T.J., McNeal, J.R., Lim, S.-H., Coat, G., Croom, H.B., Young, N.D., DePamphilis, C.W., 2007. Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. BMC Evol. Biol. 7, 248.

Bailey, W.J., Fitch, D.H.A., Tagle, D.A., Czelusniak, J., Slightom, J.L., Goodman, M., 1991. Molecular evolution of the psi-eta-globin gene locus – gibbon phylogeny and the hominoid slowdown. Mol. Biol. Evol. 8, 155–184.

Baum, D.A., Smith, S.D., Yen, A., Alverson, W.S., Nyffeler, R., Whitlock, B.A., Oldham, R.L., 2004. Phylogenetic relationships of *Malvatheca* (Bombacoideae and Malvoideae; Malvaceae sensu lato) as inferred from plastid DNA sequences. Am. J. Bot. 91, 1863–1871.

Bell, C.D., Soltis, D.E., Soltis, P.S., 2010. The age and diversification of the angiosperms re-revisited. Am. J. Bot. 97, 1296–1303.

Bellot, S., Renner, S.S., 2013. Pollination and mating systems of Apodanthaceae and the distribution of reproductive traits in parasitic angiosperms. Am. J. Bot. 100, 1083–1094.

Bellot, S., Renner, S.S., 2014. The systematics of the worldwide endoparasite family Apodanthaceae (Cucurbitales), with a key, a map, and color photos of most species. Phytokeys 36, 41–57.

Bromham, L., Cowman, P.F., Lanfear, R., 2013. Parasitic plants have increased rates of molecular evolution across all three genomes. BMC Evol. Biol. 13, 126.

Coates, A.G., Collins, L.S., Aubry, M.-P., Berggren, W.A., 2004. The geology of the Darien, Panama, and the late Miocene-Pliocene collision of the Panama arc with northwestern South America. Geol. Soc. Am. Bull. 116, 1327–1344.

Crisp, M.D., Cook, L.G., 2003. Phylogeny and embryo sac evolution in the endemic Australasian Papilionoid tribes Mirbelieae and Bossiaeeae. In: Klitgaard, B.B., Bruneau, A. (Eds.), Advances in Legume Systematics, Part 10, Higher Level Systematics. Royal Botanic Gardens, Kew, pp. 253–268.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772.

de Vattimo, I., 1971. Contribuçoa ao conhecimento da tribo Apodanthea R. Br. Parte I – Conspecto das espécies (Rafflesiaceae). Rodriguésia 38, 37–62.

Dornburg, A., Brandley, M.C., McGowen, M.R., Near, T.J., 2012. Relaxed clocks and inferences of heterogeneous patterns of nucleotide substitution and divergence time estimates across whales and dolphins (Mammalia: Cetacea). Mol. Biol. Evol. 29, 721–736.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol. 4, e88.

Drummond, A.J., Suchard, M.A., 2010. Bayesian random local clocks, or one rate to rule them all. BMC Biol. 8, 114.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.

Gutiérrez-García, T.A., Vázquez-Domínguez, E., 2013. Consensus between genes and stones in the biogeographic and evolutionary history of Central America. Quater. Res. 79, 311–324.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41, 95–98.

Hardy, N.B., Cook, L.G., 2012. Testing for ecological limitation of diversification: a case study using parasitic plants. Am. Nat. 180, 438–449.

Heibl, C., 2008. PHYLOCH: R Language Tree Plotting Tools and Interfaces to Diverse Phylogenetic Software Packages. <http://www.christophheibl.de/Rpackages.html>.

Herendeen, P.S., Crane, P.R., Drinnan, A.N., 1995. Fagaceous flowers, fruits, and cupules from the Campanian (Late Cretaceous) or central Georgia, USA. Int. J. Plant Sci. 156, 93–116.

Huelsenbeck, J.P., Larget, B., Swofford, D., 2000. A compound Poisson process for relaxing the molecular clock. Genetics 154, 1879–1892.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90, 791.

Katoh, S., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.

Krause, K., 2012. Plastid genomes of parasitic plants: a trail of reductions and losses. In: Bullerwell, C.E. (Ed.), Organelle Genetics. Springer, Heidelberg, Germany, pp. 79–103.

Lavin, M., Herendeen, P.S., Wojciechowski, M.F., 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the Tertiary. Syst. Biol. 54, 530–549.

Li, W.-H., Tanimura, M., 1987. The molecular clock runs more slowly in man than in apes and monkeys. Nature 326, 93–96.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, LA, pp. 1–8.

Montes, C., Cardona, A., McFadden, R., Morón, S.E., Silva, C.A., Restrepo-Moreno, S., Ramírez, D.A., Hoyos, N., Wilson, J., Farris, D., Bayona, G.A., Jaramillo, C.A., Valencia, V., Bryan, J., Flores, J.A., 2012. Evidence for middle eocene and younger land emergence in central Panama: implications for isthmus closure. Geol. Soc. Am. Bull. 124, 780–799.

Naumann, J., Salomo, K., Der, J.P., Wafula, E.K., Bolin, J.F., Maass, E., Frenzke, L., Samain, M.-S., Neinhuis, C., dePamphilis, C.W., Wanke, S., 2013. Single-copy nuclear genes place haustorial Hydnoraceae within Piperales and reveal a Cretaceous origin of multiple parasitic angiosperm lineages. PLoS One 8, e79204.

Nickrent, D.L., Starr, E.M., 1994. High rates of nucleotide substitution in nuclear small-subunit (18S) rDNA from holoparasitic flowering plants. J. Mol. Evol. 39, 62–70.

Nunome, M., Yasudo, S.P., Sato, J.J., Vogel, P., Suzuki, H., 2007. Phylogenetic relationships and divergence times among dormice (Rodentia, Gliridae) based on three nuclear genes. Zool. Scr. 36, 537–546.

R Core Team, 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.

Rambaut, A., Drummond, A.J., 2007. Tracer v1.4. <http://beast.bio.ed.ac.uk/Tracer>.

Rannala, B., Yang, Z., 2007. Inferring speciation times under an episodic molecular clock. Syst. Biol. 56, 453–466.

Rothfels, C.J., Schuettpelz, E., 2014. Accelerated rate of molecular evolution for vittarioid ferns is strong and not driven by selection. Syst. Biol. 63, 31–54.

Sanderson, M.J., 1997. A nonparametric approach to estimating divergence times in the absence of rate constancy. Mol. Biol. Evol. 14, 1218–1231.

Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.

Schaefer, H., Renner, S.S., 2011. Phylogenetic relationships in the order Cucurbitales and a new classification of the gourd family (Cucurbitaceae). Taxon 60, 122–138.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Thiele, K.R., Wylie, S.J., Maccarone, L., Hollick, P., McComb, J.A., 2008. *Pilostyles coccoidea* (Apodanthaceae), a new species from Western Australia described from morphological and molecular evidence. Nuytsia 18, 273–284.

Thorne, J.L., Kishino, H., Painter, I.S., 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15, 1647–1657.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3 – new capabilities and interfaces. Nucl. Acids Res. 40, e115.

VanDerWal, J., Falconi, L., Januchowski, S., Shoo, L., Storlie, C., 2012. SDMTools: species distribution modelling tools: tools for processing data associated with species distribution modeling exercises. R Package Version 1.1-13. <http://CRAN.R-project.org/package=SDMTools>.

Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., Davis, C.C., Latvis, M., Manchester, S.R., Soltis, D.E., 2009. Rosid radiation and the rapid

rise of angiosperm-dominated forests. Proc. Natl. Acad. Sci. USA 106, 3853–3858.
Wicke, S., Müller, K.F., de Pamphilis, C.W., Quandt, D., Wickett, N.J., Zhang, Y., Renner, S.S., Schneeweiss, G.M., 2013. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. Plant Cell 25, 3711–3725.
Yoder, A.D., Yang, Z., 2000. Estimation of primate speciation dates using local molecular clocks. Mol. Biol. Evol. 17, 1081–1090.

Table S1. List of species used in this study with author names, herbarium vouchers, and GenBank accession numbers for all sequences.

Species name (before synonymizations)	Accepted species name	Voucher (Herbarium) 18S; *matR*	Source 18S; *matR*	GenBank accession 18S	GenBank accession *matR*
Anisophylleaceae (Cucurbitales)					
Anisophyllea fallax Scott-Elliot	*Anisophyllea fallax* Scott-Elliot	Clausing s.n. (MO)	Clement et al., 2004	AF534775	-
Anisophyllea sp.	*Anisophyllea sp.*	P. Boyce 758 (K)	Qiu et al., 2010	-	GU351150
Apodanthaceae (Cucurbitales)					
Apodanthes casearіae Poiteau	*Apodanthes casearіae* Poiteau	C. Galdames 7260 (M)	This paper	KJ634128	KJ634100
Apodanthes tribracteata Rusby	*Apodanthes casearіae* Poiteau	D. N. Smith et al. 14057 (MO)	This paper	KJ634127	KJ634101
Pilostyles aethiopica Welwitsch	*Pilostyles aethiopica* Welwitsch	S. Bellot 29 (M)	This paper	KJ634129	-
Pilostyles berteroi Guill.	*Pilostyles berteroi* Guill.	C. Heibl 05-017 (M)	This paper	KJ634126	KJ634102
Pilostyles blanchetii (Gardner) Brown	*Pilostyles blanchetii* (Gardner) Brown	R. L. Dressler 3380 (MO)	This paper	-	KJ642224
Pilostyles cf. *calliandrae* (Gardner) Brown	*Pilostyles blanchetii* (Gardner) Brown	E. P. Heringer et al. 5768 (MO)	This paper	KJ634125	KJ634103
Pilostyles coccoidea K.R.Thiele	*Pilostyles coccoidea* K.R.Thiele	K. R. Thiele 3495 (PERTH)	This paper	KJ634124	KJ634104
Pilostyles collina B.Dell	*Pilostyles collina* B.Dell	K. R. Thiele 4501 (PERTH)	This paper	KJ634123	KJ634105
Pilostyles goyazensis Ule	*Pilostyles blanchetii* (Gardner) Brown	H. S. Irwin et al. 14880 (MO)	This paper	KJ634122	KJ634106
Pilostyles hamiltonii C.A.Gardner	*Pilostyles hamiltonii* C.A.Gardner	D. Dixon 1039 (PERTH)	This paper	KJ634121	KJ634107
Pilostyles haussknechtii Boissier	*Pilostyles haussknechtii* Boissier	A. Chehregani and S. Zarre 17834 (M)	This paper	KJ634120	-
Pilostyles ingae (Karsten) Hooker	*Pilostyles blanchetii* (Gardner) Brown	A. Glaziou 22029 (G)	This paper	KJ634119	KJ634108
Pilostyles mexicana (Brandegee) Rose	*Pilostyles mexicana* (Brandegee) Rose	D. E. Breedlove 27233 (NY)	This paper	KJ634118	KJ634109
Pilostyles mortoni Vattimo unpublished	*Pilostyles thurberi* Gray	F. Lyle Wynd and C. H. Mueller 256 (NY)	This paper	KJ634117	KJ634110
Pilostyles thurberi Gray	*Pilostyles thurberi* Gray	M. C. Johnston et al. 11286 (MO)	This paper	KJ634116	KJ634111
Pilostyles thurberi Gray	*Pilostyles thurberi* Gray	D. L. Nickrent 4342 (SIU)	Nickrent et al., 2004	-	AY739003
Pilostyles ulei Solms	*Pilostyles blanchetii* (Gardner) Brown	G. T. Ceccantini et al. 3593 (SPF)	This paper	KJ634114	KJ634113
Pilostyles ulei Solms	*Pilostyles blanchetii* (Gardner) Brown	M. Nee et al. 48695 (NY)	This paper	KJ634115	KJ634112

Begoniaceae (Cucurbitales)					
Begonia rex Putz.	*Begonia rex* Putz.	S. S. Renner 2815 (M)	Filipowicz and Renner, 2010	-	HM600770
Betulaceae (Fagales)					
Alnus sinuata Rydb.	*Alnus sinuata* Rydb.	R.-Q. Li 20002 357-97B (AA)	Li et al., 2004	AY263900	AY263907
Betula pendula Roth.	*Betula pendula* Roth.	[1]Unvouchered; [2]T. J. Barkman 359 (PAC)	[1]Gillman et al., 2010; [2]Barkman et al., 2004	GU476453[1]	AY453121[2]
Clusiaceae (Malpighiales)					
Clusia rosea Jacq.	*Clusia rosea* Jacq.	[1]T. Tokuoka 272 (KYO); [2]Y.-L. Qiu 05042 (MICH)	[1]Tokuoka and Tobe, 2006; [2]Qiu et al., 2010	AB233537[1]	GU351186[2]
Corynocarpaceae (Cucurbitales)					
Corynocarpus laevigatus J.R.Forst & G.Forst	*Corynocarpus laevigatus* J.R.Forst & G.Forst	[1]M. W. Chase 236 (NCU); [2]M. W. Chase s.n. (NCU)	[1]Soltis et al., 2000; [2]Zhu et al., 2007	AF206892[1]	AY121499[2]
Cucurbitaceae (Cucurbitales)					
Coccinia adoensis (Hochst. ex A.Rich.) Cogn.	*Coccinia adoensis* (Hochst. ex A.Rich.) Cogn.	M. Wilkins 367 (unvouchered)	Clement et al., 2004	AF534781	-
Cucumis sativus L.	*Cucumis sativus* L.	Soltis s.n. (WS)	Soltis et al., 2000	AF206894	-
Cucurbita pepo L.	*Cucurbita pepo* L.	[1]Soltis s.n. (WS); [2]T. J. Barkman 371 (PAC)	[1]Soltis et al., 2000; [2]Barkman et al., 2004	AF206895[1]	AY453101[2]
Datiscaceae (Cucurbitales)					
Datisca cannabina L.	*Datisca cannabina* L.	Y.-L. Qiu 97102 (IND)	Qiu et al., 2010	-	GU351197
Datisca glomerata (C.Presl) Baill.	*Datisca glomerata* (C.Presl) Baill.	A. Liston 767 (RSA)	Liston et al., 1992; Swensen et al., 1998	DGU42426	-
Fagaceae (Fagales)					
Fagus grandifolia L.	*Fagus grandifolia* L.	[1]Soltis 2521 (WS); [2]T. J. Barkman 383 (PAC)	[1]Soltis et al., 2000; [2]Barkman et al., 2004	AF206910[1]	AY453092[2]
Juglandaceae (Fagales)					
Juglans nigra L.	*Juglans nigra* L.	[1]Soltis 2520 (WS); [2]unvouchered	[1]Soltis et al., 2000; [2]Zhu et al., 2007	AF206943[1]	AF520073[2]
Nothofagaceae (Fagales)					
Nothofagus antarctica Oerst.	*Nothofagus antarctica* Oerst.	Z.-D. Chen 99028 (Z)	Li et al., 2004	AY147111	-
Nothofagus moorei (F.Muell.) Krasser	*Nothofagus moorei* (F.Muell.) Krasser	Y.-L. Qiu 98036 (Z)	Qiu et al., 2006	-	DQ401401

Barkman, T.J., Lim, S.H., Salleh, K.M., Nais,J., 2004. Mitochondrial DNA sequences reveal the photosynthetic relatives of *Rafflesia*, the world's largest flower. Proc. Natl. Acad. Sci. U.S.A. 101, 787-792.

Clement, W.L., Tebbitt, M.C., Forrest, L.L., Blair, J.E., Brouillet, L., Eriksson, T., Swensen, S.M., 2004. Phylogenetic position and biogeography of *Hillebrandia sandwicensis* (Begoniaceae): a rare Hawaiian relict. American Journal of Botany 91, 905–917.

Dong, T.T., Ma, X.Q., Clarke, C., Song, Z.H., Ji, Z.N., Lo, C.K., Tsim, K.W., 2003. Phylogeny of *Astragalus* in China: molecular evidence from the DNA sequences of 5S rRNA spacer, ITS, and 18S rRNA. J. Agric. Food Chem. 51, 6709-6714.

Gillman, L.N., Keeling, D.J., Gardner, R.C., Wright, S.D., 2010. Faster evolution of highly conserved DNA in tropical plants. Journal of Evolutionary Biology 23, 1327-1330.

Li, R.-Q., Chen, Z.-D., Lu, A.-M., Soltis, D.E., Soltis, P.S., Manos, P.S., 2004. Phylogenetic relationships in Fagales based on DNA sequences from the three genomes. Int. J. Plant Sci. 165, 311-324.

Liston, A.L., Rieseberg, L.H., Hanson, M.A., 1992. Geographic partitioning of chloroplast DNA variation in the genus *Datisca* (Datiscaceae). Plant Systematics and Evolution 181, 121-132.

Martin, P.G., Dowd, J.M., 1993. Partial sequence of ribosomal RNA of Papilionaceae and related families. Phytochemistry 33, 361-363.

Nickrent, D.L., Blarer, A., Qiu, Y.L., Vidal-Russell, R., Anderson, F.E., 2004. Phylogenetic inference in Rafflesiales: the influence of rate heterogeneity and horizontal gene transfer. BMC Evol. Biol. 4, 40.

Pinto, L.S., Andrade-Neto, M., Cunha, R.M.S., Cavada, B.S., Alves, M.A.O., Grangeiro,T.B. Submitted (27-JUN-2002) Biology, UFC, Av. Mister Hull, sn., Fortaleza, CE 60.451-970, Brazil

Qiu, Y.-L., Li, L., Wang, B., Xue, J.-Y., Hendry, T.A., Li, R.-Q., Taylor, D.W., Issa, M.J., Ronen, A.J., Vekaria, M.L., White, A.M., 2006. Reconstructing the basal angiosperm phylogeny: evaluating information content of mitochondrial genes. Taxon 55, 837-856.

Qiu, Y.-L., Li, L., Wang, B., Xue, J.-Y., Hendry, T.A., Li, R.-Q., Brown, J.W., Liu, Y., Hudson, G.T., Chen, Z.-D., 2010. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. Journal of Systematics and Evolution 48, 391-425.

Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, W.J., Nixon, K.C., Farris, J.S., 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. Botanical Journal of the Linnean Society 133, 381–461.

Sugiyama, Y., Watase, Y., Nagase, M., Makita, N., Yagura, S., Hirai, A., Sugiura,M.,2005. The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. Mol. Genet. Genomics 272, 603-615.

Swensen, S.M., Luthi, J.N., Rieseberg, L.H., 1998. Datiscaceae revisited: monophyly and the sequence of breeding system evolution. Systernntic Botany 23, 157-169.

Thiele, K.R., Wylie, S.J., Maccarone, L., Hollick, P., McComb, J.A., 2008. *Pilostyles coccoidea* (Apodanthaceae), a new species from Western Australia described from morphological and molecular evidence. Nuytsia 18, 273–284.

Tokuoka,T., Tobe, H., 2006. Phylogenetic analyses of Malpighiales using plastid and nuclear DNA sequences, with particular reference to the embryology of Euphorbiaceae sens. str. J. Plant Res. 119, 599-616.

Zhu, X.-Y., Chase, M.W., Qiu, Y.-L., Kong, H.-Z., Dilcher, D.L., Li, J.-H., Chen, Z.-D., 2007. Mitochondrial matR sequences help to resolve deep phylogenetic relationships in rosids. BMC Evolutionary Biology 7, 217.

Table S2. Primers used for amplification of *matR* and 18S.

Primer name	Sequence (5'>3')	Source
matR_F2	GCACCGTATCCATATAACTGC	Barkman et al., 2004
matR_F3	AGTGGGGAACCTTCTCGCTGG	This paper
matR_F4	CCTTCCTCATAGAAGCCACCGGG	This paper
matR_F5'	GTTTTCACACCATCGACCGACATCG	Barkman et al., 2004
matR_F5b	AGCTGAAGAAGACGAAAGGGCTG	This paper
matR_F6	ACAAGTACCCCACGGCTTCG	This paper
matR_F7	TCGTCGTATTGATGACCAAGA	This paper
matR_F8	TTTATTACGCGCGATATGCC	This paper
matR_F9	TATAGCTCATGTCGTGGTCG	This paper
matR_R1	GCAGTTATATGGATACGGTGC	Barkman et al., 2004
matR_R2	AAAGAAGGCTCGAGGGCTTG	Barkman et al., 2004
matR_R3'	CGCGGCACCTGTAGTAGGACAGAGGA	Barkman et al., 2004
matR_R3b	CGTTCGTCAAGCAGGCCACG	This paper
matR_R4	GCGGAGCGTAAGTGGCAAGC	This paper
matR_R5	AAACCTCCCCCGAAGCCGTG	This paper
matR_R6	CTTCAGCTTTGGGATATTCTGACC	This paper
matR_R7	GGAATGAAAGGCCTCTTTGC	This paper
matR_R8	CCAACGATTCCCCTAGACA	This paper
matR_R9	AATGACAACTTCCGGACACT	This paper
18S_25ef	CTGGTTGATCCTGCCAG	Nickrent and Starr, 1994
18S_1769R	CACCTACGGAAACCTTGTT	Nickrent and Starr, 1994
18S_626R	TCCAACTACGAGCTT	Nickrent and Starr, 1994
18S_1131R	CAATTCCTTTAAGTTTCAGCC	Nickrent and Starr, 1994
18S_530F	GTGCCAGC(AC)GCCGCGG	Nickrent and Starr, 1994
18S_1322F	TAACGAACGAGACCTCAGCCT	Nickrent and Starr, 1994
18S_1433R	ATCTAAGGGCAT(GC)ACAGACC	Nickrent and Starr, 1994
18S_922F	GAAACTTAAA(GT)GAATTG	Nickrent and Starr, 1994
18S_F1	TCGATGGTAGGATAGTGGCC	This paper
18S_R1	CTCCTTGGATGTGGTAGCCG	This paper
18S_F2	TCAGCATGGGATAACGACGC	This paper
18S_R2	CCTCTGACTGCGGAATACGG	This paper

Barkman, T.J., Lim, S.H., Salleh, K.M., Nais,J., 2004. Mitochondrial DNA sequences reveal the photosynthetic relatives of *Rafflesia*, the world's largest flower. P. Natl. Acad. Sci. USA 101, 787–792.

Nickrent, D.L., Starr, E.M., 1994. High rates of nucleotide substitution in nuclear small-subunit (18S) rDNA from holoparasitic flowering plants. J. Mol. Evol. 39, 62–70.

Table S3. Branch lengths and calculation of a local rate of substitution for matR and 18S in Apodanthaceae, using the branch lengths of the phylogenies in Fig. S1 and the approach described in Fig. 1B and in Section 2. bl1 refers to the branch between Fagales and Cucurbitales (see Fig. 1B), RC is the rate of Cucurbitales, and RA that of Apodanthaceae as in Fig. 1B, in subst./site/my. T2 is the divergence time between apodanths and other cucurbits as shown on Fig. 1B.

	matR UCLN	*matR* RLC	mean *matR* UCLN, RLC	18S UCLN	18S RLC	Mean 18S UCLN, RLC
Mean non-parasitic cucurbits	0.036	0.032	0.034	0.011	0.013	0.012
Stem of apodanths	0.098	0.100	0.099	0.063	0.063	0.063
Crown of apodanths	0.074	0.079	0.077	0.066	0.069	0.068
Apodanths	0.172	0.178	0.175	0.128	0.132	0.130
Ratio apodanths/ other cucurbits	5	6	5	12	10	11
bl_1	0.0071	0.0066	0.0069	0.0031	0.0040	0.0036
R_C	0.0005	0.0004	0.0004	0.0001	0.0002	0.0002
T_2	79	79	79	73	73	73
R_A	0.0022	0.0023	0.0022	0.0018	0.0018	0.0018
R_A using T_2 of 18S for *matR* and T_2 of *matR* for 18S			0.0024			0.0016
Rate used to calibrate the local clock			0.0023			0.0017

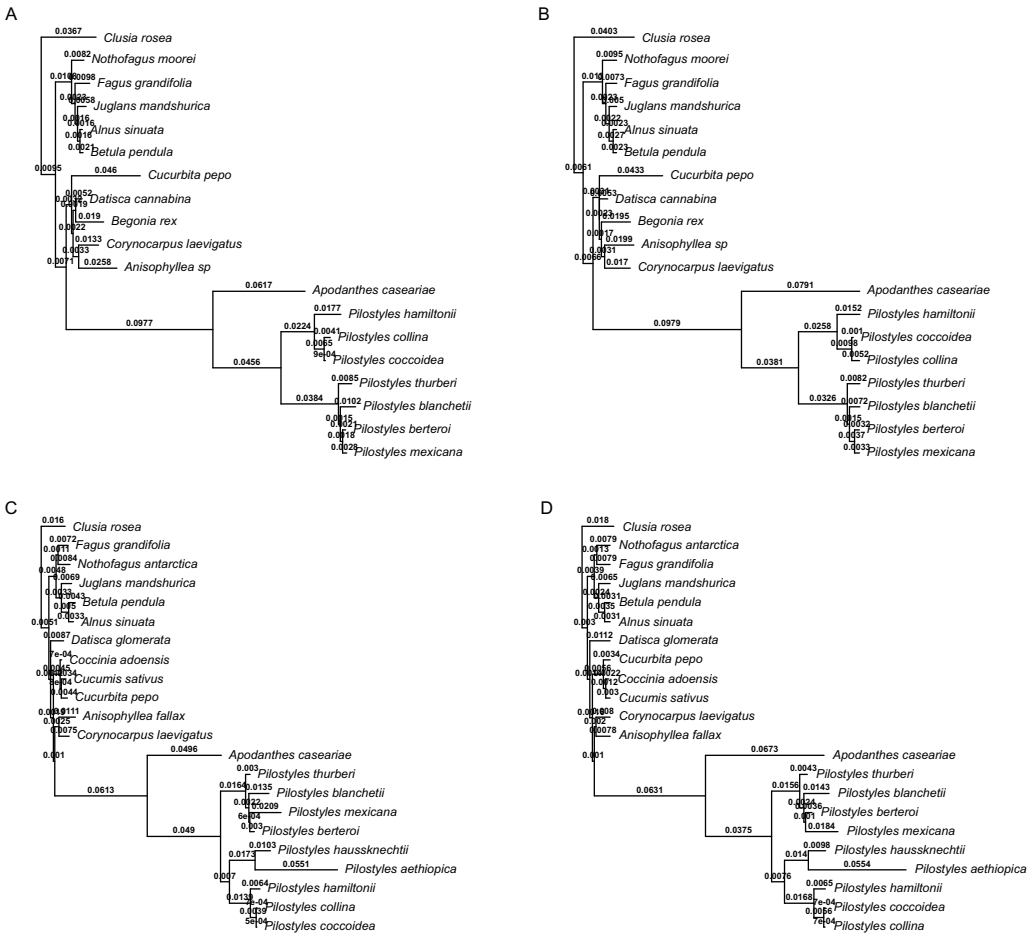

Figure S1. Branch lengths in substitutions per site inferred by Bayesian inference with BEAST (Drummond et al., 2006 and Drummond et al., 2012) for *matR* (A and B) and 18S (C and D) using the UCLN (A and C) and the RLC (B and D) approaches.

Chapter 5

INTRACELLULAR AND HORIZONTAL TRANSFER OF PLASTID GENES TO THE CHONDRIOME – ENDOPARASITISM AS THE PRECONDITION OF PLASTOME LOSS

Bellot, S. and S. S. Renner

Molecular Biology and Evolution (in review)

Intracellular and horizontal transfer of plastid genes to the chondriome – endoparasitism as the precondition of plastome loss

Sidonie Bellot and Susanne S. Renner

Department of Biology, Ludwig Maximilian University, Menzinger Straße 67, D-80638 Munich, Germany

Corresponding Author: Sidonie Bellot
Email: sido.bellot@gmail.com

Running Title: Plastome loss in endoparasites

Keywords:

ABSTRACT Analyses of the plastomes of 20 non-photosynthetic land plants from 13 genera and five families have revealed that they retain a minimal set of 22 genes. All 20 are exoparasites or mycotrophs, meaning they connect to their host plants from the outside via special organs (haustoria) or fungi. A second form of parasitism, called endoparasitism, involves plants that exist as clusters of parenchyma cells embedded inside the host, emerging only to flower. Endoparasitism has evolved in four families, including Rafflesiaceae and Apodanthaceae. The only endoparasite studied, *Rafflesia lagascae*, has completely lost a functional plastome. We analyzed the plastomes of two Apodanthaceae, the African *Pilostyles aethiopica* and the Australian *P. hamiltonii*, both parasitizing Fabaceae, to determine patterns of horizontal and intra-cellular gene transfer or loss. A clock-dated phylogeny for the ten species of Apodanthaceae shows that the two species belong to clades that separated 23-33 my ago. Using information from gene flanking regions, we found that 54 out of 82 plastid-like regions from P. *hamiltonii* are located in its mitochondrial genome; for *P. aethiopica*, this ratio is 45 out of 59. The remaining plastid-like contigs lacked flanking regions, but based on coverage are also located in the mitochondrial or nuclear genomes. In addition, 20 (24%) of the plastid-like fragments of *P. hamiltonii* and 16 (27%) of *P. aethiopica* were more similar to Fabales than to other lineages, suggesting horizontal gene transfer. Taken all together, our results suggest that endoparasitism may completely remove

selection for plastid-encoded functions and may be the precondition for the complete loss of a plastome.

INTRODUCTION

The loss of photosynthesis has occurred in parasitic lineages ranging from algae to angiosperms, but is not necessarily accompanied by a loss of plastids (Krause, 2012). Which genes may persist in no longer photosynthetically active plastids has been studied in 23 species from 13 genera of Cuscutaceae, Orobanchaceae, Rafflesiaceae, and myco-heterotrophic Orchidaceae and Petrosaviaceae (Wolfe et al., 1992; Funk et al., 2007; McNeal et al., 2007; Delannoy et al., 2011; Logacheva et al., 2011; Barrett and Davis, 2012; Wicke et al., 2013, Barrett et al., 2014; Logacheva et al., 2014; Molina et al., 2014). A model of directional gene loss has been inferred for myco-heterotrophic Orchidaceae, based on ten species of known evolutionary relationships in which certain gene-class categories, such as the *ndh* genes, are lost earlier than other categories, such as the photosystem genes (Barrett et al., 2014). Another study of a clade of parasites, namely nine holoparasitic Orobanchaceae, also found early loss of *ndh* genes, but otherwise no directional loss (Wicke et al., 2013). Instead, the individual parasite species varied greatly in the numbers and kinds of lost or retained genes.

Pooling the data from the 23 non-photosynthetic plastomes analyzed, shows that 22 genes are retained in most of them (ten ribosomal proteins, four ribosomal RNAs, seven transfer RNAs, and the gene *ycf*2). The exception is *Rafflesia lagascae,* which lacks functional copies of all 22 genes (Molina et al., 2014). This species is the only endoparasite investigated so far and completely lacks a plastid genome (Molina et al., 2014). All the other parasites analyzed are exo-parasites, meaning they connect to the host via haustoria from the outside, rather than permanently living inside the host as a network of parenchyma cells as do endoparasites (Heide-Jorgensen, 2008). Endoparasitism has evolved four times, namely in Rafflesiaceae (34 species in three genera), Cytinaceae (ca. 10 species in two genera), Mitrastemonaceae (one or two species in one genus), and Apodanthaceae (10 species in two genera, *Apodanthes* and *Pilostyles*).

Based on NGS coverage ratios, the plastome-like fragments of *Rafflesia lagascae* appear located in the nuclear or mitochondrial genomes (Molina et al., 2014), and most of them are phylogenetically closer to the Vitales genus *Tetrastigma*, the exclusive host of all species of *Rafflesia*, than to *Rafflesia lagascae*, pointing to horizontal transfers. The chondriome of the related *R. cantleyi* also contains a large amount of host (presumably

Tetrastigma) DNA (Xi et al., 2012, 2013); the plastome of this species has not been investigated. Other parasitic flowering plants have also exchanged mitochondrial genetic material with their hosts, judging from analyses of short stretches of DNA (Renner and Bellot, 2012 for a review), and large amounts of macromolecules, including mRNA, have been shown to pass through plasmodesmata between the parasite *Cuscuta* and two experimental hosts, *Arabidopsis thaliana* and *Solanum lycopersium* (Kim et al., 2014). Other experiments have demonstrated that entire organelle and even nuclear genomes can be transferred between distantly related angiosperms when their cells are in direct and prolonged contact (Stegemann and Bock, 2009; Stegemann et al., 2012; Fuentes et al., 2014). These findings are relevant to studies of parasite plastomes because the occurrence of both intra-cellular and horizontal gene transfer (IGT and HGT) complicates detecting any autochthonous parasite plastome.

Apodanthaceae occur in North and South America, Africa, Iran, and Australia, and belong in the Cucurbitales (Filipowicz and Renner, 2010; Bellot and Renner, 2014a, b). They are endoparasites, living as cell clusters inside Fabaceae and Salicaceae, and emerging only in the form of tiny (few-millimeter large) flowers that break through the host's bark once a year (color photos in Bellot and Renner, 2013). Apodanthaceae completely lack any green parts, and no chloroplasts have ever been observed in their tissues (Rutherford, 1970: *P. thurberi*; Dell et al., 1982; *P. hamiltonii*; personal observation by SB on living Apodanthaceae in Australia, Africa, and Iran, and light microscopy for *P. aethiopica*). We chose this lineage to compare organelle evolution in endoparasites with that in exo-parasites.

We had two expectations about endoparasite organellar genomes: (i) Complete degeneration of the plastome and physical but not functional transfer of its genes to one of the other two plant genomic compartments and (ii) horizontal gene transfer between parasite and host, involving the parasites' plastomes (we do not focus on the nuclear genome for lack of a suitable reference genome). Different from the study of the only other endoparasite so far analyzed, *Rafflesia lagascae*, we use the flanking regions of plastid-like genes to determine their genomic location. We analyzed two not closely related species, the African *Pilostyles aethiopica* and the Australian *P. hamiltonii,* both parasitizing Fabaceae, to compare horizontal and intra-cellular gene transfer or loss. A clock-dated phylogeny for the ten species of Apodanthaceae shows that the two species belong to clades that separated 23-33 my ago (Bellot and Renner, 2014a).

MATERIALS AND METHODS

Taxon sampling, DNA extractions and sequencing

Flower tissue from a female individual of *Pilostyles aethiopica* (voucher S. Bellot 29, deposited in the herbarium of Munich) was collected in the Mukuvisi woodlands of Harare, Zimbabwe on 29 February 2012 and kept frozen until DNA isolation with the kit DNeasy Plant Maxi Kit (Qiagen), following the manufacturer's protocol. 1 mL of the DNA (17.2 ng/µL) was sent to Eurofins MWG Operon for precipitation and sequencing. One genomic shotgun library of insert sizes 160 to 310 bp was sequenced in one channel of Illumina HiSeq 2000, yielding 236,404,172 reads of 101 bp. DNA from silica-dried flowers of male and female *Pilostyles hamiltonii* collected in Perth in October 2010 (voucher K. Dixon 1039 in the herbarium of Perth) was isolated using the same approach, and the DNA was sent to the University of Vienna (C. Schlötterer's lab) for Illumina sequencing on a Genome Analyzer IIx platform. This yielded 80,223,076 paired-end reads of 101 bp.

Quality control and pre-processing of the reads

For *P. aethiopica*, following a quality control of the reads with PRINSEQ (Schmieder and Edwards, 2011) and FASTQC (Andrews, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), adaptors were removed when necessary using fastx-toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), and the sequences were trimmed at both ends using PRINSEQ to remove polyA/T tails >5 nucleotides and all bases with a quality score <20, stopping at the first base with a quality >20. The few sequences <60 bp were then removed as were sequences of a mean quality score <30, or >1% of Ns, or an entropy <70. This left 146,860,644 reads.

Adaptors were removed from the reads of *P. hamiltonii* using Cutadapt (Martin, 2011), and the reads were then again filtered and trimmed using PRINSEQ with similar stringency thresholds as for *P. aethiopica*, which left 79,054,264 reads. The paired-end information for both species was retained.

Sorting and assembling reads

The reads of *P. aethiopica* and *P. hamiltonii* were mapped against the plastome of *Cucumis sativus* (Table S1 lists all plastid genomes used in this study), and for each protein-coding region, rRNA, or tRNA, the corresponding reads were extracted up to the first gap in the mapping. When a gap was present, two separate extractions of the reads on either side of the gap were performed. Where mapping continued beyond coding regions, all reads were

kept. After collecting the reads, those corresponding to one region were *de novo* assembled with Geneious R7 (Biomatters, http://www.geneious.com/) to identify possible multiple contigs mapping to the same region. Every contig was checked for ambiguities, and consensuses of 3, 20, 40 or 80 reads (depending on the length and the total number of contigs for a region) were retained to avoid losing plastid-like nuclear regions or a plastid genome that might be present in low copy number. In many cases, we could identify multiple contigs corresponding to a given gene. All retained contigs were then aligned against the plastome of *C. sativus*, keeping only contigs with an e-value <0.00001. This reduced the number of contigs from 3755 to 313 for *P. aethiopica* and from 873 to 215 for *P. hamiltonii*.

In parallel, *de novo* assemblies of the clean reads of *P. aethiopica* and *P. hamiltonii* were performed on the CLC Genomics Workbench **7 (**http://www.clcbio.com) and improved with SSPACE (Boetzer et al., 2011), which remaps reads using the included Bowtie assembler and scaffolds the resulting contigs using the paired-end information. This produced 952,874 contigs for *P. aethiopica* and 270,940 for *P. hamiltonii*, with N50 of 601 and 446 bp, and a maximum contig size of 61,798 and 33,472 bp. All plastid-like contigs were mapped against the *de novo* contigs using CLC, and their flanking regions were retrieved using custom scripts and the tool suites BEDtools (Quinlan and Hall, 2009) and SAMtools (Li et al., 2009). Coverage information was retrieved for the plastid-like contigs as well as for their flanking regions.

Genomic and phylogenetic origin of plastid-like DNA fragments

The contigs derived from the mapping on the *Cucumis* plastome were submitted to two analyses to identify their genomic and phylogenetic origins. First, for both species of *Pilostyles*, each contig was aligned against the NCBI nucleotide database using the BLASTn command of BLAST+ version 29 (Camacho et al., 2008; ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/2.2.28/), with the maximum number of target sequences set to 5000. A script (http://seqanswers.com/forums/showthread.php?t=40975) was used to retrieve the taxonomic description of every subject sequence, and custom Python scripts were written to record sequence bitscores, genomic compartments, and taxonomic assignment to Cucurbitales, Fagales, Rosales, Fabales, and other orders. Flanking regions of the contigs were submitted to the same analysis.

For *P. aethiopica*, a custom script was used to retrieve the homologous sequence of each contig in 16 published rosid plastomes and in the plastome of the malvid (or rosid *sensu*

lato) *Gossypium anomalum*. This produced 313 plastid alignments of which some were discarded because they were too short or harbored only identical sequences, resulting in 231 alignments containing a variable number of species. Alignments were generated with MAFFT version 7.017 (Katoh et al., 2002). Maximum likelihood phylogenetic analyses were conducted in RAxML 7.2.8 (Stamatakis, 2006) with the GTR + G substitution model and 1000 bootstrap replicates, all through Geneious R7. The placement of the *Pilostyles* sequences in the resulting phylogenies was recorded, along with bootstrap support.

To test the reliability of BLAST-based inferences of phylogenetic relationships (rather than tree-based ones), we retrieved representative nuclear 18S and mitochondrial *matR* and *nad1 Pilostyles* sequences from GenBank and submitted them to the same BLAST analyses as described above. We also split sequences in the middle to see how sequence length would influence results. The tabulation of all bitscores, genomic compartments, and taxonomic assignment is available upon request. The true genomic compartment (nuclear or mitochondrial) was always recovered, while the true taxonomic assignment (Cucurbitales) was never the first hit, although it was always among the top 5000 hits. We counted all cases where the match to Fabales (the host) was better than that to Cucurbitales. The existence of such cases in this test indicates that BLASTing can result in erroneously inferred horizontal gene transfer from Fabales to *Pilostyles*. To guard against false positives and false negatives when BLASTing our *Pilostyles* plastome-like contigs (with positive and negative referring to the inference of HGT from Fabales to *Pilostyles*), we decided to develop thresholds for trustworthy BLAST results, given currently available data in GenBank (Oct. 2014). To do this, we selected homologous regions to the *P. aethiopica* plastid-like contigs in the genome of the Fabaceae *Cicer arietinum* (source see Table S1) and submitted them to the same BLAST analyses as above to test how often they would match other Fabales or other rosid orders. We did the same with homologous *Cucumis* (Cucurbitaceae) regions, to quantify how often they would BLAST to Fabales instead of Cucurbitales (the order to which *Pilostyles* belongs). The bitscores of the different hits were used to establish the thresholds for false negatives and false positives. When the *Cicer* plastid regions were BLASTed, 75% had Fabales as their first match (once *Cicer* sequences had been removed from the hits). For *Cicer* sequences for which the first hit was not to Fabales (i.e., false negatives), the average difference between their bitscore and that to Fabales was 1.6, with a maximum of 11. When the *Cucumis* plastid sequences homologous to *Pilostyles* plastid-like fragments were BLASTed, 70% had their first hit to Cucurbitales and 29% to orders other than Cucurbitales or Fabales. Only two sequences (0.7%) had their best bitscores to Fabales or Cucurbitales.

They were considered false positives. These experiments suggest that if a contig of *Pilostyles* is BLASTed and (i) its first hit is to Fabales with no bitscore difference to Cucurbitales, HGT can be rejected, whereas if the difference is >0, HGT can be accepted, and (ii) if a contig's first hit is not to Fabales and the bitscore difference is <11, *absence* of HGT can be rejected, whereas if it is >11, *absence* of HGT can be accepted.

RESULTS

Plastid-like regions located in the mitochondrial genome

We found 59 plastid-like contigs in *P. aethiopica* and 82 in *P. hamiltonii* (Table 1). Because we focused on coding regions (usually more conserved than intergenic spacers), and because the BLAST analyses always recovered the correct genomic compartment during our tests (see *Material and Methods*), we are confident that those sequences indeed are originally coming from a plastid genome. The longest plastid-like fragment found in *P. hamiltonii* was 5,202 bp long and matched plastid reference sequences along its entire length. In *P. aethiopica*, the longest fragment was 4,151 bp long (Table 1). When considering only parts matching with plastid references (the rest not matching with anything), the longest contig in this species was 2,582 bp long.

As inferred from BLASTing, 79 of 89 flanking regions of the plastid-like contigs from P. *hamiltonii* are located in the mitochondrial genome; for *P. aethiopica*, the equivalent numbers are 72 out of 76 (Table S2). Ten flanking regions from *P. hamiltonii* and four from *P. aethiopica* failed to match anything.

When both flanking regions were mitochondrial-like, we considered the inserted plastid-contig as located in the mitochondrial genome. This was the case for 43 plastid-like contigs from *P. hamiltonii* and 41 from *P. aethiopica* (Tables 1 and S2). When one flanking region was mitochondrial and the other was from an unknown genomic location, we considered the plastid-contig as located in the mitochondrial genome if its coverage was >10, which was the case for 11 contigs from *P. hamiltonii* and 4 from *P. aethiopica* (Tables 1 and S2). The mean read coverage of the plastid-like contigs located in the mitochondrial genome of *P. aethiopica* was 243 (Table 2), with a minimum of 117 and a maximum of 394 (Table 1). For *P. hamiltonii*, the mean coverage was 115 (Table 2), with a minimum of 2 and a maximum of 280 (Table 1).

To decide where the plastid-like contigs that failed to match anything were located, we compared their coverage to that of the contigs with known flanking regions. In *P. aethiopica*, the singletons had a coverage of <10 except for one, which had a coverage of

302, similar to contigs with mitochondrial flanking regions (Table 1). In *P. hamiltonii*, the coverage of 12 of the 28 singletons was ≤10, for 13 it was between 10 and 100, and for 3 it was >100, with the highest coverage being 381. *De novo* assemblies of the singletons did not concatenate them into larger contigs.

Horizontal gene transfer from Fabales to Cucurbitales

Using our experimentally developed thresholds for false positives and negatives (*Materials and Methods*), we determined the phylogenetic affinity of each plastid-like contig and its flanking region. We discarded plastid-like singletons that matched only with Fabales and that might represent contamination of parasite DNA with host DNA. Table S3 shows the results from the phylogenetic approach and the BLAST approach for all plastid-like contigs, including those discarded as possible contamination. The accepted topology for angiosperm families and orders was recovered with 21 of the 231 contig alignments. Conflict between the two approaches was limited to a few cases. In the nine cases in which the phylogeny had >70% ML bootstrap support for the conflicting placement, we scored the contig's phylogenetic origin as unknown. For unsupported conflicts between the phylogenetic and the BLAST approach, we accepted the BLAST result as correct. The phylogenetic affinities of all alignments are provided in Table S3 and summarized in Table 1.

Overall, 20 (24.4%) of the plastid-like fragments of *P. hamiltonii* and 16 (27%) of *P. aethiopica* were more similar to Fabales than to other lineages. In *P. aethiopica*, four contigs (7%) were chimeras, with some parts most similar to Fabales, others to Cucurbitales (1 fragment) or other rosids (3 fragments). Of the 73% of contigs from *P. aethiopica* that did not match Fabales, 25 (42%) were more similar to other rosids including one with its first hit to Cucurbitales, and 18 (31%) with first hits to orders other than rosids. Of the 75.6% of contigs from *P. hamiltonii* that did not match Fabales, 29 (35.4%) were more similar to other rosids including two with their first hit to Cucurbitales, whereas 33 (40.2%) matched an order other than rosids. Figure 1 summarizes the proportions of plastid-like contigs per genomic location and phylogenetic affinity.

Plastid genes of *P. aethiopica* and *P. hamiltonii* appear non-functional

Table S4 lists the plastid genes, tRNAs, and rRNAs found or absent in *P. aethiopica* and *P. hamiltonii* compared to the *Cucumis sativus* reference plastome. Of 79 protein-coding genes present in *Cucumis*, traces of 30 were found in *P. aethiopica* and of 35 in *P. hamiltonii*, with 18 common to both species. Each species has at least 7 of the 30 tRNAs found in

Cucumis, with none of the 7 in common. rRNAs were lacking except for one piece of rrn16 in *P. hamiltonii*, probably a bacterial contamination (BLAST results unclear). For each gene found in the plastid-like contigs of *P. aethiopica* and *P. hamiltonii*, the size of the largest region without a stop codon aligning with the homologous open reading frame (ORF) of the reference is reported in Table S4. For *P. aethiopica*, only the ORFs of *pet*G, *psb*N and *psb*Z were longer than 2/3 of the reference ORFs, with only *psb*Z having an equivalent length (62 vs 63 amino-acids). The two other genes were missing 12 or 14 amino acids compared to *Cucumis*. For *P. hamiltonii*, only *rpl*36, *rpo*C1-e1 (but not *rpo*C1-e2) and *ycf*4 had an ORF longer than 2/3 of the reference ORF, with only *rpl3*6 having a comparable length (37 vs 38 amino-acids), the two other genes each missing 37 amino-acids compared to the reference.

Figure 2 shows the gene losses in *P. aethiopica* and *P. hamiltonii* compared to the only other endoparasite so far analyzed, all exo-holoparasites and exo-holomycotrophs, and all hemiparasites and partial mycotrophs, with a basic non-parasitic/non-mycotrophic photosynthetic angiosperm plastome added as a reference. It is clear from this figure that all genes involved in photosynthesis have been lost in at least one heterotroph, whereas 22 genes (cf. *Introduction*) have been kept in all exoparasites but are lost in endoparasites. The only two possibly functional plastid genes in *P. aethiopica* (*psb*Z) and *P. hamiltonii* (*rpl*36) are not part of the minimal set of genes found in exo-holoparasites/mycotrophs.

Differences between *P. aethiopica* and *P. hamiltonii*

A non-stringent BLAST search (maximum e-value set to 10, with 5 target sequences allowed) of the plastid-like contigs from *P. hamiltonii* against the ones from *P. aethiopica* yielded seven matches, all with e-values $<10^{-20}$. But when we aligned (global alignment) contigs from *P. hamiltonii* with matching contigs from *P. aethiopica*, only the regions previously found to match in the BLAST could be aligned, with no extension of the alignment. In the same way, when we aligned the entire *de novo* contigs, we could not elongate the original BLAST-based alignment, neither did we find other matching parts in the flanking regions.

DISCUSSION

No functional plastome in either species, and most remnants located in the chondriome

In photosynthetic plants, the read coverage of sequences obtained by random sequencing of genomic DNA provides an indication of their physical location because the copy number of plastome sequences is one or two orders of magnitude higher than that of

mitochondrial sequences, which in turn is one to two orders of magnitude higher than that of the nuclear genome (Zoschke et al., 2007). These ratios can be modified by endopolyploidization (Barow, 2006), ontogeny, at least for the chondriome (Preuten et al., 2010), and the presence of non-photosynthetic tissues (Isono et al., 1997). We therefore additionally relied on flanking regions to infer the organellar location of the plastid-like contigs obtained from *P. aethiopica* and *P. hamiltonii*. This revealed the mitochondrial location of most of the plastid-like contigs except for 14 in *P. aethiopica* and 28 in *P. hamiltonii*. The coverage of the 13 (*P. aethiopica*) and 12 (*P. hamiltonii*) non-placed contigs was more than one order of magnitude lower than the mean coverage of the chondriome-located sequences. It is therefore likely that these contigs are sitting in the nuclear genome. The coverage of the remaining singletons (1 in *P. aethiopica* and 16 in *P. hamiltonii*) was on the order of that of typical mitochondrial regions, so they could be sitting in the chondriome. None of the contigs without flanking regions could be assembled into a longer plastome contig.

An alternative explanation for mitochondrial flanking regions of plastid-like genes might be a plastome hosting numerous mitochondrial regions. However, incorporation of mitochondrial sequences into plastomes is exceedingly rare (for examples see Iorizzo et al., 2012a, 2012b; Straub et al., 2013), while the inverse scenario, where plastid-like regions have accumulated in chondriomes, is common (Notsu et al., 2002; Goremykin et al., 2008; Rice et al., 2013).

Differences between the plastomes of exo-holoparasites and those of endoparasites

Plastids fulfill metabolic functions other than photosynthesis, and of the ca. 116 land plant plastome genes, only ca. 50 are involved in photosynthesis while ca. 60 are involved in the modification of the RNAs or proteins encoded by the first 50 (Bock, 2007; Wicke et al., 2011). Including the two species studied here, the plastid genomes of 23 non-photosynthetic plants have now been investigated (Table 3). Of these, 20 are "normal" exo-holoparasites and mycotrophs (Table 3) and their plastomes still contain 32 to 92 protein-coding genes and RNAs, depending on lineage (Barrett et al., 2014). The retained plastome genes are *ycf*2, seven tRNAs (*trn*D-GUC, *trn*E-UUC, *trn*fM-CAU, *trn*I-CAU, *trn*Q-UUG, *trn*W-CCA and *trn*Y-GUA), ten ribosomal proteins and four ribosomal RNAs. The plastid-encoded *trn*E is involved in the biosynthesis of a mitochondrial heme and might be essential (Barbrook et al., 2006) whereas the function of *ycf*2 remains unknown. Why the other tRNAs are maintained is unclear, but the presence of *trn*E and of *ycf*2 may explain the conservation of ten ribosomal

protein genes and four ribosomal RNAs (Fig. 2). Three additional genes have also been considered as the possible "raison d'être" (Krause, 2012) of a plastome in non-photosynthetic plants, namely *acc*D, *clp*P, and *ycf*1. The *acc*D gene is involved in fatty-acids biosynthesis, *clp*P is likely a protease and also involved in the import of proteins into the plastid (Krause, 2012), and *ycf*1 is involved in "photosynthetic protein import, and [is] therefore essential for plant viability" (Kikuchi et al., 2013: p. 573). However, even these genes have been lost or pseudogenized in some parasitic Orobanchaceae (Li et al., 2013; Wicke et al., 2013).

As illustrated in Figure 2, all these putatively essential genes have been lost from endoparasites. In *P. aethiopica* and *P. hamiltonii*, we found non-functional fragments of *acc*D, *ycf*1 and *ycf*2, whereas *trn*E and a pseudogenized *clp*P were only found in *P. hamiltonii*. In both species, the fragments of these genes are located in the mitochondrial genome, although in *P. hamiltonii*, copies of the *acc*D gene may be located in the nuclear genome, judging from their coverage (Table 1). The endoparasite *Rafflesia lagascae* also lacks functional copies of these genes (Fig. 2).

Fate of the plastome in relation to the age and type of parasitism

Our data confirm that plastid genomes can become completely lost (Molina et al., 2014) and reveal that it has happened in all three endoparasites so far studied (Table 3), but not in any of the 20 exo-parasites (from 13 genera and 5 families). This suggests that endoparasitism may be the precondition for the complete loss of a plastome. This extreme mode of life may completely remove selection for plastid-encoded functions. The four families of endoparasites (Apodanthaceae, Cytinaceae, Mitrastemonaceae and Rafflesiaceae) are not closely related to any exo-parasites, and there is so far no scenario for how endoparasitism evolved, however, molecular clocks indicate that it evolved a long time ago. The crown age of Rafflesiaceae has been estimated as 82 (69-96) million years [Ma] old (Bendiksby et al., 2010) or minimally 37 (18-56) Ma (Naumann et al., 2013), that of Apodanthaceae as 36 to 57 Ma (Bellot and Renner, 2014a). Crown ages are not available for Cytinaceae and Mitrastemonaceae, but their stem ages of 72 (52-93) Ma and 78 (56-98) Ma (Naumann et al., 2013; Table 3) are similar to the stem age of Rafflesiaceae estimated to 95 (83-109) Ma (Bendiksby et al., 2010) or 65 (46-84) Ma (Naumann et al., 2013), and to that of Apodanthaceae estimated at 81 (62-98) Ma (Bellot and Renner, 2014a; Table 3). The stem age of the parasitic Orobanchaceae is 32 (13-52) Ma (Naumann et al., 2013; Table 3), while mycoheterotrophy in *Neottia*, *Corallorhiza* and *Rhizanthella* is probably younger than 49 Ma (Lovisa et al., 2010; Table 3). The origin of parasitism in the exo-parasitic lineages

investigated so far thus is younger than the origin of the endoparasitic life style of Apodanthaceae or Rafflesiaceae. It would be interesting to investigate the plastomes of ancient exo-holoparasites, such as the Hydnoraceae with a crown age of 58 (30-87) Ma, or the Balanophoraceae with a stem age of 110 (99-120) Ma (Naumann et al., 2013; Table 3). If such ancient exo-holoparasites still retained their plastomes this would lend indirect support to our hypothesis that endoparasitism is indeed the precondition for complete plastome loss.

Horizontal gene transfer from Fabaceae to Apodanthaceae

Some 20–27% of the plastid-like regions of *P. aethiopica* and 24% of those of *P. hamiltonii* appear to come from these plants' exclusive Fabaceae hosts. To reduce the risk of contamination artifacts, we filtered out contigs that resembled Fabales along their entire length; fragments that only partially matched Fabales were not excluded as long as their coverage was high. Any contamination with host plastid DNA would have been recognizable because it would have included functional plastid genes. Ours is not the first study to report HGT in Apodanthaceae. Barkmann et al. (2007) already reported that the mitochondrial *atp*1 gene of the Californian species *P. thurberi* is more similar to Fabaceae than to Cucurbitales. Large-scale HGT has also been documented in *Rafflesia cantleyi* (Xi et al., 2012, 2013) and *R. lagascae* (Molina et al., 2014). In the latter, 26 of 46 plastid-like contigs could be assigned to an angiosperm order, and 22 of them were most similar to Vitales, the order of the host, (*Tetrastigma*) whereas none could be assigned to the order of *Rafflesia* (Malpighiales; Molina et al., 2014).

That no angiosperm orders other than Cucurbitales and Fabales consistently came up in the numerous BLAST searches carried out for this project, fits with Apodanthaceae belonging in the Cucurbitales (Filipowicz and Renner, 2010) and with *Pilostyles* parasitizing Fabaceae (Bellot and Renner, 2014b). Sequences that could not be assigned to either Cucurbitales or Fabales were usually short and probably contained too little signal.

The dynamics of the mitochondrial genomes of the two species of Apodanthaceae

The mitochondrial genome of *Pilostyles* has absorbed 39-47 genes (Table S4) from its plastid genome before the latter disappeared. Mitochondrial genomes are particularly prone to accepting new DNA because of their constant reshuffling *via* recombination between sub-genomes (Woloszynska, 2010). The mitochondrial genomes of the Cucurbitaceae *Cucurbita pepo*, *Citrullus lanatus* and *Cucumis sativus* contain numerous plastid and nuclear-derived sequences (Alverson et al., 2010, 2011), and the dynamics of the mitochondrial genome of

the two species of *Pilostyles* studied here may thus reflect a Cucurbitales heritage rather than result from a parasitic way of life.

The numerous differences between *P. aethiopica* and *P. hamiltonii* in their chondriome-located plastid remnants are striking. These species diverged from each other 23-33 my ago (Bellot and Renner, 2014a: Table 2), and their hosts belong to different subfamilies of Fabaceae (Faboideae and Caesalpinioideae). It is plausible that the plastome of these two species was transferred to the mitochondrial genome a long time ago, with evolution since then reducing their original similarity, especially with the addition of plastid DNA from different legume hosts.

DATA ACCESS

Sequencing reads are deposited in the NCBI Sequence Read Archive.

ACKNOWLEDGEMENTS

We thank Martina Silber for technical support; K. Dixon, K. Thiele, M. Hyde and D. Plowes for plant material; M. Piednoel for teaching the first author Python; S. Wicke for technical help in 2010, and N. Cusimano for discussion. Financial support came from the German Science Foundation (RE 603/9-1).

REFERENCES

Alexander, L, K Woeste. 2014. Pyrosequencing of the northern red oak (*Quercus rubra* L.) chloroplast genome reveals high quality polymorphisms for population management. *Tree Genetics & Genomes* 10: 803–812.

Alverson, AJ, X Wei, DW Rice, DB Stern, K Barry, JD Palmer. 2010. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol Biol Evol.* 27: 1436–1448.

Alverson, AJ, DW Rice, S Dickinson, K Barry, JD Palmer. 2011. Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell* 23: 2499–513.

Aragao, FJL, APZ Mota. 07 March 2012. Direct submission to Genbank.

Atherton, RA, BJ McComish, LD Shepherd, LA Berry, NW Albert, PJ Lockhart. 2010. Whole genome sequencing of enriched chloroplast DNA using the Illumina GAII platform. *Plant Methods* 6, 22.

Barbrook, AC, CJ Howe, S Purton. 2006. Why are plastid genomes retained in non-photosynthetic organisms? *Trends Plant Sci.* 11: 101–108.

Barkman TJ, JR McNeal, SH Lim, G Coat, HB Croom, ND Young, CW dePamphilis. 2007. Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. *BMC Evol. Biol.* 7: 248.

Barow, M. 2006. Endopolyploidy in seed plants. *Bioessays* 28: 271 – 281.

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

Barrett, CF, JV Freudenstein, J Li, DR Mayfield-Jones, L Perez, JC Pires, C Santos. 2014. Investigating the path of plastid genome degradation in an early-transitional clade of heterotrophic orchids, and implications for heterotrophic angiosperms. *Mol Biol Evol* doi: 10.1093/molbev/msu252.

Bellot, S, SS Renner. 2013. Pollination and mating systems of Apodanthaceae and the distribution of reproductive traits in parasitic angiosperms. *Am. J. Bot.* 100: 1083–1094.

Bellot, S, SS Renner. 2014a. Exploring new dating approaches for parasites: The worldwide Apodanthaceae (Cucurbitales) as an example. *Mol Phyl Evol* 80: 1-10.

Bellot, S, SS Renner. 2014b. The systematics of the worldwide endoparasite family Apodanthaceae (Cucurbitales), with a key, a map, and color photos of most species. *Phytokeys* 36: 41–57.

Bendiksby, M, T Schumacher, G Gussarova, J Nais, K Mat-Salleh, N Sofiyanti, D Madulid, SA Smith, T Barkman. 2010. Elucidating the evolutionary history of the Southeast Asian, holoparasitic, giantflowered Rafflesiaceae: Pliocene vicariance, morphological convergence and character displacement. *Mol Phylogenet Evol* 57: 620–633.

Bock, R. 2007. Structure, function, and inheritance of plastid genomes. In *Cell and Molecular Biology of Plastids*, Vol. 19 of *Topics in Current Genetics*, pp. 29–63. Springer Berlin & Heidelberg, Germany

Boetzer, M, CV Henkel, HJ Jansen, D Butler, W Pirovano. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27: 578–579.

Camacho, C, G Coulouris, V Avagyan, N Ma, J Papadopoulos, K Bealer, TL Madden. 2008. BLAST+: architecture and applications. *BMC Bioinformatics* 10: 421.

Delannoy, E, S Fujii, CC des Francs, M Brundrett, I Small. 2011. Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. *Mol Biol Evol* 28: 2077–2086.

Dell, B, J Kuo, AH Burbidge. 1982. Anatomy of *Pilostyles hamiltonii* C. *A.* Gardner (Rafflesiaceae) in Stems of *Daviesia*. *Aust. J. Bot.* 30: 1–9.

Filipowicz, N, SS Renner. 2010. The worldwide holoparasitic Apodanthaceae confidently placed in the Cucurbitales by nuclear and mitochondrial gene trees. *BMC Evol. Biol.* 10: 219.

Fuentes, I, S Stegemann, H Golczyk, D Karcher, R Bock. 2014. Horizontal genome transfer as an asexual path to the formation of new species. *Nature* 511: 232–235.

Funk, H, S Berg, K Krupinska, U Maier, K Krause. 2007. Complete DNA sequences of the plastid genomes of two parasitic flowering plant species, *Cuscuta reflexa* and *Cuscuta gronovii*. *BMC Plant Biol* **7**: 45.

Goremykin, VV, F Salamini, R Velasco, R Viola. 2009. Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. Mol. Biol. Evol. 26: 99 – 110.

Heide-Jorgensen, HS. 2008. Parasitic flowering plants. Brill, The Netherlands.

Iorizzo, M, D Grzebelus, D Senalik, M Szklarczyk, D Spooner, P Simon. 2012a. Against the traffic: the first evidence for mitochondrial DNA transfer into the plastid genome. *Mob Genet Elements* 2: 261–266.

Iorizzo, M, D Senalik, M Szklarczyk, D Grzebelus, D Spooner, P Simon. 2012b. De novo assembly of the carrot mitochondrial genome using next generation sequencing of whole genomic DNA provides first evidence of DNA transfer into an angiosperm plastid genome. *BMC Plant Biol.* 12: 61.

Isono, K, Y Niwa, K Satoh, H Kobayashi. 1997. Evidence for transcriptional regulation of plastid photosynthesis genes in *Arabidopsis thaliana* roots. *Plant Physiol* 114: 623–630.

Jansen, RK, MF Wojciechowski, E Sanniyasi, SB Lee, H Daniell. 2008. Complete plastid genome sequence of the chickpea (*Cicer arietinum*) and the phylogenetic distribution of rps12 and clpP intron losses among legumes (Leguminosae). *Mol. Phylogenet. Evol.* 48: 1204–1217.

Jansen, RK, C Saski, SB Lee, AK Hansen, H Daniell. 2011. Complete plastid genome sequences of three Rosids (*Castanea*, *Prunus*, *Theobroma*): evidence for at least two independent transfers of rpl22 to the nucleus. *Mol. Biol. Evol.* 28: 835–847.

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

Kazakoff, SH, M Imelfort, D Edwards, J Koehorst, B Biswas, J Batley, PT Scott, PM Gresshoff. 2012. Capturing the biofuel wellhead and powerhouse: the chloroplast and mitochondrial genomes of the leguminous feedstock tree *Pongamia pinnata*. *PLoS ONE* 7: e51687.

Kikuchi, S, J Bédard, M Hirano, Y Hirabayashi, M Oishi, M Imai, M Takase, T Ide, M Nakai. 2013. Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* 339: 571–574.

Kim, G, ML LeBlanc, EK Wafula, CW dePamphilis, JH Westwood. 2014. Genomic-scale exchange of mRNA between a parasitic plant and its hosts. Science 345: 808–811.

Krause, K. 2012. Plastid genomes of parasitic plants: A trail of reductions and losses. In CE Bullerwell (ed.), Organelle Genetics Springer-Verlag Berlin Heidelberg pp. 79–103.

Li, H, B Handsaker, A Wysoker, T Fennell, J Ruan, N Homer, G Marth, G Abecasis, R Durbin, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25: 2078–2079.

Li, X, T-C Zhang, Q Qiao, Z Ren, J Zhao, T Yonezawa, M Hasegawa, MJC Crabbe, J Li, Y Zhong. 2013. Complete chloroplast genome sequence of holoparasite *Cistanche deserticola* (Orobanchaceae) reveals gene loss and horizontal gene transfer from its host Haloxylon ammodendron (Chenopodiaceae). *PLOS One* 8: e58747.

Logacheva, MD, MI Schelkunov, AA Penin. 2011. Sequencing and analysis of plastid genome in mycoheterotrophic orchid *Neottia nidus-avis*. *Genome Biol. Evol.* 3: 1296–1303.

Logacheva, MD, MI Schelkunov, MS Nuraliev, TH Samigullin, AA Penin. 2014. The plastid genome of mycoheterotrophic monocot *Petrosavia stellaris* exhibits both gene losses and multiple rearrangements. *Genome Biol. Evol.* 6: 238–246.

Lovisa, A, S Gustafsson, CF Verola, A Antonelli. 2010. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). *BMC Evol Biol.* 10: 177.

Malé, PJ, L Bardon, G Besnard, E Coissac, F Delsuc, J Engel, E Lhuillier, C Scotti-Saintagne, A Tinaut, J Chave. 2014. Genome skimming by shotgun sequencing helps resolve the phylogeny of a pantropical tree family. *Mol Ecol Resour* 14. 966–975.

McNeal, JR, J Kuehl, J Boore, CW dePamphilis. 2007. Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. *BMC Plant Biol* 7: 57.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17: 10–12.

Molina, J, KM Hazzouri, D Nickrent, M Geisler, RS Meyer, MM Pentony, JM Flowers, P Pelser, J Barcelona, SA Inovejas, I Uy, W Yuan, O Wilkins, CI Michel, S Locklear, GP Concepcion, MD Purugganan. 2014. Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Mol. Biol. Evol*. 31: 793–803.

Naumann, J, K Salomo, JP Der, EK Wafula, JF Bolin, E Maass, L Frenzke, M-S Samain, C Neinhuis, CW dePamphilis, S Wanke. 2013. Single-copy nuclear genes place haustorial Hydnoraceae within Piperales and reveal a Cretaceous origin of multiple parasitic angiosperm lineages. *PLoS ONE* 8: e79204.

Notsu, Y, S Masood, T Nishikawa, N Kubo, G Akiduki, M Nakazono, A Hirai, K Kadowaki. 2002. The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Mol. Genet. Genomics* 268: 434 – 445.

Plader, W, Y Yukawa, M Sugiura, S Malepszy. 2007. The complete structure of the cucumber (*Cucumis sativus* L.) chloroplast genome: its composition and comparative analysis. *Cell. Mol. Biol. Lett.* 12: 584-594.

Preuten, T, E Cincu, J Fuchs, R Zoschke, K Liere, T Börner. 2010. Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells. *The Plant Journal* 64: 948–959.

Quinlan, R, IM Hall. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841–842.

Ravi, V, JP Khurana, AK Tyagi, P Khurana. 2006 The chloroplast genome of mulberry: complete nucleotide sequence, gene organization and comparative analysis. *Tree Genet. Genomes* 3: 49–59.

Renner, SS, S Bellot. 2012. Horizontal gene transfer in eukaryotes, focusing on fungi-to-plant and plant-to-plant transfers. Pp. 223-235 in R. Bock & V. Knoop, eds., Genomics of chloroplasts and mitochondria. Springer, Heidelberg.

Rice, DW, AJ Alverson, AO Richardson, GJ Young, MV Sanchez-Puerta, J Munzinger, K Barry, JL Boore, Y Zhang, CW dePamphilis, EB Knox, JD Palmer. 2013. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. *Science* 342: 1468–1473.

Rivarola, M, JT Foster, AP Chan, AL Williams, DW Rice, X Liu, A Melake-Berhan, H Huot Creasy, D Puiu, MJ Rosovitz, HM Khouri, SM Beckstrom-Sternberg, GJ Allan, P Keim, J Ravel, PD Rabinowicz. 2011. Castor bean organelle genome sequencing and worldwide genetic diversity analysis. *PLoS ONE* 6: e21743.

Rodriguez-Moreno, L, VM Gonzalez, A Benjak, MC Marti, P Puigdomenech, MA Aranda, J Garcia-Mas.2011. Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. *BMC Genomics* 12: 424.

Rutherford, RJ. 1970. The Anatomy and Cytology of *Pilostyles thurberi* Gray (Rafflesiaceae). *Aliso* 7: 263–288.

Salamone, I, R Govindarajulu, S Falk, M Parks, A Liston, TL Ashman. 2013. Bioclimatic, ecological, and phenotypic intermediacy and high genetic admixture in a natural hybrid of octoploid strawberries. *Am. J. Bot.* 100: 939–950.

Schmieder R, R Edwards. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27: 863–864.

Shang, MZ, KB Wang, JP Hua, F Liu, CY Wang, YH Wang, X Zhang, SH Li. 10 Feb. 2011. Direct submission to NCBI.

Stamatakis, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.

Stegemann, S, R Bock. 2009. Exchange of genetic material between cells in plant tissue grafts. *Science* 324: 649–651.

Stegemann, S, M Keuthe, S Greiner, R Bock. 2012. Horizontal transfer of chloroplast genomes between plant species. *Proc. Natl. Acad. Sci. USA* 109: 2434–2438.

Straub, SCK, RC Cronn, C Edwards, M Fishbein, M Liston. 2013. Horizontal transfer of DNA from the mitochondrial to the plastid genome and its subsequent evolution in milkweeds (Apocynaceae). *Gen. Biol. and Evol*. 10: 1872–1885.

Tangphatsornruang, S, P Uthaipaisanwong, D Sangsrakru, J Chanprasert, T Yoocha, N Jomchai, S Tragoonrung. 2011. Characterization of the complete chloroplast genome of *Hevea brasiliensis* reveals genome rearrangement, RNA editing sites and phylogenetic relationships. *Gene* 475: 104–112.

Tuskan, GA, S Difazio, S Jansson, J Bohlmann, I Grigoriev, U Hellsten, N Putnam, S Ralph, S Rombauts, A Salamov, J Schein, L Sterck, A Aerts, RR Bhalerao, RP Bhalerao, D Blaudez, W Boerjan, A Brun, A Brunner, V Busov, M Campbell, J Carlson, M Chalot, J Chapman, GL Chen, D Cooper, PM Coutinho, J Couturier, S Covert, Q Cronk, R Cunningham, J Davis, S Degroeve, A Dejardin, C Depamphilis, J Detter, B Dirks, I Dubchak, S Duplessis, J Ehlting, B Ellis, K Gendler, D Goodstein, M Gribskov, J Grimwood, A Groover, L Gunter, B Hamberger, B Heinze, Y Helariutta, B Henrissat, D Holligan, R Holt, W Huang, N Islam-Faridi, S Jones, M Jones-Rhoades, R Jorgensen, C Joshi, J Kangasjarvi, J Karlsson, C Kelleher, R Kirkpatrick, M Kirst, A Kohler, U Kalluri, F Larimer, J Leebens-Mack, JC Leple, P Locascio, Y Lou, S Lucas, F Martin, B Montanini, C Napoli, DR Nelson, C Nelson, K Nieminen, O Nilsson, V Pereda, G Peter, R Philippe, G Pilate, A Poliakov, J Razumovskaya, P Richardson, C Rinaldi, K Ritland, P Rouze, D Ryaboy, J Schmutz, J Schrader, B Segerman, H Shin, A Siddiqui, F Sterky, A Terry, CJ Tsai, E Uberbacher, P Unneberg, J Vahala, K Wall, S Wessler, G Yang, T Yin, C Douglas, M Marra, G Sandberg, Y Van de Peer, D Rokhsar. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596–1604.

Wicke, S, GM Schneeweiss, CW dePamphilis, KF Müller, D Quandt. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol* **7**6: 273–297.

Wicke, S, KF Müller, dePamphilis CW, D Quandt, NJ Wickett, Y Zhang, SS Renner, GM Schneeweiss. 2013. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *The Plant Cell* 25: 3711–3725.

Wickett, NJ, Y Zhang, SK Hansen, JM Roper, JV Kuehl, SA Plock, PG Wolf, CW dePamphilis, JL Boore, B Goffinet. 2008. Functional gene losses occur with minimal size reduction in the plastid genome of the parasitic liverwort *Aneura mirabilis*. *Mol Biol Evol* 25: 393–401.

Wolfe, KH, CW Morden, JD Palmer. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc Natl Acad Sci USA* 89: 10648–10652.

Woloszynska M. 2010. Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes — though this be madness, yet there's method in't. *J Exp Bot* 61: 657–671.

Xi, Z, RK Bradley, KJ Wurdack, KM Wong, M Sugumaran, K Bomblies, JS Rest, CC Davis. 2012. Horizontal transfer of expressed genes in a parasitic flowering plant. *BMC Genomics* 13: 227.

Xi, Z, Y Wang, RK Bradley, M Sugumaran, CJ Marx, JS Rest, CC Davis. 2013. Massive mitochondrial gene transfer in a parasitic flowering plant clade. *PLoS Genet* 9: e1003265.

Zoschke, R, K Liere, T Börner. 2007. From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. *Plant Journal* 50: 710–722

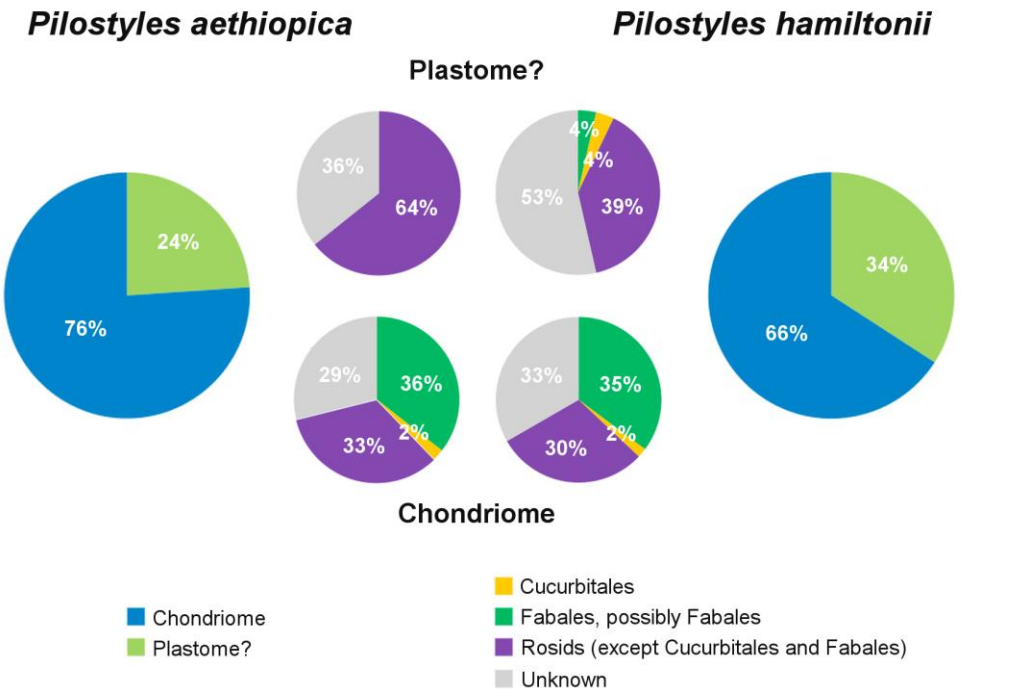

Figure 1. Proportions of plastid-like regions found in *Pilostyles aethiopica* and *P. hamiltonii*, with respect to their genomic location and phylogenetic affinity. Genomic location was inferred from flanking regions and coverage. Regions indicated as "Plastome?" are the ones lacking flanking regions (see *Discussion*). Phylogenetic affinity was inferred from BLAST searches for both species, and from phylogenetic reconstructions for *P. aethiopica* (*Material and Methods*). The near-absence of Fabales-like sequences among contigs lacking flanking regions ("Plastome?") is partly because sequences completely matching Fabales were excluded as possible contaminations (*Material and Methods*)

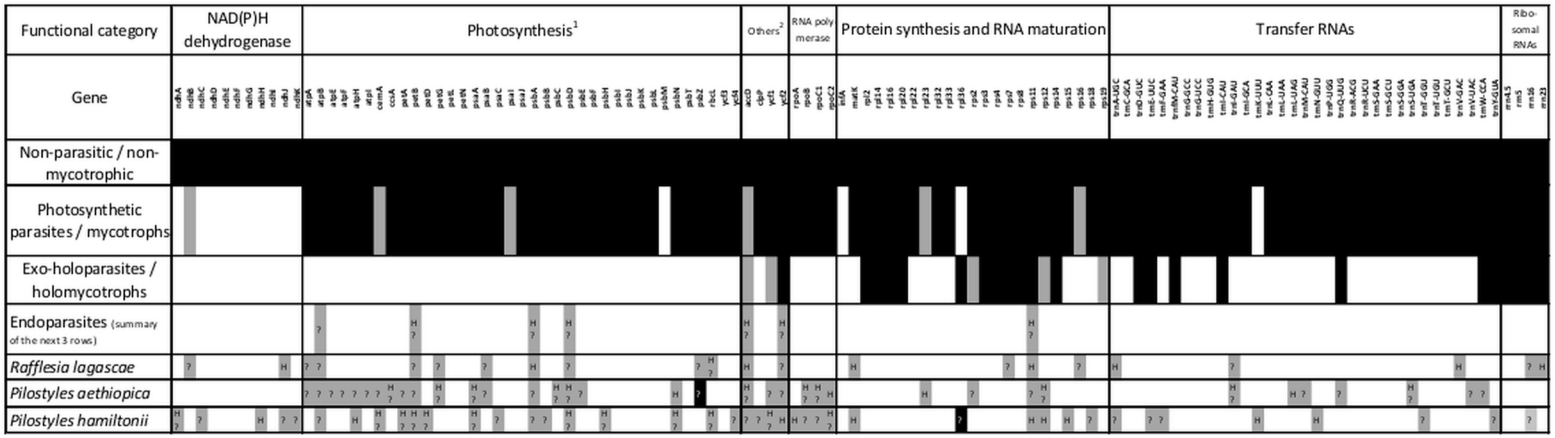

Figure 2. Summary of the plastome gene content of parasitic and mycotrophic plants, adapted from Barrett et al. (2014), Molina et al. (2014), and Li et al. (2013). Species are grouped by type of parasitism or mycotrophy. White fields indicate that a gene is lost in at least one lineage of a given kind of parasitism, whereas clear fields indicate gene pseudogenized at least once. Black fields indicate that a gene is functional in all lineages of a parasitism category. For endoparasites, genes transferred to other genomic compartments are included. "H" means that the gene has been acquired from the host, whereas a question mark indicates a probable autochthonous origin. The hatched box indicates a possible bacterial contamination.

Table 1. Genomic location, phylogenetic affinity, and mean coverage of all plastid-like sequences found in *Pilostyles aethiopica* and *P. hamiltonii*. C: Cucurbitales, F: Fabales, pF: probably Fabales as inferred from the thresholds for HGT assessment (see *Material and Methods*) R: another order of rosids *sensu lato*, O: another order. CP?: no flanking regions, MT: located in the chondriome.

Pilostyles hamiltonii						*Pilostyles aethiopica*					
Plastid fragment	Size (bp)	Reference genes matched	Phylo-genetic affinity	Geno-mic location	Mean cover-age	Plastid fragment	Size (bp)	Reference genes matched	Phylo-genetic affinity	Geno-mic location	Mean cover-age
Ph_4	184	*rpoB*	O	CP?	2	Pa_7	146	*rpoB*	R	CP?	3
Ph_11	157	*orf224*	O	CP?	2	Pa_59	528	*petB*	O	CP?	4
Ph_14	123	*ndhK, ndhC*	R	CP?	2	Pa_1	143	*atpA*	O	CP?	4
Ph_2	115	*rpoC1, rpoB*	O	CP?	3	Pa_2	228	*atpH*	O	CP?	4
Ph_10	153	*atpB*	O	CP?	3	Pa_11	171	*petA*	R	CP?	4
Ph_12	484	*rpoB*	O	CP?	3	Pa_3	145	*rps11*	O	CP?	6
Ph_9	104	*accD*	O	CP?	4	Pa_5	153	*atpA*	R	CP?	6
Ph_16	116	*psbB*	R	CP?	4	Pa_9	189	*atpE, atpB*	R	CP?	6
Ph_5	175	*ycf4*	O	CP?	5	Pa_12	378	*psaA*	R	CP?	6
Ph_7	209	*accD*	O	CP?	6	Pa_4	409	*trnM*^{CAU}, *psaB*	O	CP?	7
Ph_3	239	*accD*	O	CP?	7	Pa_8	238	*psaA*	R	CP?	7
Ph_18	711	*ycf3*_intron	R	CP?	10	Pa_13	101	*rpoB*	R	CP?	8
Ph_6	318	*accD*	O	CP?	40	Pa_10	313	*petA*	R	CP?	9
Ph_20	217	*clpP*	R	CP?	41	Pa_6	546	*rpoC1, rpoB*	R	CP?	302
Ph_1	236	*atpB*	C	CP?	43	Pa_39	259	*ycf3*_intron	R	MT	117
Ph_19	289	*psbB*	R	CP?	52	Pa_31	179	*psbN*	pF	MT	140
Ph_24	151	*ndhJ*	O	CP?	61	Pa_38	170	*psbD, psbC, trnS*^{UGA}	pF	MT	143
Ph_13	592	*rpoB*	O	CP?	62	Pa_54	500	*ycf1, orf224*	O, R	MT	147
Ph_23	2006	*ndhA*-e2	O	CP?	64	Pa_57	153	*rpoC2*	pF	MT	151
Ph_26	950	*ndhJ, trnF*^{GAA}	R	CP?	72	Pa_49	153	*atpA*	R	MT	164
Ph_21	723	*ndhK, ndhC*	R	CP?	73	Pa_42	133	*trnI*	F	MT	166
Ph_28	396	*ndhK, ndhC*	R	CP?	75	Pa_36	286	*trnQ*^{UUG}	R	MT	172
Ph_15	352	*trnF*^{GAA}	R	CP?	84	Pa_58	1254	*atpE, atpB*	O	MT	175
Ph_27	515	*ycf2*	pF	CP?	90	Pa_34	353	*petG, trnW*^{CAA}	C	MT	181
Ph_25	561	*ndhJ, ndhK, ndhC*	O	CP?	94	Pa_17	308	*psbE*	R	MT	185
Ph_22	517	*psbB*	R	CP?	140	Pa_26	381	*psaA*	R	MT	189
Ph_8	421	*rpl36*	O	CP?	263	Pa_16	190	*trnL*	F	MT	211
Ph_17	804	*psbA*	R	CP?	381	Pa_50	2460	*rpoB, rpoC1*	pF, R	MT	212
Ph_40	84	*rbcL*	pF	MT	2	Pa_30	2898	*psbD, psbC, trnS*^{UGA}, *psbZ*	R, O	MT	214
Ph_59	76	*rpoC2*	O	MT	16	Pa_25	606	*psaA*	F	MT	221
Ph_61	154	*rrn16*	C	MT	40	Pa_33	218	*petG*	pF	MT	229

Ph_65	116	*trnN*^{GUU}	pF	MT	42
Ph_73	228	*psaA*	pF	MT	44
Ph_41	97	*rbcL*	pF	MT	44
Ph_56	543	*trnY*^{GUA}, *trnE*^{UUC}	R	MT	53
Ph_60	351	*rps12*	pF	MT	58
Ph_38	151	*rbcL*	pF	MT	58
Ph_64	499	*ycf1*, *orf224*	O	MT	59
Ph_80	1125	*trnT*^{GGU}	R	MT	60
Ph_57	145	*trnM*^{CAU}	pF	MT	60
Ph_51	148	*rbcL*	pF	MT	68
Ph_50	259	*ycf2*	pF	MT	74
Ph_29	174	*rpoC2*, *rpoC1*-e2	O	MT	77
Ph_62	319	*rps18*	O	MT	84
Ph_77	916	*psbB*, *clpP*	O	MT	86
Ph_54	755	*psaA*	R	MT	88
Ph_49	278	*ycf3*_intron	O	MT	90
Ph_47	636	*psaA*	R	MT	91
Ph_58	263	*rpoC2*	O	MT	94
Ph_71	2352	*trnK*, *matK*	F	MT	97
Ph_39	229	*ndhJ*	O	MT	102
Ph_63	186	*ycf3*_intron	R	MT	102
Ph_76	680	*psbB*	R	MT	103
Ph_34	392	*petA*	O	MT	104
Ph_42	142	*trnV*^{GAC}_IG	F	MT	106
Ph_82	590	*rbcL*	F	MT	106
Ph_37	1788	*rpoC2*	O	MT	110
Ph_78	2264	*clpP*	R	MT	110
Ph_48	383	*atpH*	pF	MT	111
Ph_70	881	*rbcL*	O	MT	112
Ph_67	4027	*rpoC1*-e1, *rpoB*	O	MT	114
Ph_52	211	*psbN*, *psbH*, *petB*, *petD*	R	MT	114
Ph_53	672	*psaC*	R	MT	115
Ph_30	348	*clpP*-e2	R	MT	119
Ph_72	560	*rbcL*	R	MT	139
Ph_44	675	*cemA*, *petA*	F	MT	144
Ph_46	284	*trnA*^{UGC}	R	MT	149
Ph_66	313	*psbB*	R	MT	154
Ph_32	508	*psbD*	F	MT	157
Ph_69	5202	*psbN*, *psbH*, *petB*, *petD*, *rpoA*, *rps11*	F	MT	159
Pa_41	293	*rpoB*	O	MT	233
Pa_37	574	*accD*	pF, R	MT	236
Pa_29	432	*trnR*^{ACG}-*IG*	R	MT	237
Pa_23	266	*rpl23*	pF	MT	239
Pa_21	1961	*ycf2*, *trnI*, *atpE*, *atpB*	O	MT	240
Pa_43	1050	*rps2*	O	MT	241
Pa_44	586	*accD*	R	MT	247
Pa_51	245	*rps12*	F	MT	253
Pa_28	4151	*ccsA*, *trnL*^{UAG}	R, F	MT	259
Pa_53	206	*trnV*, *trnM*^{CAU}	O	MT	265
Pa_15	251	*trnI*	O	MT	265
Pa_27	225	*cemA*	R	MT	272
Pa_40	587	*rpoB*	R	MT	275
Pa_48	455	*petA*	O	MT	279
Pa_46	217	*psbD*, *psbC*, *trnS*^{UGA}	pF	MT	290
Pa_32	3443	*atpI*, *atpA*, *atpF*	UK	MT	291
Pa_24	330	*psbD*, *psbC*, *trnS*^{UGA}	O	MT	293
Pa_14	418	*trnI*	R	MT	300
Pa_47	304	*psbD*, *psbC*, *trnS*^{UGA}	O	MT	300
Pa_20	319	*ycf3_intron*	F	MT	301
Pa_18	1124	*trnV*, *trnM*^{CAU}	R	MT	304
Pa_56	2582	*psbD*, *psbC*, *trnS*^{UGA}, *rpoB*	F, C	MT	311
Pa_19	368	*trnM*^{CAU}, *psaB*	R	MT	314
Pa_22	355	*trnI*	UK	MT	317
Pa_52	585	*psbA*	R	MT	321
Pa_35	1561	*atpE*, *atpB*	R	MT	324
Pa_45	562	*psbD*, *psbC*, *trnS*^{UGA}	F	MT	335
Pa_55	531	*rps12*	R	MT	394

Ph_43	**580**	***rpoC2***	**F**	**MT**	**164**
Ph_79	**836**	***rpoB***	**O**	**MT**	**164**
Ph_33	**486**	***rpl16*_intron**	**O**	**MT**	**177**
Ph_45	**294**	***ycf3*_intron**	**O**	**MT**	**183**
Ph_31	**697**	***psbD***	**O**	**MT**	**184**
Ph_35	**637**	***cemA, petA***	**R**	**MT**	**190**
Ph_55	**1798**	***rpoC2, rpoC1*-e2**	**R**	**MT**	**196**
Ph_75	**544**	***ycf1***	**F**	**MT**	**200**
Ph_74	**2099**	***ndhA, ndhH, rps15***	**F**	**MT**	**201**
Ph_36	**2439**	***ycf4, accD, rbcL***	**O**	**MT**	**206**
Ph_68	**1320**	***ndhK, ndhC***	**R**	**MT**	**245**
Ph_81	**1115**	***accD***	**O**	**MT**	**280**

Table 2. Coverage of the plastid-like sequences of *Pilostyles* according to their genomic location. sd: standard deviation.

	Pilostyles aethiopica		*Pilostyles hamiltonii*	
	mean	sd	mean	sd
Plastome?	26.8	79.2	60.2	84.4
Chondriome	243.4	62.9	114.9	58.1

Table 3. Plastome content of non-photosynthetic land plants and age of parasitism/mycotrophy. [1]The number of genes in a typical angiosperm is 116 (from Barrett et al., 2014), whereas for photosynthetic *Aneura* it is 121 (Wickett et al., 2008). [2]The age of parasitism/mycotrophy is younger than the stem age of the respective parasite/mycotrophic clade. References for molecular clock-inferred ages are given in the text (*Discussion*).

Name	Type of parasitism	Plastome	Number of functional genes/RNAs[1]	Source	Age of parasitism or holomycotrophy[2] (Ma)
Pilostyles aethiopica	Endo-holoparasite	Not found	Maybe one	This study	< 81 (62-98)
Pilostyles hamiltonii	Endo-holoparasite	Not found	Maybe one	This study	< 81 (62-98)
Rafflesia lagascae	Endo-holoparasite	Not found	Not found	Molina et al., 2014	< 95 (83-109)
Aneura mirabilis	Exo-holomycotroph	Yes	92	Wickett et al., 2008	?
Corallorhiza maculata ssp. *maculata*	Exo-holomycotroph	Yes	89	Barrett et al., 2014	?
Corallorhiza maculata ssp. *occidentalis*	Exo-holomycotroph	Yes	88	Barrett et al., 2014	?
Corallorhiza mertensiana	Exo-holomycotroph	Yes	90	Barrett et al., 2014	?
Corallorhiza striata	Exo-holomycotroph	Yes	82	Barrett and Davis, 2012	?
Neottia nidus-avis	Exo-holomycotroph	Yes	59	Logacheva et al., 2011	?
Petrosavia stellaris	Exo-holomycotroph	Yes	72	Logacheva et al., 2014	?
Rhizantella gardneri	Exo-holomycotroph	Yes	32	Delannoy et al., 2011	?
Boulardia latisquama	Exo-holoparasite	Yes	54	Wicke et al., 2013	< 32 (13-52)
Cistanche deserticola	Exo-holoparasite	Yes	62	Li et al., 2013	< 32 (13-52)
Cistanche phelypaea	Exo-holoparasite	Yes	60	Wicke et al., 2013	< 32 (13-52)
Conopholis americana	Exo-holoparasite	Yes	49	Wicke et al., 2013	< 32 (13-52)
Cuscuta gronovii	Exo-holo?parasite	Yes	89	Funk et al., 2007	< 35 (13-57)
Cuscuta obtusiflora	Exo-holo?parasite	Yes	92	McNeal et al., 2007	< 35 (13-57)
Epifagus virginiana	Exo-holoparasite	Yes	51	Wolfe et al., 1992	< 32 (13-52)
Myzorrhiza californica	Exo-holoparasite	Yes	78	Wicke et al., 2013	< 32 (13-52)
Orobanche crenata	Exo-holoparasite	Yes	63	Wicke et al., 2013	< 32 (13-52)
Orobanche gracilis	Exo-holoparasite	Yes	58	Wicke et al., 2013	< 32 (13-52)
Phelipanche purpurea	Exo-holoparasite	Yes	60	Wicke et al., 2013	< 32 (13-52)
Phelipanche ramosa	Exo-holoparasite	Yes	57	Wicke et al., 2013	< 32 (13-52)
Hydnoraceae	**Exo-holoparasite**	**?**	**?**	**Na**	**< 101 (77–124)**
Balanophoraceae	**Exo-holoparasite**	**?**	**?**	**Na**	**< 110 (99–120)**

Table S1. Plastid genomes used in this study with their GenBank accession numbers, authors, size and gene content.

Species	Order	Family	GenBank accession	Source	Size (bp)	Number of protein-coding genes	Number of RNAs
Castanea mollissima	Fagales	Fagaceae	NC_014674	Jansen et al., 2011	160799	83	45
Cicer arietinum	Fabales	Fabaceae	NC_011163	Jansen et al., 2008	125319	75	33
Corynocarpus laevigata	Cucurbitales	Corynocarpaceae	NC_014807	Atherton et al., 2012	159202	83	44
Cucumis melo	Cucurbitales	Cucurbitaceae	NC_015983	Rodriguez-Moreno et al., 2011	156017	88	45
Cucumis sativus	Cucurbitales	Cucurbitaceae	NC_007144	Plader et al., 2007	155293	85	45
Fragaria chiloensis	Rosales	Rosaceae	NC_019601	Salamone et al., 2013	155603	85	45
Gossypium anomalum	Malvales	Malvaceae	NC_023213	Shang et al., 2011	159507	86	45
Hevea brasiliensis	Malpighiales	Euphorbiaceae	NC_015308	Tangphatsornruang et al., 2011	161191	84	45
Hirtella racemosa	Malpighiales	Chrysobalanaceae	NC_024060	Male et al., 2014	162891	83	45
Licania alba	Malpighiales	Chrysobalanaceae	NC_024064	Male et al., 2014	162467	83	45
Millettia pinnata	Fabales	Fabaceae	NC_016708	Kasakoff et al., 2012	152968	83	45
Morus indica	Rosales	Moraceae	NC_008359	Ravi et al., 2006	158484	84	45
Populus trichocarpa	Malpighiales	Salicaceae	NC_009143	Tuskan et al., 2006	157033	98	45
Prunus persica	Rosales	Rosaceae	NC_014697	Jansen et al., 2011	157790	85	45
Quercus rubra	Fagales	Fagaceae	NC_020152	Alexander and Woeste, 2012	161304	89	48
Ricinus communis	Malpighiales	Euphorbiaceae	NC_016736	Rivarola et al., 2011	163161	86	45
Vigna unguiculata	Fabales	Fabaceae	NC_018051	Arago and Mota, 2012	152415	84	46

Table S2. Plastid-like sequences found in *P. aethiopica* and *P. hamiltonii*, as well as their flanking regions when mapped on de novo contigs, with their phylogenetic affinity, coverage and BLAST-based genomic location. The definitive genomic location is based on BLAST results, as well as on the coverage and genomic location of the flanking regions. For the latter, the definitive genomic location was assessed after checking for possible misleading gene transfers in the BLAST hits. Full data available upon request.

Pilostyles hamiltonii								*Pilostyles aethiopica*							
De *novo* contig	Start on the *de novo* contig	End on the *de novo* contig	Name	Genomic compart-ment[1]	Genomic compart-ment[2]	Phylo-genetic affinity[3]	Mean cove-rage	De *novo* contig	Start on the *de novo* contig	End on the *de novo* contig	Name	Genomic compart-ment[1]	Genomic compart-ment[2]	Phylo-genetic affinity[3]	Mean cove-rage
3	1	21536	Flanking	MT	MT	O	100	1	1	9199	Flanking	MT	MT	R	345
3	21850	22024	Ph_29	na	MT	O	77	1	9199	9617	Pa_14	na	MT	R	300
3	22024	27447	Flanking	MT	MT	R	102	1	9617	9993	Flanking	MT	MT	R	272
5	1	5478	Flanking	MT	MT	R	139	1	9993	10244	Pa_15	na	MT	O	265
5	5478	5826	Ph_30	na	MT	R	119	1	10244	35732	Flanking	MT	MT	R	317
5	5826	23345	Flanking	MT	MT	R	93	8	1	6486	Flanking	MT	MT	O	291
9	1	3489	Flanking	MT	MT	O	189	8	6486	6676	Pa_16	na	MT	F	211
9	3489	4186	Ph_31	na	MT	O	184	8	6676	31670	Flanking	MT	MT	O	307
9	4186	19255	Flanking	MT	MT	O	137	10	1	21175	Flanking	MT	MT	R	306
9	19847	20160	Ph_66	na	MT	R	154	10	21175	21483	Pa_17	na	MT	R	185
10	1	3306	Flanking	MT	MT	O	125	10	21483	31555	Flanking	MT	MT	R	313
10	3660	10707	Flanking	MT	MT	R	122	17	1	9570	Flanking	MT	MT	R	310
10	10707	11215	Ph_32	na	MT	F	157	17	9570	10694	Pa_18	na	MT	R	304
10	11215	19715	Flanking	MT	MT	O	173	17	10694	29040	Flanking	MT	MT	O	307
17	1	134	Flanking	UK	UK	NA	128	20	1	31903	Flanking	MT	MT	C	323
17	134	620	Ph_33	na	MT	O	177	20	31903	32271	Pa_19	na	MT	R	314
17	620	7534	Flanking	MT	MT	O	157	20	32271	39970	Flanking	MT	MT	R	281
17	7534	7926	Ph_34	na	MT	O	104	24	1	31	Flanking	UK	UK	NA	417
17	7926	18145	Flanking	MT	MT	O	131	24	187	11435	Flanking	MT	MT	O	50
22	1	13048	Flanking	MT	MT	R	152	24	11435	11754	Pa_20	na	MT	F	301
22	13048	17075	Ph_67	na	MT	O	114	24	11754	14612	Flanking	MT	MT	O	327
25	1	10345	Flanking	MT	MT	O	84	24	14612	16573	Pa_21	na	MT	O	240
25	10345	10982	Ph_35	na	MT	R	190	24	16573	31604	Flanking	MT	MT	R	304
25	10982	16830	Flanking	MT	MT	O	109	29	1	25438	Flanking	MT	MT	R	295
31	1	4088	Flanking	MT	MT	O	162	29	25438	25793	Pa_22	na	MT	UK	317
31	4088	6527	Ph_36	na	MT	O	206	29	25793	37296	Flanking	MT	MT	R	318
31	6527	15331	Flanking	MT	MT	O	192	37	1	11284	Flanking	MT	MT	R	315
32	1	10256	Flanking	MT	MT	R	130	37	11284	11550	Pa_23	na	MT	pF	239
32	10256	12044	Ph_37	na	MT	O	110	37	11718	12469	Flanking	MT	MT	O	315
32	12044	13513	Flanking	MT	MT	O	90	37	12469	12799	Pa_24	na	MT	O	293
32	13513	14833	Ph_68	na	MT	R	245	37	12799	25170	Flanking	MT	MT	O	331
33	1	7102	Flanking	MT	MT	R	120	42	1	18421	Flanking	MT	MT	O	288
33	7102	7253	Ph_38	na	MT	pF	58	42	18421	19027	Pa_25	na	MT	F	221
33	7253	14770	Flanking	MT	MT	R	128	42	19027	19030	Flanking	MT	MT	O	293
38	1	8843	Flanking	MT	MT	R	127	42	19030	19411	Pa_26	na	MT	R	189
38	9014	14216	Ph_69	na	MT	F	159	42	19411	24464	Flanking	MT	MT	O	246
42	1	5508	Flanking	MT	MT	O	81	50	1	13117	Flanking	MT	MT	R	325
42	5508	5737	Ph_39	na	MT	O	102	50	13117	13342	Pa_27	na	MT	R	272
42	5737	5766	Flanking	MT	MT	O	56	50	13342	13372	Flanking	UK	UK	NA	303
43	1	571	Flanking	MT	MT	O	241	50	13614	15280	Flanking	MT	MT	O	366
43	571	655	Ph_40	na	MT	pF	2	50	15405	18905	Flanking	MT	MT	R	202
43	655	13592	Flanking	MT	MT	C	96	50	18905	23056	Pa_28	na	MT	R, F	259
44	1	8612	Flanking	MT	MT	O	124	51	1	4278	Flanking	NC	MT	O	307
44	8612	8709	Ph_41	na	MT	pF	44	51	4278	5532	Pa_58	na	MT	O	175
44	8709	13444	Flanking	MT	MT	O	104	51	5532	10620	Flanking	MT	MT	O	326
46	1	12426	Flanking	MT	MT	O	126	51	10620	11052	Pa_29	na	MT	R	237
46	12426	13307	Ph_70	na	MT	O	112	51	11052	14561	Flanking	MT	MT	R	382
47	1	3030	Flanking	MT	MT	O	167	51	14561	17459	Pa_30	na	MT	R, O	214
47	3030	3172	Ph_42	na	MT	F	106	51	16847	16863	Flanking	MT	MT	O	247

47	3172	13241	Flanking	MT	MT	R	200	51	17459	20332	Flanking	MT	MT	O	285
52	1	7880	Flanking	MT	MT	R	210	51	20332	20511	Pa_31	na	MT	pF	140
52	8309	12362	Flanking	MT	MT	F	143	51	20511	22971	Flanking	MT	MT	R	264
52	12362	12942	Ph_43	na	MT	F	164	67	1	20924	Flanking	MT	MT	R	176
58	1	10001	Flanking	MT	MT	R	149	67	20924	24367	Pa_32	na	MT	UK	291
58	10001	12353	Ph_71	na	MT	F	97	69	1	5979	Flanking	MT	MT	F	331
62	1	4995	Flanking	MT	MT	O	82	69	5979	6197	Pa_33	na	MT	pF	229
62	4995	5670	Ph_44	na	MT	F	144	69	6197	11571	Flanking	MT	MT	O	345
62	5670	11772	Flanking	MT	MT	R	110	69	11571	11924	Pa_34	na	MT	C	181
70	1	5171	Flanking	MT	MT	R	105	69	11924	19317	Flanking	MT	MT	R	301
70	5171	5465	Ph_45	na	MT	O	183	70	1	5925	Flanking	MT	MT	O	380
70	5465	10995	Flanking	MT	MT	R	132	70	5925	7486	Pa_35	na	MT	R	324
78	1	4501	Flanking	MT	MT	O	124	70	7486	8711	Flanking	MT	MT	R	278
78	4501	4785	Ph_46	na	MT	R	149	70	8711	8997	Pa_36	na	MT	R	172
78	4785	10129	Flanking	MT	MT	O	79	70	8997	19273	Flanking	MT	MT	O	263
80	1	4220	Flanking	MT	MT	R	204	72	1	9744	Flanking	MT	MT	O	244
80	4220	4856	Ph_47	na	MT	R	91	72	9744	10318	Pa_37	na	MT	pF, R	236
80	4856	9968	Flanking	MT	MT	R	86	72	10318	19109	Flanking	MT	MT	O	312
83	1	5868	Flanking	MT	MT	O	120	73	1	17416	Flanking	MT	MT	R	282
83	5868	6251	Ph_48	na	MT	pF	111	73	17416	17586	Pa_38	na	MT	pF	143
83	6251	9173	Flanking	MT	MT	O	97	73	17586	18995	Flanking	MT	MT	R	293
83	9173	9733	Ph_72	na	MT	R	139	84	1	1476	Flanking	MT	MT	O	294
88	1	5286	Flanking	MT	MT	O	103	84	1476	1735	Pa_39	na	MT	R	117
88	5286	5564	Ph_49	na	MT	O	90	84	1735	17082	Flanking	MT	MT	O	319
88	5564	9431	Flanking	MT	MT	R	67	93	1	11536	Flanking	MT	MT	R	257
108	1	7784	Flanking	MT	MT	O	153	93	11536	12123	Pa_40	na	MT	R	275
108	7784	8012	Ph_73	na	MT	pF	44	93	12123	12128	Flanking	MT	MT	O	251
108	8012	8451	Flanking	UK	UK	R	65	93	12128	12421	Pa_41	na	MT	O	233
110	1	2965	Flanking	MT	MT	R	93	93	12421	16367	Flanking	MT	MT	O	281
110	2965	3224	Ph_50	na	MT	pF	74	95	1	793	Flanking	UK	UK	NA	3
110	3224	8338	Flanking	MT	MT	O	62	95	793	926	Pa_42	na	MT	F	166
120	1	946	Flanking	UK	UK	NA	181	95	926	14413	Flanking	MT	MT	O	303
120	946	3045	Ph_74	na	MT	F	201	95	14413	15463	Pa_43	na	MT	O	241
120	3045	3099	Flanking	UK	UK	NA	239	95	15630	16119	Flanking	MT	MT	R	360
120	3099	3643	Ph_75	na	MT	F	200	106	1	10609	Flanking	MT	MT	O	134
120	3643	7912	Flanking	MT	MT	O	147	106	10816	11402	Pa_44	na	MT	R	247
136	1	5429	Flanking	MT	MT	R	76	106	11402	25524	Flanking	MT	MT	R	390
136	5429	5577	Ph_51	na	MT	pF	68	107	1	14693	Flanking	MT	MT	O	353
136	5577	7019	Flanking	MT	MT	R	86	107	14693	15255	Pa_45	na	MT	F	335
158	1	221	Flanking	MT	MT	R	128	108	1	3099	Flanking	MT	MT	R	319
158	221	432	Ph_52	na	MT	R	114	108	3283	9175	Flanking	MT	MT	R	228
158	432	5986	Flanking	MT	MT	R	121	108	9175	9392	Pa_46	na	MT	pF	290
160	1	1376	Flanking	MT	MT	O	120	108	9392	9394	Flanking	UK	UK	NA	4
160	1376	2048	Ph_53	na	MT	R	115	108	9394	9698	Pa_47	na	MT	O	300
160	2048	5936	Flanking	MT	MT	O	94	108	9698	15212	Flanking	MT	MT	O	316
163	1	3571	Flanking	MT	MT	O	80	136	1	5928	Flanking	MT	MT	O	333
163	3571	4326	Ph_54	na	MT	R	88	136	5928	6383	Pa_48	na	MT	O	279
163	4326	5810	Flanking	MT	MT	O	89	136	6383	13413	Flanking	MT	MT	R	352
175	1	3903	Flanking	MT	MT	O	100	170	1	4622	Flanking	MT	MT	F	329
175	3903	4583	Ph_76	na	MT	R	103	170	4622	4775	Pa_49	na	MT	R	164
175	4583	4615	Flanking	BC	UK	NA	131	170	4775	11544	Flanking	MT	MT	O	323
175	4615	5531	Ph_77	na	MT	O	86	226	1	2460	Pa_50	na	MT	pF, R	212
180	1	1757	Flanking	MT	MT	O	144	226	2460	9244	Flanking	MT	MT	R	428
180	1757	3555	Ph_55	na	MT	R	196	245	1	4207	Flanking	MT	MT	R	351
180	3555	5450	Flanking	MT	MT	R	114	245	4207	4452	Pa_51	na	MT	F	253
198	1	37	Flanking	UK	UK	NA	10	245	4452	14616	Flanking	MT	MT	R	372
198	37	580	Ph_56	na	MT	R	53	267	1	3240	Flanking	MT	MT	O	305
198	580	4805	Flanking	MT	MT	O	68	267	3240	3825	Pa_52	na	MT	R	321
205	1	669	Flanking	MT	MT	O	105	267	3825	8568	Flanking	MT	MT	R	291
205	669	814	Ph_57	na	MT	pF	60	346	1	6144	Flanking	MT	MT	O	294
205	814	4705	Flanking	MT	MT	R	133	346	6144	6350	Pa_53	na	MT	O	265
251	1	928	Flanking	MT	MT	O	86	346	6350	12079	Flanking	MT	MT	O	327
251	928	1191	Ph_58	na	MT	O	94	377	1	884	Flanking	MT	MT	R	266
251	1191	1286	Flanking	MT	MT	NA	130	377	884	1384	Pa_54	na	MT	O, R	147
251	1286	1362	Ph_59	na	MT	O	16	377	1384	6619	Flanking	MT	MT	R	276
251	1362	3762	Flanking	NC	MT	O	80	420	1	5706	Flanking	MT	MT	R	280
270	1	2265	Ph_78	na	MT	R	110	420	5706	6237	Pa_55	na	MT	R	394
270	2265	3380	Flanking	MT	MT	R	62	908	1	1789	Flanking	MT	MT	O	315

315	1	2362	Flanking	MT	MT	R	115
315	2362	2713	Ph_60	na	MT	pF	58
315	2713	2865	Flanking	UK	UK	NA	48
491	1	836	Ph_79	na	MT	O	164
491	836	2035	Flanking	MT	MT	R	96
503	1	867	Flanking	UK	UK	NA	75
503	867	1021	Ph_61	na	MT	C	40
503	1021	2006	Flanking	MT	MT	O	96
506	1	858	Flanking	MT	MT	O	41
506	878	2003	Ph_80	na	MT	R	60
508	1	877	Flanking	MT	MT	O	45
508	877	1196	Ph_62	na	MT	O	84
508	1196	1997	Flanking	MT	UK	O	85
706	1	1115	Ph_81	na	MT	O	280
706	1115	1718	Flanking	MT	MT	NA	272
1024	1	546	Flanking	MT	MT	R	97
1024	546	732	Ph_63	na	MT	R	102
1024	732	1509	Flanking	MT	MT	R	94
1041	1	911	Flanking	MT	MT	R	114
1041	911	1501	Ph_82	na	MT	F	106
14490	1	238	Flanking	MT	MT	O	113
14490	238	737	Ph_64	na	MT	O	59
14490	737	777	Flanking	MT	MT	O	25
62882	1	179	Flanking	UK	UK	NA	63
62882	179	295	Ph_65	na	MT	pF	42
62882	295	509	Flanking	MT	MT	O	54
na	na	na	Ph_1	na	CP?	C	43
na	na	na	Ph_10	na	CP?	O	3
na	na	na	Ph_11	na	CP?	O	2
na	na	na	Ph_12	na	CP?	O	3
na	na	na	Ph_13	na	CP?	O	62
na	na	na	Ph_14	na	CP?	R	2
na	na	na	Ph_15	na	CP?	R	84
na	na	na	Ph_16	na	CP?	R	4
na	na	na	Ph_17	na	CP?	R	381
na	na	na	Ph_18	na	CP?	R	10
na	na	na	Ph_19	na	CP?	R	52
na	na	na	Ph_2	na	CP?	O	3
na	na	na	Ph_20	na	CP?	R	41
na	na	na	Ph_21	na	CP?	R	73
na	na	na	Ph_22	na	CP?	R	140
na	na	na	Ph_23	na	CP?	O	64
na	na	na	Ph_24	na	CP?	O	61
na	na	na	Ph_25	na	CP?	O	94
na	na	na	Ph_26	na	CP?	R	72
na	na	na	Ph_27	na	CP?	pF	90
na	na	na	Ph_28	na	CP?	R	75
na	na	na	Ph_3	na	CP?	O	7
na	na	na	Ph_4	na	CP?	O	2
na	na	na	Ph_5	na	CP?	O	5
na	na	na	Ph_6	na	CP?	O	40
na	na	na	Ph_7	na	CP?	O	6
na	na	na	Ph_8	na	CP?	O	263
na	na	na	Ph_9	na	CP?	O	4

908	1789	4371	Pa_56	na	MT	F, C	311
2390	1	1522	Flanking	MT	MT	R	392
2390	1522	1675	Pa_57	na	MT	pF	151
2390	1675	3276	Flanking	MT	MT	O	387
na	na	na	Pa_1	na	CP?	O	4
na	na	na	Pa_10	na	CP?	R	9
na	na	na	Pa_11	na	CP?	R	4
na	na	na	Pa_12	na	CP?	R	6
na	na	na	Pa_13	na	CP?	R	8
na	na	na	Pa_2	na	CP?	O	4
na	na	na	Pa_3	na	CP?	O	6
na	na	na	Pa_4	na	CP?	O	7
na	na	na	Pa_5	na	CP?	R	6
na	na	na	Pa_59	na	CP?	O	4
na	na	na	Pa_6	na	CP?	R	302
na	na	na	Pa_7	na	CP?	R	3
na	na	na	Pa_8	na	CP?	R	7
na	na	na	Pa_9	na	CP?	R	6

Table S3. Phylogenetic affinity of plastid-like sequences of *P. aethiopica* inferred from BLAST searches and from maximum likelihood phylogenetic trees including other rosids.

Putative gene(s)	Genomic location[1]	BLAST order (RosFagales: Rosales or Fagales)	BLAST-derived affinity[2]	Maximum-likelihood placement[3]	Bootstrap support (%)	Expected topology[4]	Alignment length	Maximum-likelihood-derived affinity[2]	Inferred phylo-genetic affinity[2]
atpA-2_7	MT	Cucurbitales	C	Malp	31	no	188	UK	C
atpI-2_1	MT	Cucurbitales	C	Fab	94	no	433	pF	C
petG-trnW-CAA-trnP-UGG_3	MT	Cucurbitales	C	RF	26	no	139	UK	C
rrn5451623_33	MT c?	Cucurbitales	C	na	na	no	59	na	C
atpA-2_6	MT	Cucurbitales	C	OG	0	yes	147	UK	C
rpoB_1	MT	Cucurbitales	C	Malp	59	yes but 1	1043	UK	C
ndhB-e2_1	MT	Fabales	F	na	na	no	68	na	F
psaA-4_5	MT	Fabales	F	Fab	71	no	65	pF	F
psbD-psbC-trnS-UGA_19	CP	Fabales	F	na	na	no	95	na	F
psbD-psbC-trnS-UGA_15	CP	Fabales	F	RF-C-Fab-Malp	0	no	136	UK	F
rpoA_2	CP	Fabales	F	Fab	71	no	97	pF	F
rpoB_11	CP	Fabales	F	na	na	no	93	na	F
rpoC1-e1-rpoB_5	CP	Fabales	F	Fab	64	no	182	UK	F
rrn5451623_46	MT	Fabales	F	na	na	no	85	na	F
rrn5451623_68	MT	Fabales	F	na	na	no	51	na	F
rrn5451623_29	MT	Fabales	F	Fab	37	no	331	UK	F
trnN-GUU_6	UK	Fabales	F	na	na	no	65	na	F
ccsA-2_1	MT	Fabales	F	Fab	55	no	198	UK	F
psaA-4_1	MT	Fabales	F	Fab	100	no	485	pF	F
psbD-psbC-trnS-UGA_3	MT	Fabales	F	Fab	72	no	229	pF	F
trnI-e2-trnA-e1_3	MT	Fabales	F	na	na	no	40	na	F
ycf3-e2-ycf3-e1_1	MT	Fabales	F	Fab	63	no	56	UK	F
trnM-CAU-rps14-psaB_4	MT c	Fabales	F	Malp	20	no	79	UK	F
ndhA-e2-ndhA-e1_1	MT c?	Fabales	F	C	62	no	60	UK	F
psaA-4_4	MT c?	Fabales	F	Fab	41	no	61	UK	F
psbN-psbH_1	MT c?	Fabales	F	Fab	28	no	315	UK	F
rbcL_3	MT c?	Fabales	F	na	na	no	210	na	F
rpoA_1	MT c?	Fabales	F	Fab	76	no	212	pF	F
rrn5451623_60	MT c?	Fabales	F	na	na	no	79	na	F
trnL-e1_1	MT c?	Fabales	F	C	30	no	144	UK	F
trnL-e2_1	MT c?	Fabales	F	Malp	25	no	240	UK	F
trnN-GUU_3	MT c?	Fabales	F	Fab	88	no	189	pF	F
trnR-UCU_1	MT c?	Fabales	F	Malp	45	no	101	UK	F
trnV-e2-e1-trnM-CAU_1	MT c?	Fabales	F	Fab	98	no	1267	pF	F
ycf2-trnI_2	MT c?	Fabales	F	Fab	52	no	793	UK	F
psaA-2_1	CP	Fabales	F	Fab	40	no	387	UK	F
atpA-2_1	MT	Crossosomatales	pF	Fab	72	yes	713	F	F
psbD-psbC-trnS-UGA_2	MT	Cucurbitales	pF	Fab	95	yes	1273	F	F
psbA-1_6	CP	Fabales	F	RF-C-Fab	30	yes	254	UK	F
rbcL_2	UK	Fabales	F	Fab	78	yes	826	F	F

rrn5451623_22	MT	Fabales	F	RF	57	yes	281	UK	F
ndhB-e1-rps12-e2-e3_5	MT	Fabales	F	Fab	93	yes	168	F	F
ndhB-e1-rps12-e2-e3_2	MT c	Fabales	F	Fab	94	yes	494	F	F
atpE-atpB_2	MT c?	Fabales	F	Fab	90	yes	962	F	F
trnS-GGA-rps4_1	MT c?	Fabales	F	Fab	77	yes	702	F	F
psaA-4_9	CP	Myrtales	pF	Fab	97	yes	171	F	F
cemA_1	MT c?	Fabales	F	Fab	92	yes but 1	631	F	F
petA-1_1	MT c?	Fabales	F	Fab	96	yes but 1	477	F	F
trnF-GAA_1	MT c?	Fabales	F	RF	25	yes but 1	258	UK	F
psaA-3_1	MT	Fabales	F	RF	21	yes but 1	600	UK	F
psbA-1_2	MT c	Fabales	F	Fab	94	yes but 1	386	F	F
psbA-1_1	MT c	Fabales	F	Fab	70	yes but 1	729	F	F
rrn5451623_127	UK	Acholeplasmatales	O	na	na	no	91	na	O
rrn5451623_91	UK	Alteromonadales	O	Fab	70	no	389	pF	O
rpl23-trnI-CAU_1	MT	Amborellales	O	na	na	no	92	na	O
trnM-CAU-rps14-psaB_2	MT	Amborellales	O	RF	30	no	91	UK	O
ycf3-e2_4	MT	Amborellales	O	RF	20	no	87	UK	O
trnV-e2-e1-trnM-CAU_5	MT	Asparagales	O	Malp	59	no	161	UK	O
petB-e1_1	NC?	Asterales	O	na	na	no	106	na	O
trnM-CAU-rps14-psaB_1	MT c?	Berberidopsidales	O	Malp	56	no	953	UK	O
atpA-1_2	MT c?	Buxales	O	na	na	no	260	na	O
trnR-UCU_2	MT c?	Canellales	O	na	na	no	42	na	O
ycf2-trnI_3	MT	Caryophyllales	O	C	40	no	239	UK	O
atpA-1_1	MT c?	Chloranthales	O	na	na	no	345	na	O
trnV-e2-e1-trnM-CAU_4	MT c?	Cornales	O	Malp	74	no	307	pR	O
trnN-GUU_2	MT	Ericales	O	Malp	52	no	82	UK	O
psbZ_1	MT	Ericales	O	C	18	no	332	UK	O
atpE-atpB_3	NC	Ericales	O	RF	28	no	936	UK	O
ndhF-4_1	MT	Hypocreales	O	Malp	40	no	83	UK	O
trnD-GUC_1	MT	Lamiales	O	C	23	no	73	UK	O
atpA-2_8	MT c?	Liliales	O	na	na	no	73	na	O
petA-2_1	MT	Magnoliales	O	na	na	no	182	na	O
psbD-psbC-trnS-UGA_5	MT	Magnoliales	O	na	na	no	224	na	O
atpA-2_3	MT c?	Magnoliales	O	C	30	no	236	UK	O
rrn5451623_80	NC	Mycoplasmatales	O	RF	100	no	82	pR	O
atpE-atpB_16	MT	NA	O	RF	48	no	71	UK	O
rrn5451623_2	UK	NA	O	RF-C-Fab-Malp	0	no	184	UK	O
rrn5451623_4	UK	NA	O	na	na	no	172	na	O
rrn5451623_10	UK	NA	O	na	na	no	45	na	O
rrn5451623_16	UK	NA	O	na	na	no	45	na	O
rrn5451623_86	UK	NA	O	na	na	no	62	na	O
rrn5451623_7	UK	NA	O	na	na	no	50	na	O
rrn5451623_5	UK	NA	O	na	na	no	47	na	O
rrn5451623_133	BC	NA	O	Fab	100	no	55	pF	O
rrn5451623_39	BC	NA	O	na	na	no	40	na	O
rrn5451623_27	BC	NA	O	na	na	no	94	na	O
rrn5451623_124	MC	NA	O	RF	100	no	59	pR	O
rrn5451623_52	BC	NA	O	RF	100	no	78	pR	O

rrn5451623_58	UK	NA	O	na	na	no	45	na	O
rrn5451623_96	BC	NA	O	Fab	100	no	82	pF	O
rrn5451623_57	BC	NA	O	Fab	100	no	96	pF	O
rrn5451623_138	UK	NA	O	na	na	no	221	na	O
rrn5451623_25	MC	NA	O	na	na	no	104	na	O
trnQ-UUG-1_3	UK	NA	O	Malp	55	no	60	UK	O
atpA-2_11	CP	Petrosaviales	O	Malp	54	no	132	UK	O
rpoC2-8_1	BC	Plasmodiales	O	Malp	35	no	139	UK	O
atpE-atpB_5	MT	Santalales	O	na	na	no	394	na	O
trnI-e2-trnA-e1_2	MT	Santalales	O	RF-C-Fab-Malp	0	no	102	UK	O
rbcL_1	MT c?	Santalales	O	RF-C-Fab-Malp	0	no	1429	UK	O
rpoB_9	MT	Solanales	O	na	na	no	78	na	O
rpoB_10	MT	Solanales	O	Malp	33	no	78	UK	O
rrn5451623_41	NC	Solanales	O	na	na	no	61	na	O
atpH-2_1	CP	Trochodendrales	O	RF	36	no	112	UK	O
psbD-psbC-trnS-UGA_4	MT	Trochodendrales	O	na	na	no	318	na	O
ycf1-11_1	MT	Trochodendrales	O	RF	33	no	211	UK	O
atpA-2_5	MT c?	Asparagales	O	Malp	29	yes	179	UK	O
ycf2-trnI_1	MT c?	Berberidopsidales	O	Malp	34	yes	1178	UK	O
rpoB_4	MT	Buxales	O	Malp	23	yes	181	UK	O
rrn5451623_32	UK	NA	O	C	15	yes	1307	UK	O
atpA-2_4	MT c?	Petrosaviales	O	Malp	60	yes	139	UK	O
atpA-2_2	MT c?	Chloranthales	O	RF-C-Fab	19	yes but 1	265	UK	O
psbD-psbC-trnS-UGA_7	MT c?	Cornales	O	Malp	26	yes but 1	137	UK	O
rps2_1	MT	Trochodendrales	O	RF-C	43	yes but 1	1001	UK	O
trnN-GUU_5	MT	Apiales	pF	na	na	no	68	na	pF
trnH-GUG_1	MT	Asterales	pF	na	na	no	79	na	pF
accD-4_4	CP?	Asterales	pF	RF	66	no	87	UK	pF
ndhI_1	MT c	Asterales	pF	RF	55	no	137	UK	pF
atpE-atpB_7	MT c	Brassicales	pF	na	na	no	167	na	pF
ndhB-e1-rps12-e2-e3_6	MT c?	Caryophyllales	pF	na	na	no	163	na	pF
accD-3_2	MT	Crossosomatales	pF	Fab	64	no	193	UK	pF
rpoB_7	MT	Crossosomatales	pF	Fab	70	no	224	pF	pF
psbD-psbC-trnS-UGA_17	CP	Dipsacales	pF	Fab	76	no	177	pF	pF
rrn5451623_12	MT c?	Ericales	pF	RF	28	no	561	UK	pF
rrn5451623_147	CP	Geraniales	pF	na	na	no	79	na	pF
rrn5451623_47	MT	Laurales	pF	na	na	no	119	na	pF
psaA-3_5	CP	Malpighiales	pF	na	na	no	111	na	pF
psaA-3_2	CP	Malpighiales	pF	Fab	69	no	324	UK	pF
psbD-psbC-trnS-UGA_18	CP	Malpighiales	pF	Fab	42	no	128	UK	pF
rrn5451623_66	MT	Malpighiales	pF	RF	100	no	82	pR	pF
trnM-CAU-rps14-psaB_7	CP	Malpighiales	pF	Malp	10	no	298	UK	pF
accD-3_3	MT c?	Malpighiales	pF	Fab	31	no	54	UK	pF
psbA-1_7	MT c?	Malpighiales	pF	RF	49	no	60	UK	pF
psbB-2_2	MT c?	Malpighiales	pF	na	na	no	108	na	pF
rrn5451623_59	MT c?	Malpighiales	pF	na	na	no	67	na	pF
rrn5451623_48	MT c?	Malpighiales	pF	na	na	no	197	na	pF
ycf3-e1_1	MT c?	Malpighiales	pF	na	na	no	83	na	pF
atpE-atpB_11	CP	Myrtales	pF	Fab	51	no	244	UK	pF

trnN-GUU_4	MT	Myrtales	pF	na	na	no	66	na	pF
trnI-e1_2	CP	NA	pF	Malp	34	no	93	UK	pF
rpl2-rpl23_2	MT	Oxalidales	pF	na	na	no	92	na	pF
rrn5451623_118	CP	Polypodiales	pF	na	na	no	57	na	pF
rrn5451623_148	CP	Polypodiales	pF	na	na	no	52	na	pF
ndhB-e2_2	CP	RosFagales	pF	na	na	no	138	na	pF
psaA-3_4	CP	RosFagales	pF	na	na	no	137	na	pF
rpoC1-e1-rpoB_4	CP	RosFagales	pF	na	na	no	127	na	pF
psbD-psbC-trnS-UGA_8	MT	RosFagales	pF	C	12	no	146	UK	pF
ndhJ-ndhK-ndhC_1	MT c?	RosFagales	pF	Fab	98	no	1602	pF	pF
rbcL_4	MT c?	RosFagales	pF	na	na	no	113	na	pF
psbN-psbH_2	MT	Sapindales	pF	RF	49	no	73	UK	pF
psbD-psbC-trnS-UGA_13	CP	Saxifragales	pF	Malp	44	no	105	UK	pF
trnS-GGA-rps4_2	CP	Saxifragales	pF	Fab	28	no	97	UK	pF
ycf2-trnI_676	CP	Saxifragales	pF	Fab	42	no	255	UK	pF
atpE-atpB_14	CP?	Saxifragales	pF	na	na	no	102	na	pF
psbD-psbC-trnS-UGA_6	MT	Saxifragales	pF	Fab	33	no	149	UK	pF
rpoC2-2_2	MT	Saxifragales	pF	Fab	53	no	130	UK	pF
rrn5451623_37	MT	Solanales	pF	Fab	38	no	208	UK	pF
rrn5451623_53	MT	Solanales	pF	na	na	no	108	na	pF
rrn5451623_44	MT	Solanales	pF	na	na	no	148	na	pF
trnV-e2-e1-trnM-CAU_7	MT	Solanales	pF	na	na	no	62	na	pF
petG_1	MT	Trochodendrales	pF	C	49	no	93	UK	pF
rrn5451623_51	MT	Solanales	pF	Fab	44	yes but 1	134	UK	pF
ndhD-3_3	MT	Brassicales	R	na	na	no	197	na	R
trnV-e2-e1-trnM-CAU_6	MT	Brassicales	R	na	na	no	66	na	R
ndhD-3_2	MT	Brassicales	R	Malp	75	no	93	pR	R
psbJ-psbL-psbF-psbE_2	MT	Geraniales	R	RF-C-Fab-Malp	0	no	160	UK	R
trnL-CAA_1	MT c?	Geraniales	R	Malp	20	no	598	UK	R
atpA-1_3	CP	Malpighiales	R	na	na	no	101	na	R
atpE-atpB_9	MT	Malpighiales	R	na	na	no	68	na	R
atpE-atpB_12	CP	Malpighiales	R	Malp	45	no	197	UK	R
cemA_2	MT	Malpighiales	R	na	na	no	98	na	R
petG-trnW-CAA-trnP-UGG_5	MT	Malpighiales	R	RF-C-Fab	76	no	88	pR	R
petG-trnW-CAA-trnP-UGG_1	MT	Malpighiales	R	C	59	no	207	UK	R
psaA-3_3	CP	Malpighiales	R	Malp	48	no	224	UK	R
rpoB_12	CP	Malpighiales	R	na	na	no	74	na	R
trnT-GGU_1	MT	Malpighiales	R	Malp	100	no	40	pR	R
accD-4_2	MT	Malpighiales	R	Malp	83	no	651	pR	R
rpoB_3	MT	Malpighiales	R	Malp	71	no	598	pR	R
atpE-atpB_6	MT c?	Malpighiales	R	Malp	15	no	200	UK	R
atpE-atpB_4	MT c?	Malpighiales	R	Malp	51	no	703	UK	R
psbA-1_4	MT c?	Malpighiales	R	RF	15	no	106	UK	R
psbB-2_1	MT c?	Malpighiales	R	na	na	no	230	na	R
rpoB_8	MT c?	Malpighiales	R	Malp	21	no	92	UK	R
rpoC1-e1-rpoB_2	MT c?	Malpighiales	R	Malp	32	no	577	UK	R
trnV-e2-e1-trnM-CAU_3	MT c?	Malpighiales	R	Malp	40	no	241	UK	R
trnQ-UUG-1_1	MT	Malvales	R	na	na	no	63	na	R

rpoC2-2_1	MT c	Malvales	R	na	na	no	440	na	R
rpl23-trnI-CAU_2	MT	RosFagales	R	na	na	no	77	na	R
ycf2-trnI_10	MT	RosFagales	R	na	na	no	78	na	R
ccsA-1_1	MT	RosFagales	R	Fab	84	no	834	pF	R
trnL-UAG-ccsA_1	MT	RosFagales	R	Fab	47	no	219	UK	R
ycf4-2_1	MT c?	RosFagales	R	Fab	41	no	133	UK	R
trnM-CAU-rps14-psaB_3	MT	Sapindales	R	RF	67	no	261	UK	R
rpoC1-e1-rpoB_3	CP	Saxifragales	R	Malp	34	no	427	UK	R
cemA_3	MT	Saxifragales	R	Malp	67	no	233	UK	R
orf224-2_1	MT	Saxifragales	R	RF	25	no	146	UK	R
psaA-4_2	MT	Saxifragales	R	Fab	43	no	139	UK	R
trnV-e2-e1-trnM-CAU_2	MT	Saxifragales	R	Fab	49	no	1159	UK	R
rrn5451623_19	MT c?	Saxifragales	R	C	44	no	377	UK	R
atpE-atpB_10	MT	Vitales	R	na	na	no	47	na	R
ndhD-3_1	MT	Vitales	R	Malp	94	no	197	pR	R
petG-trnW-CAA-trnP-UGG_4	MT	Vitales	R	RF	85	no	63	pR	R
atpA-2_9	MT	Vitales	R	na	na	no	127	na	R
trnI-e2-trnA-e1_1	MT	Vitales	R	RF	38	no	318	UK	R
trnR-ACG_1	MT	Vitales	R	na	na	no	411	na	R
ycf3-e2_6	MT	Vitales	R	na	na	no	75	na	R
ycf3-e2-ycf3-e1_2	MT	Vitales	R	na	na	no	98	na	R
rpoB_6	UK	Cucurbitales	C	Malp	77	yes	201	R	R
psbB-1_1	MT c?	Malpighiales	R	Malp	88	yes	363	R	R
accD-1_1	MT	Sapindales	R	RF-C-Fab-Malp	0	yes	237	UK	R
psbA-1_3	MT	Malpighiales	R	Fab	56	yes but 1	484	UK	R
atpF-e1-1_1	MT c?	Malpighiales	R	Malp	83	yes but 1	162	R	R
rpoB_5	MT c?	Malpighiales	R	Malp	43	yes but 1	206	UK	R
ndhB-e1-rps12-e2-e3_1	MT	Malvales	R	Fab	55	yes but 1	965	UK	R
ndhF-2_1	MT c?	Malvales	R	RF-C-Fab-Malp	0	yes but 1	294	UK	R
rpoC1-e2-2_1	MT	RosFagales	R	RF-C-Fab	38	yes but 1	411	UK	R
rpoB_2	MT c?	Saxifragales	R	Malp	91	yes but 1	1319	R	R
petG-trnW-CAA-trnP-UGG_2	MT	Vitales	R	RF-C-Fab	63	yes but 1	262	UK	R
psbJ-psbL-psbF-psbE_1	MT	Vitales	R	RF-C-Fab	10	yes but 1	699	UK	R
rpoB_14	UK	Vitales	R	Malp	14	yes but 1	153	UK	R
atpE-atpB_1	MT	Vitales	R	RF-C-Fab	0	yes but 1	1446	UK	R
psbD-psbC-trnS-UGA_1	MT	Vitales	R	RF-C-Fab	58	yes but 1	1978	UK	R
ndhF-6_1	MT c	Vitales	R	C	59	yes but 1	285	UK	R
rrn5451623_31	MT	Brassicales	pF	Malp	100	no	412	pR	UK
rrn5451623_109	MT	NA	pF	Malp	74	no	86	pR	UK
trnI-e1_1	MT	NA	pF	Malp	77	no	86	pR	UK
atpA-atpF-e2_1	MT	Ericales	O	Fab	98	yes	973	F	UK
rpoC1-e1-rpoB_1	MT	RosFagales	R	Fab	91	yes	1955	F	UK
accD-4_1	MT c?	Cucurbitales	C	Fab	76	yes but 1	679	F	UK
rpl14-rpl16_1	MT c?	Cucurbitales	C	Fab	81	yes but 1	816	F	UK
accD-3_1	MT	Fabales	F	Malp	86	yes but 1	307	R	UK
rpl2-rpl23_1	MT c?	Saxifragales	R	Fab	84	yes but 1	1193	F	UK

Table S4. Genes and RNAs present or absent in *P. aethiopica* and *P. hamiltonii* relative to their autotrophic relative *Cucumis sativus*, and size comparisons as an assessment of the functionality of protein-coding genes.

Gene	Length in *Cucumis sativus* (bp)	Length in *Cucumis sativus* (aa)	2/3 of length (aa)	Present in *P. aethiopica*	Present in *P. hamiltonii*	*P. aethiopica* longest ORF(aa)	*P. hamiltonii* longest ORF (aa)
accD	1,578	526	351	yes	yes	52	136
atpA	1,521	507	338	yes	no	86	0
atpB	1,497	499	333	yes	yes	127	27
atpE	411	137	91	yes	no	83	0
atpF-e1	144	48	32	no	no	0	0
atpF-e2	408	136	91	yes	no	38	0
atpH	246	82	55	yes	yes	31	24
atpI	744	248	165	yes	no	77	0
ccsA	969	323	215	yes	no	91	0
cemA	690	230	153	yes	yes	15	58
clpP-e1	69	23	15	no	yes	0	0
clpP-e2	291	97	65	no	yes	0	44
clpP-e3	225	75	50	no	no	0	0
matK	1,539	513	342	no	yes	0	99
ndhA-e1	540	180	120	no	yes	0	78
ndhA-e2	537	179	119	no	yes	0	0
ndhB-e1	777	259	173	no	no	0	0
ndhB-e2	756	252	168	no	no	0	0
ndhC	363	121	81	no	yes	0	43
ndhD	1,503	501	334	no	no	0	0
ndhE	306	102	68	no	no	0	0
ndhF	2,247	749	499	no	no	0	0
ndhG	531	177	118	no	no	0	0
ndhH	1,182	394	263	no	yes	0	65
ndhI	483	161	107	no	no	0	0
ndhJ	477	159	106	no	yes	0	48
ndhK	681	227	151	no	yes	0	119
orf224	672	224	149	yes	yes	39	69
petA	957	319	213	yes	yes	43	21
petB-e1	6	2	1	yes	yes	1	0
petB-e2	642	214	143	no	yes	0	25
petD-e1	6	2	1	no	yes	0	0
petD-e2	474	158	105	no	yes	0	47
petG	114	38	25	yes	no	26	0
petL	96	32	21	no	no	0	0
petN	90	30	20	no	no	0	0
psaA	2,253	751	501	yes	yes	80	78
psaB	2,205	735	490	yes	no	88	0

psaC	246	82	55	no	yes	0	36
psaI	114	38	25	no	no	0	0
psaJ	171	57	38	no	no	0	0
psbA	1,062	354	236	yes	yes	78	96
psbB	1,527	509	339	no	yes	0	121
psbC	1,386	462	308	yes	no	82	0
psbD	1,062	354	236	yes	yes	126	122
psbE	252	84	56	yes	no	51	0
psbF	120	40	27	no	no	0	0
psbH	222	74	49	no	yes	0	42
psbI	111	37	25	no	no	0	0
psbJ	123	41	27	no	no	0	0
psbK	186	62	41	no	no	0	0
psbL	117	39	26	no	no	0	0
psbM	105	35	23	no	no	0	0
psbN	132	44	29	yes	yes	30	23
psbT	114	38	25	no	no	0	0
psbZ	189	63	42	yes	no	62	0
rbcL	1,431	477	318	no	yes	0	142
rpl14	369	123	82	no	no	0	0
rpl16-e1	9	3	2	no	no	0	0
rpl16-e2	399	133	89	no	no	0	0
rpl20	354	118	79	no	no	0	0
rpl23	282	94	63	yes	no	31	0
rpl2-e1	390	130	87	no	no	0	0
rpl2-e2	432	144	96	no	no	0	0
rpl32	165	55	37	no	no	0	0
rpl33	198	66	44	no	no	0	0
rpl36	114	38	25	no	yes	0	37
rpoA	1,002	334	223	no	yes	0	85
rpoB	3,216	1072	715	yes	yes	119	109
rpoC1-e1	432	144	96	yes	yes	0	107
rpoC1-e2	1,611	537	358	yes	yes	69	0
rpoC2	4,173	1391	927	yes	yes	23	132
rps11	417	139	93	yes	yes	19	88
rps12-e1	114	38	25	no	no	0	0
rps12-e2	231	77	51	yes	yes	43	17
rps12-e3	24	8	5	no	no	0	0
rps14	303	101	67	no	no	0	0
rps15	273	91	61	no	yes	0	52
rps16-e1	39	13	9	no	no	0	0
rps16-e2	213	71	47	no	no	0	0
rps18	315	105	70	no	yes	0	16
rps19	279	93	62	no	no	0	0
rps2	711	237	158	yes	no	133	0
rps22	489	163	109	no	no	0	0
rps3	657	219	146	no	no	0	0
rps4	606	202	135	no	no	0	0
rps7	468	156	104	no	no	0	0
rps8	405	135	90	no	no	0	0

rrn16	1,492	na	na	no	yes?	na	na
rrn23	2,810	na	na	no	no	na	na
rrn4.5	103	na	na	no	no	na	na
rrn5	121	na	na	no	no	na	na
trnA-UGC-e1	38	na	na	no	no	na	na
trnA-UGC-e2	35	na	na	no	yes	na	na
trnC-GCA	72	na	na	no	no	na	na
trnD-GUC	74	na	na	no	no	na	na
trnE-UUC	73	na	na	no	yes	na	na
trnF-GAA	73	na	na	no	yes	na	na
trnG-GCC	71	na	na	no	no	na	na
trnG-UCC-e1	48	na	na	no	no	na	na
trnG-UCC-e2	23	na	na	no	no	na	na
trnH-GUG	75	na	na	no	no	na	na
trnI-CAU	74	na	na	no	no	na	na
trnI-GAU-e1	37	na	na	no	no	na	na
trnI-GAU-e2	35	na	na	yes	no	na	na
trnK-UUU-e1	37	na	na	no	no	na	na
trnK-UUU-e2	35	na	na	no	yes	na	na
trnL-CAA	81	na	na	no	no	na	na
trnL-UAA-e1	35	na	na	no	no	na	na
trnL-UAA-e2	50	na	na	no	no	na	na
trnL-UAG	80	na	na	yes	no	na	na
trnM-CAU	74	na	na	no	no	na	na
trnM-CAU	73	na	na	yes	no	na	na
trnN-GUU	72	na	na	no	yes	na	na
trnP-UGG	74	na	na	no	no	na	na
trnQ-UUG	73	na	na	yes	no	na	na
trnR-ACG	74	na	na	no	no	na	na
trnR-UCU	72	na	na	no	no	na	na
trnS-GCU	88	na	na	no	no	na	na
trnS-GGA	85	na	na	no	no	na	na
trnS-UGA	93	na	na	yes	no	na	na
trnT-GGU	72	na	na	no	yes	na	na
trnT-UGU	73	na	na	no	no	na	na
trnV-GAC	72	na	na	no	no	na	na
trnV-UAC-e1	41	na	na	yes	no	na	na
trnV-UAC-e2	35	na	na	yes	no	na	na
trnW-CAA	74	na	na	yes	no	na	na
trnY-GUA	84	na	na	no	yes	na	na
ycf1	5,640	1880	1253	yes	yes	37	63
ycf2	5,973	1991	1327	yes	yes	39	65
ycf3-e1	126	42	28	no	no	0	0
ycf3-e2	228	76	51	no	no	0	0
ycf3-e3	153	51	34	no	no	0	0
ycf4	555	185	123	no	yes	0	148

Chapter 6

HORIZONTAL GENE TRANSFER IN EUKARYOTES, FOCUSING ON FUNGI-TO-PLANT AND PLANT-TO-PLANT TRANSFERS.

Renner S. S. and S. Bellot

Pp. 223-235 in R. Bock & V. Knoop, eds.,
Genomics of Chloroplasts and Mitochondria.
Springer, Heidelberg.

Chapter 10

Horizontal Gene Transfer in Eukaryotes: Fungi-to-Plant and Plant-to-Plant Transfers of Organellar DNA

Susanne S. Renner* and Sidonie Bellot
Systematic Botany and Mycology, University of Munich, Menzinger Straße 67, D-80638 Munich, Germany

Summary

This review focuses on horizontal gene transfer (HGT) involving bacteria, fungi, and plants (Viridiplantae). It highlights in particular the persistent challenge of recognizing HGT, which requires a combination of methods from bioinformatics, phylogenetics, and molecular biology. Non-phylogenetic methods rely on compositional structure, such as G/C content, dinucleotide frequencies, codon usage biases, or co-conversion tracts, while phylogenetic methods rely on incongruence among gene trees, one of which is taken to represent the true organismal phylogeny. All methods are handicapped by short sequence lengths with limited or highly uneven substitution signal; the statistical problems of working with taxon-rich alignments of such sequences include low support for inferred relationships, and difficult orthology assessment. Plant-to-plant HGT is known from two dozen mitochondrial genes and species of phylogenetically and geographically widely separated ferns, gymnosperms, and angiosperms, with seven cases involving parasitic plants. Only one nuclear HGT has

*Author for correspondence, e-mail: renner@lrz.uni-muenchen.de

R. Bock and V. Knoop (eds.), Genomics of Chloroplasts and Mitochondria, Advances in Photosynthesis and Respiration 35, pp. 223–235,
DOI 10.1007/978-94-007-2920-9_10,

come to light, and extremely few fungi-to-plant transfers. Plant mitochondrial genomes, especially in tracheophytes, are prone to take up foreign DNA, but evolutionary consequences of this are still unclear.

I. Introduction

Horizontal gene transfer (HGT) refers to movement of genetic material between organisms that does not follow the normal pathway of vertical transmission from parent to offspring. Horizontal gene transfer is sometimes seen as synonymous with lateral gene transfer, a term better restricted to *within*-species sequence copying, such as group II intron retrotransposition or the massive migration of promiscuous cpDNA into mitochondria of seed plants. With the 2003 discoveries of HGT involving eukaryotes (Bergthorsson et al. 2003; Won and Renner 2003), the availability of full genome sequences, and new insights into transposable elements, HGT has become an important issue also in plant science. Recent reviews of the topic include those of Andersson (2005), Richardson and Palmer (2007), Keeling and Palmer (2008), Keeling (2009a, b), and Bock (2010), and the paradigm is rapidly becoming that HGT is "a highly significant process in eukaryotic genome evolution" (Bock 2010).

The present review focuses on glaucophytes, red algae, green algae, and land plants. Besides briefly summarizing recent findings relevant to plant genomes, it will highlight the persistent challenge of recognizing horizontal gene transfer. This challenge stems largely from the still relatively crude methods for finding matching DNA strings in databases and the inability of phylogenetic algorithms to infer correct relationships from short sequences. Especially the latter problem is often underappreciated in the context of HGT. We therefore begin our review by discussing the combination of bioinformatics, phylogenetics, and molecular biology that forms the basis for inferring and evaluating HGT. We then discuss the evidence for gene transfer between bacteria or fungi and plants, plant-to-plant transfer, and transposable element transfer, and follow with a section on problematic or erroneous earlier inferences of HGT. We end by addressing what is known about the mechanisms of HGT among plants and by providing a perspective on ongoing research that aims at unsolved questions in HGT.

Abbreviations: BLAST – Basic local alignment search tool; cpDNA; – Plastid DNA; DNA – Deoxyribonucleic acid; EST – Expressed sequence tag; HGT – Horizontal gene transfer; HTT – Horizontal transposon transfer; mt(DNA) – Mitochondrial (DNA); MULE – *Mu*-like elements (Mu is *mutator* in corn); My – Million years; ORF – Open reading frame; PCR – Polymerase chain reaction; RNA – Ribonucleic acid; T-DNA – Transferred DNA; TE – Transposable element; Ti-plasmid – Tumor-inducing plasmid

II. Detecting and Evaluating Cases of Horizontal Gene Transfer

A. Bioinformatic Approaches for Detecting HGT

Genome-wide studies of eukaryotes typically will involve a BLAST search (Altschul et al. 1990) to identify genes matching bacterial genes or to find unusual (unique) genes that could be of bacterial origin. Another step is to employ known genes as queries and test for consistency of ORFs or to BLAST against a local database containing well-annotated genomic sequences from model organisms. All these steps rely on BLAST results. It is well understood, however, that BLAST e-values are based on the expected background noise, depend on the sequences in the database at any one time, and are not a reliable indicator of evolutionary relatedness (Koski and Golding 2001). Recent genomics studies have used pair-wise syntenic alignments and BLAST score statistical tests (e.g., Ma et al. 2010).

Other non-phylogenetic methods depend on compositional structure, such as G/C content, dinucleotide frequencies or codon usage biases, but the length of a horizontally transferred gene may be too short to reliably reveal these differences. Methods based on atypical nucleotide or amino acid composition also may only detect recent transfers because donor sequence characteristics will gradually become erased. Moreover, the reliability of these methods is difficult to assess statistically (Ragan et al. 2006). Snir and Trifonov (2010) have proposed using an additional approach that involves comparing just two genomes. With two genomes of a given length one can calculate the probabilities of identical regions (under a chosen model of substitution). To detect HGTs, the method makes use of the expectation that the flanking regions of an inserted region will normally be non-homologous and then uses a sliding window algorithm to detect these HGT borders, essentially searching for sharp borders (or walls). The method has been applied to simulated data and real bacterial genomes.

B. Phylogenetic Approaches for Detecting HGT

Phylogenetic trees are time-consuming to construct because they require a trustworthy sequence alignment. Nevertheless, many workers consider phylogenetic tree incongruence the best indicator of HGT, perhaps especially ancient HGT. When conflicts are found between two or more gene trees, HGT can be introduced as one possible explanation (for an insightful discussion concerning tree incongruency due to HGT in the microbial world, see Boto 2010). Like the bioinformatics approaches discussed in the previous section, the phylogenetic method for identifying HGT faces several challenges. First, it is incapable of coping with events residing in non-homologous regions since all tree inference methods presume character homology in the underlying sequence alignment. It also requires assumptions about where to seek the HGT events, in other words, assumptions about which tree reflects the true organismal history. There is reason to think that methods that detect HGT using atypical genomic composition ("signatures") are better at finding recent transfers whereas "phylogenetic incongruence" methods may be better at detecting older HGTs because of the increasing mutational signal over time, until saturation (Ragan et al. 2006; Cohen and Pupko 2010). Whether this generalization holds will depend on details of the substitution process since all phylogenetic methods, whether parsimony, maximum likelihood, or Bayesian inference, require sufficient mutational signal.

The statistical cut-off deemed acceptable for particular splits in a tree is a matter of debate. Among phylogeneticists, accepted cut-offs values are >75% under parsimony and likelihood optimization, and 98% under Bayesian tree sampling, values rarely reached in trees used to infer HGT because of taxon-rich alignments and short sequences. A sense of the amount of signal needed for statistical support can be gained from Felsenstein's (1985) demonstration that three non-homoplastic substitutions suffice for a bootstrap support (for a node) at the 95% level. These statistical reasons imply that well-supported phylogenies usually require concatenated multi-locus alignments. One then faces the question of which loci can safely be combined. For plants, one solution has been to accept combined plastid gene phylogenies as "true" and to view phylogenies from mitochondrial genes as HGT-prone (Cho et al. 1989a, b; Bergthorsson et al. 2003; Burger et al. 2003; Hao et al. 2010; Archibald and Richards 2010; compare Sect. VII). This is based on the rationale that no evidence has so far come to light of HGT involving plastid genes of Viridiplantae.

Statistical tests for tree incongruence, such as the Incongruence Length Difference test (Farris et al. 1994), require sufficient mutational signal and usually cannot reliably identify nodes in phylogenies due to HGT as long as the trees are based on single genes. This leaves workers in a bind, and many HGT studies have therefore inferred incongruence by eyeballing more or less unsupported trees or by contrasting an unresolved

gene tree with an organismal tree supported by other evidence, for example, morphological and/or genetic data analyzed in other studies. A software to detect HGT from tree incongruence alone is SPRIT (Hill et al. 2010), but it requires assuming that all splits in the trees being compared are true.

A second difficulty with phylogenetic approaches for detecting HGT is that gene phylogenies may be incongruent because of biases in the sequence data and not (only) because of HGT. Well known biases include uneven nucleotide frequencies (Embley et al. 1993; Foster et al. 2009; Stiller 2011), long-branch attraction (Felsenstein 1978), codon bias, and model over-parameterization. Long branch attraction is a systematic error, corresponding to the inconsistency of a statistical procedure (namely maximum parsimony), and leads to the convergence towards an incorrect answer as more and more data are analyzed. It occurs when two (or more) sequences in a phylogeny have unusually high substitution rates, resulting in their having much longer branches than the remaining sequences. Long-branch attraction cannot be resolved by adding more characters, and it is a severe and under-appreciated problem in HGT detection. (Removing one of the long branches can sometimes eliminate the problem; e.g., Goremykin et al. 2009).

A third difficulty in identifying HGT is to distinguish it from ancestral gene duplication and differential gene loss (Stanhope et al. 2001; Gogarten and Townsend 2005; Noble et al. 2007). Duplication and loss in gene families affects especially nuclear genes, and since relatively few densely sampled and deep (i.e., going back millions of years) phylogenies have been built with nuclear genes, lineage sorting has so far not been a major discussion point in HGT (but see Noble et al. 2007).

A recent study involving fungi and angiosperms, illustrates the problems of detecting HGT. To test for plant/fungi gene exchange, Richards et al. (2009) generated automated gene-by-gene alignments and phylogenies for 4,866 genes identified in analyses of the *Oryza* genome and in BLAST comparisons. Visual inspection of the phylogenies used two criteria for HGT: Either a plant gene sequence branching within a cluster of sequences from fungal taxa (or vice versa) or a phylogeny that demonstrated a diverse plant-specific gene family absent from all other taxa except a narrow taxonomic group of fungi (or vice versa). Using these criteria, Richards et al. detected 38 plant-fungi HGT candidates, of which two were detected using the rice genome-specific analysis, 35 were detected using the BLAST-based survey, and one was detected using both search protocols. However, when these authors added more sequences (taxa) from GenBank and expressed sequence tag (EST) databases, only 14 of the putative HGTs remained because increasing taxon sampling decreased the number of isolated or wrongly placed suspected HGT sequences. The number of suspected HGT events was then further reduced to nine by reconstructing phylogenies with better fitting maximum likelihood substitution models that accounted for rate heterogeneity. The study beautifully illustrates the risk of overestimating the frequency of HGT from insufficient taxon sampling and poorly fitting substitution models, with rate heterogeneity being the single most important model parameter (Yang 1994).

As is generally true for tree inference, also the dynamics of gene gains and losses in gene families are probably better inferred using maximum likelihood than parsimony optimization of the minimal number of gains and losses needed to explain the distribution of a group of orthologous genes in a phylogeny (Mirkin et al. 2003; Richards et al. 2009; Cohen and Pupko 2010). These and other studies (Cusimano et al. 2008; Goremykin et al. 2009; Ragan and Beiko 2009; Ferandon et al. 2010) all caution against inferring rampant HGT from phylogenetic incongruence among gene trees, at least as long as the trees are based on short sequences (analyzed under parsimony or, worse, neighbor-joining) from genetically distant organisms with millions of years of evolution separating them.

C. Footprints and Signatures of HGT

The third way of identifying HGT is to look for signatures or "footprints" of the HGT events themselves (Adams et al. 1998; Cho

et al. 1998; Cho and Palmer 1999; Sanchez-Puerta et al. 2008). Such footprints might be the co-conversion tracts of group I introns, which are short stretches of flanking exon sequence (>50 bp into the 5′ exon and <25 bp into the 3′ exon) that may be converted to the donor DNA sequence during intron insertion or excision (Lambowitz and Belfort 1993; Lambowitz and Zimmerly 2004). If the flanking exon stretches in the donor and recipient differ, then co-conversion will create a footprint that can stay even after the intron itself is lost again. The first study using the molecular footprint approach focused on a group I intron in the mt *cox1* gene and inferred 3–5 HGT events in a small clade of Araceae (Cho and Palmer 1999). This was inferred although a parsimony reconstruction favored a vertical transmission history with one intron gain, followed by two losses, that is, three evolutionary events, rather than five (Cho and Palmer 1999). Subsequently, reliance on co-conversion tracts as inconvertible footprints led to the extrapolation of at least 1,000 HGTs of the *cox1* intron among living angiosperms, based on a survey of the intron's distribution that suggested 32 separate cases of intron acquisition from unknown donors to account for the intron's presence in 48 of 281 species from 278 genera (Cho et al. 1998b).

Seven cases of chimeric sequences between foreign and native mt gene copies have been described (Vaughn et al. 1995; Adams et al. 1998: *Peperomia polybotrya cox1* intron; Bergthorsson et al. 2003: *Sanguinaria canadensis rps11*; Barkman et al. 2007: *Pilostyles thurberi atp1*; Hao et al. 2010: *Ternstroemia stahlii atp1*; *Hedychium coronarium matR*; *Boesenbergia rotunda matR*; Mower et al. 2010: *Plantago macrorhiza atp1*). In some cases, the chimeric sequences appear functional, in others they are unexpressed pseudogenes. A recent re-analysis of these cases based on a new recombination search algorithm developed specifically for plant mitochondrial genomes showed that detecting HGT-generated chimeras requires dense taxonomic sampling (Hao 2010). Using the new algorithm, Hao and Palmer (2009) also identified nine putative cases of short-patch gene conversion of native, functional plant mt *atp1* genes by homologous *atpA* genes of chloroplast origin. If confirmed, these cases of recombination between mitochondrial and chloroplast genes provide unique evidence for the creation of functional chimeric genes across the ca. one-billion-year divide between chloroplast and mitochondrial genes.

For transposable elements (TEs), detection of horizontal transposon transfer (HTT) may sometimes be possible by comparisons of the rates of synonymous substitution, the Ks values, observed in TEs with those in orthologous genes (Sanchez-Gracia et al. 2005; Schaack et al. 2010). If the presence of a TE in two hosts is due to horizontal transfer, then it will be younger than the hosts and will have accumulated fewer synonymous mutations than the host genes. With many complete genome sequences now available, this approach can be implemented in a robust statistical framework taking into account the Ks value distribution of hundreds of host genes to define the Ks threshold under which the presence of a TE is considered to be the result of HTT. The approach has been applied to closely related species, such as *Drosophila melanogaster* and *Drosophila simulans*, which diverged less than 5 My ago (Schaack et al. 2010).

III. DNA Transfers Among Bacteria or Fungi and Plants

The classic example of HGT from prokaryotes to multi-cellular eukaryotes is the transfer of DNA from the *Agrobacterium* Ti plasmid to plants (reviewed by Gelvin 2009). Other bacterial species, such as *Sinorhizobium meliloti* and *Mesorhizobium loti*, when harboring modified Ti plasmids, can also transfer them to plants (Broothaerts et al. 2005). During transformation, the transferred DNA (T-DNA) is moved through the plasma membrane via a channel formed by a bacterial protein that also participates in coating of the T-DNA during its transfer to the nucleus (Dumas et al. 2001). The extent of natural recent incorporation of prokaryotic genetic material into plants is

unclear, although bacterial chromosomal DNA apparently is introduced into the nuclei of transgenic plants occasionally (Ülker et al. 2008).

So far, there is one reported case of the horizontal acquisition of a group II intron in the plastid *psbA* gene of the green alga *Chlamydomonas sp.* that appears to come from a cyanobacterium (Odom et al. 2004). From red algae, two genes, *rpl36* and an unusual rubisco operon, *rbcLS*, may have been transferred from bacterial donors to the common ancestor of red algae (*rbcLS*) or the common ancestor of cryptophytes and haptophytes (*rpl36*) (Keeling and Palmer 2008, and references therein).

Genetic exchange between plants and fungi is exceedingly rare, particularly in angiosperms (Richards et al. 2009). Richards et al. compared the genomes of six plant species (*Arabidopsis thaliana*, *Populus trichocarpa*, *Sorghum bicolor*, *Oryza sativa*, *Selaginella moellendorffii*, and *Physcomitrella patens*) with those of 159 prokaryotes and non-plant eukaryotes. Comprehensive phylogenetic analyses of the data, using methods that account for site-specific substitution rate heterogeneity, supported only nine HGTs between plants and fungi (methods used in this study were discussed above in Sect. II.B). Five were fungi-to-bryophyte and fungi-lycophyte transfers and four were plant-to-fungi transfers. An older report of the transfer of a group I intron from the angiosperm *Youngia japonica* (Asteraceae) into the 18S rRNA of its pathogenic fungus *Protomyces inouyei* (Nishida and Sugiyama 1995) has yet to be followed-up.

IV. Plant-to-Plant DNA Transfers

Exchange of genetic material between mitochondria of land plants has been inferred for diverse taxa. The species involved come from phylogenetically and geographically widely separate clades of ferns, gymnosperms, and angiosperms, suggesting that HGT among plants may be relatively widespread. The known cases involve the following mitochondrial sequences and taxa:

- The *rps2* gene in the dicot *Actinidia arguta* coming from a monocot (Bergthorsson et al. 2003),
- *rps11* in an unidentified *Lonicera* (Caprifoliaceae) coming from Ranunculaceae/Berberidaceae; in the dicot *Sanguinaria canadensis* from a monocot; and in two unidentified Betulaceae from an unidentified non-Betulaceae donor (Bergthorsson et al. 2003),
- *atp1* in *Amborella trichopoda* (Amborellaceae) from an unknown Asteridae (Bergthorsson et al. 2003); in *Ternstroemia* (Pentaphylaceae) from Ericaceae, and in *Bruinsmia* (Styracaceae) from Cyrillaceae (Schönenberger et al. 2005),
- The *nad1* second intron in *Gnetum* (Gymnospermae) coming from an unknown Asteridae, that is, a flowering plant (Won and Renner 2003),
- The *nad1* second intron plus *atp1* in two parasitic species of Rafflesiaceae from their respective host plants (Davis and Wurdack 2004; Barkman et al. 2007),
- The same intron plus *matR* in the fern *Botrychium virginianum* from an unknown Loranthaceae root-parasite (Davis et al. 2005),
- *atp1* in *Pilostyles thurberi* (Apodanthaceae) from its legume host, *Psorothamnus emoryi*; in *Mitrastema yamamotoi* (Mitrastemonaceae) from its host *Quercus subsericea* (Fagaceae; Barkman et al. 2007), and
- *atp1*, *atp6* and *matR* in species of *Plantago* (Plantaginaceae) from parasitic *Cuscuta* (Convolvulaceae) and *Bartsia* (Orobanchaceae; Mower et al. 2004, 2010).

The transferred mitochondrial genes appear to sit in the hosts' mitochondrial genomes, and most are non-functional pseudogenes. Seven cases of chimeric sequences between foreign and native mt gene copies (see especially Mower et al. 2010) were already discussed above (Sect. II). The putative HGT of the mitochondrial *cox1* intron across thousands of flowering plants, either from plant to plant or via unknown fungal donors (Adams et al. 1998; Cho et al. 1998b; Cho and Palmer 1999; Sanchez-Puerta et al. 2008) is discussed in Sects. II and VI.

An additional report about mitochondrial HGT on a massive scale involves the basal angiosperm *Amborella*, which may have acquired one or more copies of 26 mitochondrial protein genes from other land plants. Twenty foreign gene sequences appear to come from other angiosperms, six from moss donors. The transferred genes seem to be intact, but have not been shown to be functional (Bergthorsson et al. 2004). The report has attracted criticism (Martin 2005; Goremykin et al. 2009; see also Sect. VI). Large-scale genome sequencing of *Amborella* is ongoing and may resolve the controversy.

A single HGT event probably can involve multiple mitochondrial genes as made plausible by the results for *Cuscuta* and *Plantago* of Mower et al. (2010). This study also suggests a complicated history of the transferred genes within *Plantago* subsequent to their acquisition via HGT, with additional transfers (including intracellular transfer), gene duplication and differential loss and mutation-rate variation (Mower et al. 2010). Resolving this history will probably require complete mitochondrial and nuclear genome sequencing from multiple individuals.

So far, only one nuclear plant-plant HGT event has come to light. It involves the parasitic Orobanchaceae *Striga hermonthica,* for which BLAST searches between an EST database of *Striga* and plant genome databases, sequencing of a 6,423 bp-long genomic region and Southern blotting collectively imply recent uptake of genetic material from an unknown monocot host (Yoshida et al. 2010). The transferred gene encodes a 448 amino acid-long protein of unknown function, is phylogenetically closer to *Sorghum* than to its *Brachypodium* ortholog, and was acquired recently, that is, after the divergence between *Striga* and *Orobanche* (both in Orobanchaceae) but before the divergence of *S. hermonthica* and *S. gesnerioides*.

From the above it emerges that most plant-to-plant HGT events involve mitochondrial DNA and that close physical association, as exists, for example, between parasitic plants and their hosts, apparently facilitates plant-to-plant HGT. See Sect. VII for possible reasons why plant mitochondria may incorporate foreign DNA more readily than other genomes.

V. Transposable Elements

There are some 200 putative cases of transposable elements (TEs) moving horizontally in eukaryotes, but such events appear to be rare among plants. The first report of the horizontal transfer of a nuclear TE between plants was that of a *Mutator*-like element between the plant genera *Setaria* and *Oryza* (Diao et al. 2006). For clades other than Viridiplantae, it has been argued that introduction of transposable elements by horizontal transfer in eukaryotic genomes has been a major force propelling genomic variation and biological innovation (Sanchez-Gracia et al. 2005; Gilbert et al. 2010; Schaack et al. 2010). Whether there is any correlation between the horizontal transfer of TEs and the horizontal transfer of functional genes is unclear. Although TEs have not yet been shown to transfer host genes between different species in eukaryotes, they are capable of capturing and transducing sequences at high frequency within a species (Schaack et al. 2010). Of 3,000 analyzed TEs in rice, many contained gene fragments of genomic DNA that apparently had been captured, rearranged and amplified over millions of years (Jiang et al. 2004). Other examples of gene duplication and exon shuffling by transposons come from *Zea mays* (Morgante et al. 2005).

VI. Problematic, Controversial, and Erroneous Reports of HGT Involving Plants

Claims of HGT require considerable supporting evidence and caution (Kurland et al. 2003; Martin 2005; Richards et al. 2009), with a case in point being the problems with

the early reports of massive HGT in the draft human genome (Lander et al. 2001) and their later dismissal (Salzberg et al. 2001; Stanhope et al. 2001). It is therefore not surprising that a few reports of HGT have been discussed controversially or turned out to be erroneous. Thus, the report of HGT between unknown Malvaceae and the parasitic species *Pilostyles thurberi* (Nickrent et al. 2004), after re-sequencing of the relevant gene region (18S RNA), turned out to be due to contaminated DNA sequences (Filipowicz and Renner 2010).

An example of putative HGTs being discussed critically is the mt *cox1* intron, which occurs in hundreds of species of flowering plants (Vaughn et al. 1995; Cho et al. 1998a, b; Cho and Palmer 1999; Cusimano et al. 2008; Sanchez-Puerta et al. 2008, 2011). Phylogenetic analysis of the *cox1* intron does not result in statistically supported trees because the intron contains too few phylogenetically informative mutations (Cusimano et al. 2008: sequence similarity among 110 *cox1* introns from throughout angiosperms ranges from 91% to 100%). Even so, the *cox1* tree for the flowering plants matches accepted relationships of orders, families and, in a few cases, genera (Cusimano et al. 2008). A parsimonious explanation is that the *cox1* intron was horizontally acquired once or a few times during the history of flowering plants, followed by vertical inheritance and numerous losses (Cusimano et al. 2008; also Ragan and Beiko 2009; Richards et al. 2009; Inda et al. 2010; Ferandon et al. 2010). Distinct mutations in co-conversion tracts, however, can lead to a scenario of intron insertions from hundreds or thousands of unknown fungal donors (Cho et al. 1998b; Sanchez-Puerta et al. 2008; fungi-to-angiosperm gene transfers are otherwise exceedingly rare: Richards et al. 2009). Resolving the issue will require a better understanding of the mechanisms of intron homing, specifically the creation and decay of co-conversion tracts (Wolf et al. 2001; Belshaw and Bensasson 2006; Ragan and Beiko 2009).

The controversy surrounding Bergthorsson et al.'s report (2004) of rampant HGT of the mtDNA of *Amborella trichopoda* has already been mentioned (Martin 2005; Goremykin et al. 2009). It is clear also from the difficult interpretation of the history of the elongation factor genes in the green algal lineage (Noble et al. 2007; Rogers et al. 2007) that greater taxon sampling can sometimes lead to a scenario more consistent with multiple losses than horizontal gains. Both processes are likely to have played important roles, and knowledge of the function of putatively transferred genes and of the biology of the involved species should help formulate testable hypotheses.

VII. Mechanisms of Plant-to-Plant HGT

The means of DNA exchange between unrelated organisms could theoretically be (1) vectors, such as bacteria, fungi or phloem-sucking bugs; (2) transfer of entire mitochondria through plasmodesmata, when there is plant-to-plant contact; (3) illegitimate pollination followed by elimination of most foreign DNA except for a few mitochondria that might fuse with native mitochondria (below) or (4) natural transformation. Of the 10–36 cases of plant-to-plant HGT (listed in Sect. IV; the numerical range depends on whether the 26 *Amborella* mt genes putatively taken up from other flowering plants and mosses are included; Bergthorsson et al. 2004), at least seven involve parasitic plants (namely Apodanthaceae: *Pilostyles*; Convolvulaceae: *Cuscuta*; unknown root-parasitic Loranthaceae; Mitrastemonaceae: *Mitrastema*; Orobanchaceae: *Bartsia*, *Striga*, *Orobanche*, *Phelipanche*; the common ancestor of the Rafflesiaceae). This ratio suggests that direct contact between donor and recipient facilitates HGT. The host plants can be the donor (Mower et al. 2004, 2010; Davis et al. 2005) or the recipient (Davis and Wurdack 2004; Barkman et al. 2007; Yoshida et al. 2010). The apparent high frequency of HGT involving parasitic plants fits with the experimental demonstration of DNA moving through a graft junction between different lines of tobacco (Stegemann and Bock 2009, although the transferred DNA stayed in the

graft zone). That messenger RNA can pass through plasmodesmata is well documented (Roney et al. 2007; Lucas et al. 2009), but whether paired DNA or entire organelles can pass through plasmodesmata remains to be investigated. Alternatively, vesicle transport of DNA or organelles from cell to cell could be involved in the horizontal transfer of genetic material (Bock 2010).

All but one of the known plant-to-plant HGTs involve mitochondrial DNA, the exception being the nuclear gene taken up by *Striga hermonthica* probably from a monocot host (Yoshida et al. 2010). The propensity of plant mitochondria to incorporate foreign DNA is remarkable, since among thousands of animal mitochondrial genomes sequenced, no convincing evidence of HGT has been found, and embryophyte (land plant) plastid genomes also apparently are devoid of horizontally transferred foreign DNA. So why are plant mitochondrial genomes so open towards foreign DNA? One explanation may be that plant mitochondria are capable of importing RNA and double-stranded DNA (Koulintchenko et al. 2003). Another explanation may be the great propensity of plant mitochondria to fuse with one another (Arimura et al. 2004; Sheahan et al. 2005) and the high recombinational activity of mtDNA throughout tracheophyte evolution (Grewe et al. 2009; Hecht et al. 2011). This may have set the stage for the integration of foreign DNA in plant mt genomes, also amply documented by the frequent integration of chloroplast DNA laterally transferred into seed plant mtDNAs. Interestingly, bryophyte mt genomes lacking similarly active DNA recombination may be sources, but not acceptors for HGTs (Knoop et al. 2011; p. 18).

It is not known whether the horizontally transferred genetic material is DNA or RNA. While it was earlier hypothesized that mitochondrial HGT might largely be an RNA-mediated process (Bergthorsson et al. 2003), transfer of double-stranded DNA, which is much more stable, may be more likely (Henze and Martin 2001; Mower et al. 2010). Whether the transferred mtDNA tends to integrate into the recipients' mitochondrial genomes or, instead, becomes transferred to the nucleus is mostly unclear (Martin 2005; Goremykin et al. 2009; Hao et al. 2010). Keeling and Palmer (2008) have suggested that most transferred genes probably are non-functional and coexist with a native, functional homologue.

In addition to the barriers that can prevent the horizontal transfer and integration of foreign DNA in a recipient, it is worth considering the barriers that prevent its spread in a population. In prokaryotes, and probably also in eukaryotes, one such barrier can be the perturbation of gene dosage and expression in the host. An experimental study of the transferability of thousands of genes within *Escherichia coli* by Sorek at al. (2007) showed that toxicity to the host and changed (increased) gene dosage and expression probably are predominant causes for transfer failure. On the other hand, over-expression of an RNA polymerase experimentally transferred from *Bacillus subtilis* to *E. coli* appeared to entail no immediate fitness costs (Omer et al. 2010).

VIII. Perspective

There are many unsolved questions regarding the transfer of genetic material among phylogenetically distinct clades or species of plants. How can genetic material arrive in a new genome and function there if it lacks active promoters and appropriate downstream sequences for RNA 3' processing and stabilization? Does most transferred DNA consist of complete gene cassettes including functional expression elements? Unless a transferred gene has a homolog in the recipient, it should function only if expression can be properly regulated by the recipient or if it is an "independent gene" as appears to be true of a horizontally transferred antifreeze protein in fish (Graham et al. 2008). Gene conversion between foreign and native genes could have deleterious consequences, for example by perturbing the function of the encoded protein (Ragan and Beiko 2009).

Whether inter-specific HGT has an important role in the evolution of plants is still unclear. Plausible examples of positive evolutionary impacts are the inferred HGTs from fungi to the lycophyte *Selaginella moellendorffii* of a putative membrane transporter gene and from fungi to the moss *Physcomitrella patens* of a putative sugar transporter gene (Richards et al. 2009). Otherwise, beneficial impacts of HGT have been demonstrated or proposed mainly for prokaryotes, unicellular eukaryotes, and animals (Graham et al. 2008; Marchetti et al. 2009; Danchina et al. 2010).

More molecular-biological investigations and better experimental systems in the lab are sorely needed to understand the role(s) of HGT in plants. Horizontal gene transfer in Viridiplantae may be especially difficult to detect because most events seem to involve mtDNA, which at the substitution level evolves extremely slowly, creating a challenge for the phylogenetic approach of inferring events from contradictory gene trees. The warning of Keeling and Palmer (2008) that the picture may be getting more complex with increasingly denser sampling of taxa, genes and genomes so far is borne out (for plants at least), and we are still far from a satisfactory understanding of the mechanisms, vectors and evolutionary significance of natural horizontal gene transfer.

References

Adams KL, Clements MJ, Vaughn JC (1998) The *Peperomia* mitochondrial *coxI* group I intron: timing of horizontal transfer and subsequent evolution of the intron. J Mol Evol 46:689–696

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Andersson JO (2005) Lateral gene transfer in eukaryotes. Cell Mol Life Sci 62:1182–1197

Archibald JM, Richards TA (2010) Gene transfer: anything goes in plant mitochondria. BMC Biol 8:147

Arimura S, Yamamoto J, Aida GP, Nakazono M, Tsutsumi N (2004) Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. Proc Natl Acad Sci USA 101:7805–7808

Barkman TJ, McNeal JR, Lim S-H, Coat G, Croom HB, Young ND, dePamphilis CW (2007) Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. BMC Evol Biol 7:248

Belshaw R, Bensasson D (2006) The rise and falls of introns. Heredity 96:208–213

Bergthorsson U, Adams KL, Thomason B, Palmer JD (2003) Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature 424: 197–201

Bergthorsson U, Richardson AO, Young GJ, Goertzen LR, Palmer JD (2004) Massive horizontal transfer of mitochondrial genes from diverse land plant donors to the basal angiosperm *Amborella*. Proc Natl Acad Sci USA 101:17747–17752

Bock R (2010) The give-and-take of DNA: horizontal gene transfer in plants. Trends Plant Sci 15:11–22

Boto L (2010) Horizontal gene transfer in evolution: facts and challenges. Proc R Soc B 277(1683): 819–827

Broothaerts W, Mitchell HJ, Weir B, Kaines S, Smith LMA, Yang W, Mayer JE, Roa-Rodrıguez C, Jefferson RA (2005) Gene transfer to plants by diverse species of bacteria. Nature 433:629–633

Burger G, Gray MW, Lang BF (2003) Mitochondrial genomes: anything goes. Trends Genet 19:709–716

Cho Y, Palmer JD (1999) Multiple acquisitions via horizontal transfer of a group I intron in the mitochondrial *coxl* gene during evolution of the Araceae family. Mol Biol Evol 16:1155–1165

Cho Y, Adams KL, Qiu Y-L, Kuhlman P, Vaughn JC, Palmer JD (1998a) A highly invasive group I intron in the mitochondrial *cox1* gene. In: Moller I-M, Gardestrom P, Glimelius K, Glaser E (eds) Plant mitochondria: from gene to function. Backhuys, Leiden, pp 19–23

Cho Y, Qiu Y-L, Kuhlman P, Palmer JD (1998b) Explosive invasion of plant mitochondria by a group I intron. Proc Natl Acad Sci USA 95:14244–14249

Cohen O, Pupko T (2010) Inference and characterization of horizontally transferred gene families using stochastic mapping. Mol Biol Evol 27:703–713

Cusimano N, Zhang L-B, Renner SS (2008) Reevaluation of the *cox1* group I intron in Araceae and angiosperms indicates a history dominated by loss rather than horizontal transfer. Mol Biol Evol 25:1–12

Danchina EGJ, Rossoa M-N, Vieiraa P, de Almeida-Englera J, Coutinhob PM, Henrissat B, Abad P (2010) Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. Proc Natl Acad Sci USA 107:17651–17656

Davis CC, Wurdack KJ (2004) Host-to-parasite gene transfer in flowering plants: phylogenetic evidence from Malpighiales. Science 305:676–678

Davis CC, Anderson WR, Wurdack KJ (2005) Gene transfer from a parasitic flowering plant to a fern. Proc R Soc B 272:2237–2242

Diao X, Freeling M, Lisch D (2006) Horizontal transfer of a plant transposon. PLoS Biol 4:e5

Dumas F, Duckely M, Pelczar P, Van Gelder P, Hohn B (2001) An *Agrobacterium* VirE2 channel for transferred-DNA transport into plant cells. Proc Natl Acad Sci USA 98:485–490

Embley TM, Thomas RH, Williams RAD (1993) Reduced thermophilic bias in the 16S rDNA sequence from *Thermus ruber* provides further support for a relationship between *Thermus* and *Deinococcus*. Syst Appl Microbiol 16:25–29

Farris JS, Kallersjo M, Kluge AG, Bult C (1994) Testing significance of incongruence. Cladistics 10:315–319

Felsenstein J (1978) Cases in which parsimony or compatability methods will be positively misleading. Syst Zool 27:401–410

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783–791

Ferandon C, Moukha S, Callac P, BenedettoJ-P CM, Barroso G (2010) The *Agaricus bisporus cox1* gene: the longest mitochondrial gene and the largest reservoir of mitochondrial group I introns. PLoS ONE 5:e14048

Filipowicz N, Renner SS (2010) The worldwide holoparasitic *Apodanthaceae* confidently placed in the Cucurbitales by nuclear and mitochondrial gene trees. BMC Evol Biol 10:219

Foster PG, Cox CJ, Embley TM (2009) The primary divisions of life: a phylogenomic approach employing composition-heterogeneous methods. Philos Trans R Soc Lond B 364:2197–2207

Gelvin SB (2009) *Agrobacterium* in the genomics age. Plant Physiol 150:1665–1676

Gilbert C, Schaack S, Pace JK, Brindley PJ, Feschotte C (2010) A role for host-parasite interactions in the horizontal transfer of transposons across phyla. Nature 464:1347–1352

Gogarten JP, Townsend JP (2005) Horizontal gene transfer, genome innovation and evolution. Nat Rev 3:679–687

Goremykin VV, Salamini F, Velasco R, Viola R (2009) Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. Mol Biol Evol 26:99–110

Graham LA, Lougheed SC, Ewart KV, Davies PL (2008) Lateral transfer of a lectin-like antifreeze protein gene in fishes. PLoS One 3:e2616

Grewe F, Viehoever P, Weisshaar B, Knoop V (2009) A trans-splicing group I intron and tRNA-hyperediting in the mitochondrial genome of the lycophyte *Isoetes engelmannii*. Nucleic Acids Res 37: 5093–5104

Hao W (2010) OrgConv: detection of gene conversion using consensus sequences and its application in plant mitochondrial and chloroplast homologs. BMC Bioinformatics 11:114

Hao W, Palmer JD (2009) Fine-scale mergers of chloroplast and mitochondrial genes create functional, transcompartmentally chimeric mitochondrial genes. Proc Natl Acad Sci USA 106:16728–16733

Hao W, Richardson AO, Zheng Y, Palmer JD (2010) Gorgeous mosaic of mitochondrial genes created by horizontal transfer and gene conversion. Proc Natl Acad Sci USA 107:21576–21581

Hecht J, Grewe F, Knoop V (2011) Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of Selaginella moellendorffii mitochondria: the root of frequent plant mtDNA recombination in early tracheophytes. Genome Biol Evol 3:344–358

Henze K, Martin W (2001) How do mitochondrial genes get into the nucleus? Trends Genet 17: 383–387

Hill T, Nordström KJV, Thollesson M, Säfström TM, Vernersson AKE, Fredriksson R, Schiöth HB (2010) SPRIT: identifying horizontal gene transfer in rooted phylogenetic trees. BMC Evol Biol 10:42

Inda LA, Pimentel M, Chase MW (2010) Contribution of mitochondrial *cox1* intron sequences to the phylogenetics of tribe Orchideae (Orchidaceae): do the distribution and sequence of this intron in orchids also tell us something about its evolution? Taxon 59:1053–1064

Jiang N, Bao Z, Zhang X, Eddy SR, Wessler SR (2004) Pack-MULE transposable elements mediate gene evolution in plants. Nature 431:569–573

Keeling PJ (2009a) Functional and ecological impacts of horizontal gene transfer in eukaryotes. Curr Opin Genet Dev 19:613–619

Keeling PJ (2009b) Role of horizontal gene transfer in the evolution of photosynthetic eukaryotes and their plastids. In: Gogarten MB et al (eds) Horizontal gene transfer: genomes in flux, vol 532, Methods in molecular biology., pp 501–515

Keeling PJ, Palmer JD (2008) Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet 9:605–618

Knoop V, Volkmar U, Hecht J, Grewe F (2010) Mitochondrial genome evolution in the plant lineage. In: Kempken F (ed) Plant mitochondria, vol 1, Advances in plant biology., pp 3–29

Koski LB, Golding B (2001) The closest BLAST hit is often not the nearest neighbor. J Mol Evol 52: 540–542

Koulintchenko M, Konstantinov Y, Dietrich A (2003) Plant mitochondria actively import DNA via the permeability transition pore complex. EMBO J 22:1245–1254

Kurland CG, Canback B, Berg OG (2003) Horizontal gene transfer: a critical view. Proc Natl Acad Sci USA 100:9658–9662

Lambowitz AM, Belfort M (1993) Introns as mobile genetic elements. Annu Rev Biochem 62:587–622

Lambowitz AM, Zimmerly S (2004) Mobile group II introns. Annu Rev Genet 38:1–35

Lander ES, The International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. Nature 409: 860–921

Lucas WJ, Ham B-K, Kim J-Y (2009) Plasmodesmata – bridging the gap between neighboring plant cells. Trends Cell Biol 19:495–503

Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim WB, WoloshukC XX, Xu J-R, Antoniw J, Baker SE, Bluhm BH, Breakspear A, Brown DW, Butchko RAE, Chapman S, Coulson R, Coutinho PM, Danchin EGJ, Diener A, Gale LR, Gardiner DM, Goff S, Hammond-Kosack KE, Hilburn K, Van Hua A, Jonkers W, Kazan K, Kodira CD, Koehrsen M, Kumar L, Lee Y-H, Li L, Manners JM, Miranda-Saavedra D, Mukherjee M, Park G, Park J, Park SY, Proctor RH, Regev A, Ruiz-Roldan MC, Sain D, Sakthikumar S, Sykes S, Schwartz DC, Turgeon BG, Wapinski I, Yoder O, Young S, Zeng Q, Zhou S, Galagan J, Cuomo CA, Kistler HC, Rep M (2010) Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367–373

Marchetti A, Parker MS, Moccia LP, Lin EO, Arrieta AL, Ribalet F, Murphy MEP, Maldonado MT, Armbrust EB (2009) Ferritin is used for iron storage in bloom-forming marine pennate diatoms. Nature 457(22):467–470

Martin W (2005) Lateral gene transfer and other possibilities. Heredity 94:565–566

Mirkin BG, Fenner TI, Galperin MY, Koonin EV (2003) Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes. BMC Evol Biol 3:2

Morgante M, Brunner S, Pea G, Fengler K, Zuccolo A, Rafalski A (2005) Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. Nat Genet 37:997–1002

Mower JP, Stefanovic S, Young GJ, Palmer JD (2004) Gene transfer from parasitic to host plants. Nature 432:165–166

Mower JP, Stefanovic S, Hao W, Gummow JS, Jain K, Ahmed D, Palmer JD (2010) Horizontal acquisition of multiple mitochondrial genes from a parasitic plant followed by gene conversion with host mitochondrial genes. BMC Biol 8:150

Nickrent DL, Blarer A, Qiu YL, Vidal-Russell R, Anderson FE (2004) Phylogenetic inference in Rafflesiales: the influence of rate heterogeneity and horizontal gene transfer. BMC Evol Biol 4:40

Nishida H, Sugiyama J (1995) A common group I intron between a plant parasitic fungus and its host. Mol Biol Evol 12:883–886

Noble GP, Rogers MB, Keeling PJ (2007) Complex distribution of EFL and EF-1alpha proteins in the green algal lineage. BMC Evol Biol 7:82

Odom OW, Shenkenberg DL, Garcia JA, Herrin DL (2004) A horizontally acquired group II intron in the chloroplast *psbA* gene of a psychrophilic *Chlamydomonas*: in vitro self-splicing and genetic evidence for maturase activity. RNA 10:1097–1107

Omer S, Kovacs A, Mazor Y, Gophna U (2010) Integration of a foreign gene into a native complex does not impair fitness in an experimental model of lateral gene transfer. Mol Biol Evol 27:2441–2445

Ragan MA, Beiko RG (2009) Lateral genetic transfer: open issues. Philos Trans R Soc Lond B 364: 2241–2251

Ragan MA, Harlow TJ, Beiko RG (2006) Do different surrogate methods detect lateral genetic transfer events of different relative ages? Trends Microbiol 14:4–8

Richards TA, Soanes DM, Foster PG, Leonard G, Thornton CR, Talbot NJ (2009) Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. Plant Cell 21:1897–1911

Richardson AO, Palmer JD (2007) Horizontal gene transfer in plants. J Exp Bot 58:1–9

Rogers MB, Watkins RF, Harper JT, Durnford DG, Gray MW, Keeling PJ (2007) A complex and punctate distribution of three eukaryotic genes derived by lateral gene transfer. BMC Evol Biol 7:89

Roney JK, Khatibi PA, Westwood JH (2007) Cross-species translocation of mRNA from host plants into the parasitic plant dodder. Plant Physiol 143:1037–1043

Salzberg SL, White O, Peterson J, Eisen JA (2001) Microbial genes in the human genome: lateral transfer or gene loss? Science 292:1903–1906

Sanchez-Gracia A, Maside X, Charlesworth B (2005) High rate of horizontal transfer of transposable elements in *Drosophila*. Trends Genet 21:200–203

Sanchez-Puerta MV, Cho Y, Mower JP, Alverson AJ, Palmer JD (2008) Frequent, phylogenetically local horizontal transfer of the *cox1* group I intron in flowering plant mitochondria. Mol Biol Evol 25:1762–1777

Sanchez-Puerta MV, Abbona CC, Zhuo S, Tepe EJ, Bohs L, Olmstead RG, Palmer JD (2011) Multiple recent horizontal transfers of the cox1 intron in Solanaceae and extended coconversion of flanking exons. BMC Evol Biol 11:277

Schaack S, Gilbert C, Feschotte C (2010) Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. Trends Ecol Evol 25:537–546

Schönenberger J, Anderberg AA, Sytsma KJ (2005) Molecular phylogenetics and patterns of floral evolution in the ericales. Int J Plant Sci 166:265–288

Sheahan MB, McCurdy DW, Rose RJ (2005) Mitochondria as a connected population: ensuring continuity of the mitochondrial genome during plant cell dedifferentiation through massive mitochondrial fusion. Plant J 44:744–755

Snir S, Trifonov E (2010) A novel technique for detecting putative horizontal gene transfer in the sequence space. J Comput Biol 17:1535–1548

Sorek R, Zhu Y, Creevey CJ, Francino MP, Bork P, Rubin EP (2007) Genome-wide experimental determination of barriers to horizontal gene transfer. Science 318:1449–1452

Stanhope MJ, Lupas A, Italia MJ, Koretke KK, Volker C, Brown JR (2001) Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates. Nature 411:940–944

Stegemann S, Bock R (2009) Exchange of genetic material between cells in plant tissue grafts. Science 324:649–651

Stiller JW (2011) Experimental design and statistical rigor in phylogenomics of horizontal and endosymbiotic gene transfer. BMC Evol Biol 11:259

Ülker B, Li Y, Rosso MG, Logemann E, Somssich IE, Weisshaar B (2008) T-DNA–mediated transfer of *Agrobacterium tumefaciens* chromosomal DNA into plants. Nat Biotechnol 26:1015–1017

Vaughn JC, Mason MT, Sper-Whitis GL, Kuhlman P, Palmer JD (1995) Fungal origin by horizontal transfer of a plant mitochondrial group I intron in the chimeric *coxI* gene of *Peperomia*. J Mol Evol 41: 563–572

Wolf YI, Kondrashov FA, Koonin EV (2001) Footprints of primordial introns on the eukaryotic genome: still no clear traces. Trends Genet 17:499–501

Won H, Renner SS (2003) Horizontal gene transfer from flowering plants to *Gnetum*. Proc Natl Acad Sci USA 100:10824–10829

Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol 39: 306–314

Yoshida S, Maruyama S, Nozaki H, Shirasu K (2010) Horizontal gene transfer by the parasitic plant *Striga hermonthica*. Science 328:1128

Chapter 7

General Discussion

Taxonomy of Apodanthaceae

When I began this research in late 2010, there were 36 species names available in Apodanthaceae, 18 of them simple nomenclatural synonyms because of transfers of names between *Apodanthes* and *Pilostyles* (there was almost no name that had not been placed in both genera by different authors). My work based on loans of 407 collections, with about 159 duplicates, from 11 major herbaria lead to the synonymizing of 7 species names and the acceptance of 10 names as probably referring to biological species. I am following a morphological-geographic species concept, supported by genetic data (but with only sparse within-species genetic sampling), in which specimens with similar flowers and a coherent geographic range are considered a biological species. Based on data from the collection labels, flower measurements, and photos of living flowers, I produced a key to the species (Bellot and Renner, 2014a: p. 49), a distribution map (Fig. 2 in Chapter 2), and the first host inventory for the family (Fig. 3 and Table I in Chapter 2). Dissections of 123 flowers from the American and African species (Suppl. material 2 in Chapter 2) revealed much variation in tepal number in some species, for instance in *P. berteroi*, a species otherwise well differentiated by its host and geographic ranges (parasitizing only *Adesmia* in Chile). No consistent differences could be observed between flowers of type collections or other historic collections of *P. ulei*, *P. ingae*, *P. goyazensis*, and *P. blanchetii*, and I therefore synonymized these names.

Using Apodanthaceae-specific primers I could sequence mitochondrial *mat*R and nuclear 18S for at least one representative of, respectively, 13 and 15 of the 18 previously accepted species, including the types of *P. coccoidea* and *A. minarum*. Where I had multiple sequences from specimens with similar flowers, the DNA phylogenies matched the morphological and geographic species boundaries I had established. Topologies from the mitochondrial and nuclear marker were mostly congruent, but both left the relationships between Central and South American species unresolved. The DNA phylogeny also allowed me to assign a species name to five specimens previously unidentified because of ambiguous flower morphology or very few poorly preserved flowers. Sequences from the type of *Apodanthes minarum*, confirmed the synonymy of this name under *A. caseariae*.

As visible on the specimen-based map (Fig. 2 in Chapter 2), East Africa is under-collected for Apodanthaceae. I have seen material of only one African species, *Pilostyles aethiopica*, represented by 41 collections mostly from Zimbabwe, but in July 1907, W. Holz may have collected another species in Tanzania, described as *P. holzii* by Engler. The type of

this species was destroyed during WWII, and I synonymized the name under *P. aethiopica* based solely on its protologue (Engler, 1912). Another problematic case concerns a collection investigated in Chapter 2 (*R. Callejas et al. 8062*), which appears to be the only collection of Apodanthaceae from Colombia so far available in the herbaria of Europe and USA. Recent studies by González and Pabón-Mora (2014a, b) report new collections of Apodanthaceae in Colombia that represent the first record of the genus *Pilostyles* in that country. The authors describe this material as the new species *P. boyacensis* (González and Pabón-Mora, 2014b), but based on its morphology, this name is a synonym of *P. berteroi* (see note in Bellot and Renner, 2014a: p. 55).

After the publication of my taxonomic revision, I had the opportunity to dissect and sequence a specimen of *Pilostyles stawiarskii*, a species previously known only from its type specimen collected in 1948 in Paraná, Brazil (Vattimo, 1950). The curator of the largest herbarium of Brazil (acronym R), R. Alves, recollected the species at the type location during the summer of 2014. Adding a *matR* fragment of this specimen to my family-wide *matR* matrix revealed that *P. stawiarskii* groups with a clade formed by representatives of the widespread *P. blanchetii* (from Argentina to Costa-Rica, Fig. 2 of Chapter 2). Based on the description of the flowers of *P. stawiarskii* (Vattimo, 1950) and the *matR* sequences, the name *P. stawiarskii* is a synonym of the older name *P. blanchetii* as I had suggested based on the name's protologue (Bellot and Renner, 2014a: p. 54). *Pilostyles blanchetii* is widespread in South America, and I was able to study 43 flowers from 24 collections (Suppl. material 2 in Chapter 2; Fig. 2A and 2B in Chapter 4).

The geographic ranges of eight of the ten species that I recognize in the family do not overlap: the exceptions are *Apodanthes caseariae* and *P. blanchetii* (Fig. 2 in Chapter 2), which have very different morphologies (ca. 1cm-long orange/yellow vs. smaller purple/brown flowers; Figures 5, 6D and 6E in Chapter 2) and host ranges (Salicaceae vs. Fabaceae; Fig. 3 in Chapter 2) and thus probably do not exchange genes even when sympatric.

Biogeography of Apodanthaceae

The molecular phylogenies based on 18S and *mat*R (Fig. 2A and 2B in Chapter 4) support the mutual monophyly of *Apodanthes* and *Pilostyles*, which is consistent with their different host families, Salicaceae and Fabaceae (Fig. 3 in Chapter 2). The 18S tree further shows that the African *P. aethiopica* is embedded in the *Pilostyles* clade, a lucky

circumstance since the name of the supposed African genus "*Berlinianche*" was never validly published (as explained in Bellot and Renner, 2013 and 2014a). The 18S tree also shows a closer relationship between the African/Iranian and the Australian species than any of these have with the American ones. In the Americas, the relationships between species remain unclear and differ between the *mat*R and 18S trees as regards the relative positions of the Chilean *P. berteroi* and the Central American *P. mexicana*, which is reflected by a yet different placement of these species in the combined tree (Fig. 2C in Chapter 4). The species from USA and Mexico, *P. thurberi*, is consistently in basal position compared to the other, more southern, Neotropical species (Fig. 2A, 2B and 2C in Chapter 4).

The stem age of Apodanthaceae was inferred as 65 to 81 million years (Ma; Table 2 of Chapter 4), younger than the breakup of Africa and South America, which occurred about 100 Ma ago (Jokat et al., 2003), hence the spread of Apodanthaceae between these continents must have involved transoceanic dispersal (although at 80 Ma, America and Africa were still relatively close). Since the sister group of Apodanthaceae within Cucurbitales is unknown (Filipowicz and Renner, 2010; Schaefer and Renner, 2011), one cannot infer the early biogeographic history of the family. The seeds of Apodanthaceae are tiny, and embedded in small fleshy fruits (Bouman and Meijer, 1994; Bellot, pers. obs. of fruits from herbarium specimens), and modern Apodanthaceae are dispersed by frugivorous birds or mammals. Because of the tiny seeds and their immediate need of a host to survive and germinate, Apodanthaceae must have achieved any long-distance dispersal inside a host, probably as parts of floating islands (Van Duzer, 2004, 2006; Vidal et al., 2008). The presence of *P. blanchetii* in Jamaica and the Cayman islands is evidence for the ability of Apodanthaceae to disperse overseas.

The diversification of Apodanthaceae in Central and South America involved colonization of new habitats (semi-arid vs. tropical) and hosts (Fabaceae vs Salicaceae or *Bauhinia* ssp. vs *Dalea* ssp.), which could have been facilitated by the closing of the Panamanian Isthmus starting between 20 and 3.5 Ma ago (Montes et al., 2012; Coates et al., 2004; Gutiérrez-García and Vázquez-Domínguez, 2013), roughly coinciding with my estimates under the RLC model of 6 Ma (Table 2 in Chapter 4) for the divergence between the Mexican and South American *Pilostyles* species.

Reproductive system of Apodanthaceae

My flower dissections revealed only unisexual (male or female) flowers (Suppl.

material 2 in Chapter 2), but there is evidence for anatomical (but not functional) bisexuality in *P.* (*ingae*) *blanchetii* (Brasil, 2010; photos of stained sections from Brasil's research are available online at http://www.ispa.cnr.it/ftp/vurro/Presentations/Ceccantini.pdf). My systematic screening of 20 infected hosts in a population of *P. aethiopica* in Zimbabwe and 26 and 27 infected hosts in, respectively, two populations of *P. haussknechtii* in Iran revealed dioecy: in most of the cases, parasite flowers of only one sex were present on one host individual (Table 2 in Chapter 3). However, in each species, I found one case where a host bore parasite flowers of both sexes, on the same branch (Fig. 2C in Chapter 3). Such variation has also been reported for *P. thurberi* and *P. blanchetii* (Brasil, http://www.ispa.cnr.it/ftp/vurro/Presentations/Ceccantini.pdf). It is therefore not yet clear whether Apodanthaceae are strictly dioecious (Poiteau, 1824; Vattimo 1950, 1971, 1973) or sometimes monoecious. Thiele et al. (2008) describe the two Australian species *P. collina* and *P. coccoidea* as monoecious, but it is unclear if this implies the colonization of the same host by two individuals of different sexes or a unique colonization by a hermaphrodite individual.

The African *Pilostyles aethiopica* is pollinated by flies from the family Calliphoridae, as inferred from observation of visitors of male and female flowers (Fig. 3A and 3B in Chapter 3), as well as from their shallow nectar cushion, purple color and sweet but unpleasant scent (Fig. 2D and 2E in Chapter 3). On the same species and near the same location where I carried out my work, Blarer et al. (2004) observed flies, albeit of other families (Drosophilidae, Sciaridae, Psychodidae, and Cecidomyiidae), but without information on behavior or photos showing the insects on the flowers. The Iranian *P. haussknechtii* has the same morphological pollination syndrome (Fig. 2C in Chapter 3), and I observed flies of the family Ulidiidae visiting the flowers (Fig. 3G in Chapter 3). Wasps have been reported as visiting the flowers of two other species (Dell and Burbidge, 1981: Australia; Brasil, 2010: Brazil), and the wasp *Pepsis menechma* Lepeletier 1845 was observed visiting *A. casearіae* in Panama (C. Galdames, Smithsonian Tropical Research Institute, Panama, pers. com. with pictures, July 2012).

Biological traits of parasitic angiosperms

I generated the first global species-level database of sexual systems, pollinators, dispersal agents, growth form and longevity of parasitic angiosperms (Appendix S1 of Chapter 3), part of it now integrated in the "Tree of Sex" database (Tree of sex consortium,

2014). The main result is that 99.5% of parasitic angiosperms species are animal-pollinated against ca. 80% of non-parasitic flowering plants. The only parasite that may be wind pollinated is the dwarf mistletoe *Arceuthobium* (Viscaceae), but even that is doubtful (Player, 1979). I also found a weak tendency of dioecy being over-represented (10% vs. 6% in non-parasitic angiosperms), similar to Conn et al. (1980) for the flora of Carolina and Renner and Ricklefs (1995) for angiosperm genera. Unfortunately, the lack of species-level phylogenies precludes estimating the number of independent acquisitions of dioecy.

The sister groups of parasitic angiosperms always contain animal-pollinated species (Table 1 in Chapter 3), reflecting the high proportion of animal-pollination in angiosperms. Parasites may thus simply have inherited animal pollination. However, the tendency of parasites to grow close to their hosts, hidden in their foliage or in the soil litter may require active, and thus biotic, pollination, which would fit with an absence of wind-pollination in parasites. Given the low proportion of dioecious species among non-parasitic angiosperms, the (slight) over-representation of dioecy among parasites is interesting. However, as explained above, the sexual systems of parasites (whether monoecious or dioecious) are not straightforward to assess, and the data for most of them are poor.

From my review (Chapter 3) it appears that most parasitic angiosperms are perennials and that no annual parasitic species are dioecious, which fits with the findings of Renner and Ricklefs (1995) that dioecy is more likely in shrubs and trees (perennials) than in herbs (shorter longevity). The higher proportion of dioecy in parasitic angiosperms than non-parasitic angiosperms could thus reflect the high proportion of perennials among parasites (Table 1 in Chapter 3), with longevity reducing the problem of having to colonize a new host individual every year. Another hypothesis is that dioecy could have evolved with host-specialized parasitism because both traits require an efficient seed dispersal mechanism (Wilson and Harder, 2003).

In conclusion, the finding of Molau (1995) that no biological traits correlate with parasitism in angiosperms has to be modified because my data indicate a clear preponderance of animal-pollination in parasites and a weak overrepresentation of dioecy. More species-level data are necessary to test hypotheses on evolutionary interactions between biological traits and (certain kinds of) parasitism, such as the links between host specialization, dioecy, longevity, animal dispersal, and animal pollination.

Substitution rates in the genomes of parasitic angiosperms, and their handling in molecular clock models

Uneven branch lengths in molecular phylogenies can be symptomatic of incomplete sampling or species extinctions (Crisp and Cook, 2009) or they can indicate different rates of nucleotide substitution. In my work, the *mat*R and 18S phylogenetic trees for Cucurbitales and Fabales (Fig. 2A and 2B in Chapter 4) show a long branch leading to the Apodanthaceae and less branch length heterogeneity within Apodanthaceae. These relatively species-poor trees thus offered a worthwhile opportunity to compare the behaviours of the uncorrelated lognormal (Drummond et al., 2006) and the random local clock (Drummond and Suchard, 2010) models implemented in BEAST since v. 1.5.4.

The results of these experiments reflect the different ways in which each clock model deals with rate heterogeneity. In terms of relative rates (in substitutions per site), the uncorrelated lognormal (UCLN) clock model inferred more, but less intense, rate jumps than the random local clock model (RLC; Figure S1 and Table S3 of Chapter 4), as expected from their theoretical basis (see *General Introduction* and Drummond et al., 2006; Drummond and Suchard, 2010; Ho and Duchène, 2014). The different prior distributions on the number of permitted rate changes in the RLC model revealed their great impact on the posteriors. As visible in Fig. 3B, 3C, 3E and 3F of Chapter 4, when the prior distribution assigned a 50% probability to the occurrence of a rate jump (default Poisson distribution), the inferred changes were fewer and stronger than when this probability was increased to follow a uniform distribution on a large interval (0 to 20 rate changes). In the latter case, it was possible to recover a rate distribution similar to the one obtained with the UCLN model.

An advantage of the RLC models (with Poisson or uniform priors) is that they can handle correlated rates and thus be used as a test for a global clock. This offers more flexibility than the UCLN model, which cannot fit cases of identical rates in adjacent branches. Setting a uniform prior distribution for the number of rate changes in the RLC model may thus have the advantages, but not the drawbacks, of the UCLN model. However, this needs to be tested using simulated data (Drummond and Suchard, 2010).

I calibrated my clock models with a cupule fossil reliably attributed to Fagales (Herendeen et al., 1995) and representing the minimum age of the divergence between Fagales and Cucurbitales; the fossil was assigned a range of 85-105 Ma. With this calibration, the UCLN and RLC models yielded very different node ages within Apodanthaceae (Table 2 in Chapter 4), underlining the importance of priors and model

comparison. The UCLN gave the older ages, whereas the RLC with the default Poisson prior on permitted rate jumps resulted in the youngest estimates (Fig. 3A, 3B, 3D and 3E in Chapter 4). The RLC with a prior that allowed less strong, but more frequent rate changes resulted in intermediate ages (Fig. 3C and 3F in Chapter 4).

My study benefited from the recent implementation in BEAST 1.7 (Drummond et al., 2012) of new estimators of the marginal likelihood, namely the path sampling and stepping-stone sampling methods, which both gave the same results and are more reliable than previous methods (Baele et al., 2012). For both 18S and *mat*R, Bayes factors strongly preferred the UCLN model over the RLC model with a Poisson prior, with strong preference defined as in the Bayes factor scale of Kass and Raftery (1995). Bayes factors favoured the UCLN over the RLC model with a uniform prior for 18S, whereas this preference was not significant for *mat*R (Tables 3 and 4 in Chapter 4). It appears from this that rate heterogeneity among lineages is distributed differently in *mat*R and 18S, matching findings that some genes of holoparasites show a higher rate of substitution than their autotrophic relatives, whereas others do not (Su and Hu, 2012; also the *Introduction* of this thesis).

Besides the fossil calibration, I also experimented using host divergence times as calibration points for host-specific parasites. Apodanthaceae species are often found on more than one host species (see Table I in Chapter 2), and I therefore used the minimal age of an entire host genus to constrain the age of a parasite species specialized on these hosts (Fig. 1A in Chapter 4). This may have contributed to huge confidence intervals (Table 2 in Chapter 4). Despite those difficulties, this method could be a valuable cross-validation approach in cases of more specific parasites than Apodanthaceae, such as the Rafflesiaceae, where the entire family grows only on the genus *Tetrastigma* (Vitaceae).

C-values, chromosome numbers, and plastid genomes of Apodanthaceae

I observed flower sections of *Pilostyles aethiopica* and *P. haussknechtii* under a light microscope but could not detect any chloroplasts. This is in accordance with the work of Rutherford (1970) and Dell et al. (1982) who also only observed undifferentiated plastids in the cells of *P. thurberi* and *P. hamiltonii*.

To infer the coverage of the genome shotgun sequencing, our lab obtained 1C-value measurements for *P. aethiopica*, *P. haussknechtii*, and *P. hamiltonii* from freshly collected flowers. The results were 1.53 pg for a female flower of *P. aethiopica*, 1.50 pg for a male flower of that species; 1.09 pg for a female flower of *P. haussknechtii* and 1.26 pg for a male

flower of that species; and 5.08 pg for young fruits of the Australian *P. hamiltonii* (E. Temsch, University of Vienna, personal communications in March, May, and June 2012). The latter result is surprising but may have to be taken with caution because the flowers were already young fruits, and there may be endopolyploid tissues in this species.

Using flower buds of *P. aethiopica* and *P. haussknechtii* fixed in 3:1 ethanol/glacial acetic acid, I squashed their young anthers with a drop of 45% acetic acid on microscopic slides until I found pollen mother cells. The best slides were stained with DAPI (2 µg/ml) and analysed by fluorescent light microscopy. For *P. haussknechtii*, I could use meiotic cells with chromosomes in prophase I: I found haploid numbers of $n = 32$ (2 cells), $n = 33$ (one cell, see Fig. 1c in the *Appendix*), $n = 34$ (2 cells), $n = 36$ (2 cells), and $n = 37$ (one cell). I also estimated diploid chromosome numbers from mitotic cells and found three cells with $2n = 66$ (Figures 1a and 1b in the *Appendix*), one cell with $2n = 65$, one cell with $2n = 65/66$, one cell with $2n = 64/66$ and one cell with $2n = 67$ chromosomes. For *P. aethiopica*, the counts were consistent between 4 cells with chromosomes in late anaphase II (end of meiosis) and revealed a haploid chromosome number of $n = 29$ (Figure 2a in the *Appendix*), whereas one cell showed an ambiguous number of $n = 29/30$, and 2 cells showed numbers of $n = 26/27$ and $n = 27/28$. More counts are needed to confirm haploid numbers of $n = 30$ for *P. aethiopica* and $n = 33$ for *P. haussknechtii*. Rutherford (1970) reported a haploid chromosome number of $n = 30$ for the Mexican *P. thurberi*.

Based on the results presented in Chapter 5, *P. aethiopica* and *P. hamiltonii* completely lack a plastid genome. By combining Illumina and 454 data, I expected to obtain high coverage for the plastome as well as enough information to assemble regions that would not be collinear to the reference plastome (*Cucumis sativus*). However, all I could detect were a few plastid-like contigs, most of which had mitochondrial-like flanking regions (Fig. 1 in Chapter 5). It was not possible to assemble those contigs in longer fragments, and no functional plastid gene was found. I conclude that *Pilostyles aethiopica* and *P. hamiltonii* probably lost their plastome. This makes Apodanthaceae the second angiosperm lineage in which a plastome could not be found using high throughput sequencing of genomic DNA, the first one being another endoparasite, *Rafflesia lagascae* (Molina et al., 2014).

The plastid-like sequences detected in *P. aethiopica* and *P. hamiltonii* are different from each other, which points to independent plastome decay in the two species. I found the *trn*E and a pseudogenized *clp*P only in *P. hamiltonii* and non-functional fragments of *acc*D, *ycf*1 and *ycf*2 in *P. hamiltonii* and *P. aethiopica*. Most fragments of these genes were flanked by mitochondrial regions (Fig. 1 and Table 1 in Chapter 5). *Rafflesia lagascae* genomic DNA

contains neither the *trn*E gene, nor functional copies of *acc*D, *clp*P, *ycf*1, and *ycf*2, and in *R. leonardi*, *trn*E seems to have been relocated to the nuclear genome (Molina et al., 2014), which goes against the hypothesis that *trn*E is essential for plant plastomes (Barbrook et al., 2006)

In exo-holoparasites (parasites that penetrate the host from the outside via haustoria), the loss or pseudogenization following acquisition of parasitism may depend on whether a gene encodes proteins involved directly in photosynthesis or whether it is involved in the housekeeping of those proteins. Wicke et al. (2013) found a highly idiosyncratic gene loss and pseudogenization in nine parasite species of Orobanchaceae, while Barrett et al. (2014) report a directional pattern of loss in eight species (including three subspecies for a total of 10 genomes) of the mycoheterotroph orchid genus *Corallorhiza* in which certain gene-class categories, such as the NAD(P)H dehydrogenase genes, are lost earlier than other categories, such as the photosystem genes. Regardless of whether there is pattern to their plastome gene loss or not, all 19 exo-holoparasitic angiosperms sequenced by now retain 10 ribosomal proteins, four ribosomal RNAs, seven transfer RNAs and the gene of unknown function ycf2, genes that are thus candidates for a minimal plastome "indispensable for plastid function at least in the angiosperms included in this comparison" (Barrett et al., 2014, p. 11). By contrast, all three endoparasites investigated, namely *P. aethiopica*, *P. hamiltonii* (Chapter 5 of this dissertation), and *Rafflesia lagascae* (Molina et al., 2014), did not retain this minimal set of plastome genes (Fig. 2 of Chapter 5). If any of the four lineages of endoparasites (Apodanthaceae, Cytinaceae, Mitrastemonaceae, Rafflesiaceae) evolved gradually from exo-parasites, this is no longer traceable today (none of them is phylogenetically close to any exoparasite lineages). If their molecular clock-inferred stem ages are taken as a maximal age of these lineages' endoparasitic way of life, then this type of life history appears to be older than are the exo-parasitic groups so far investigated in terms of their plastomes (Table 3 of Chapter 5). It would be interesting to know if older exo-holoparasites, such as Hydnoraceae or Balanophoraceae (Naumann et al., 2013), still retain a plastome. This would support endoparasitism as a precondition for complete plastome lost.

Internal and horizontal transfers in parasitic angiosperms

Pilostyles aethiopica and *P. hamiltonii* may not have plastid genomes but they do still contain plastid-like sequences as it is the case for *Rafflesia* (Molina et al., 2014). Similar to Molina et al. (2014), I based my search for the plastome on BLAST alignments to recover

plastid sequences of Cucurbitales and other rosids. A large number of plastomes are now available (530 plastid genomes of land plants were available in Genbank on the 14 November 2014), and any sequence similar to plastid DNA, including those located in the nuclear or the mitochondrial genomes, will be picked up in a BLAST search. This allows recovering any plastid-like sequence from genomic DNA but requires complementary analyses to know their genomic location. Molina et al. (2014) based their inference that most of the plastid sequences found in *Rafflesia lagascae* are in the nuclear genome only on the low coverage of those regions and do not provide information about possible flanking regions that could be identified as nuclear. I instead evaluated all flanking regions of the plastid-like regions found in *P. aethiopica* and *P hamiltonii* and found them to be mostly mitochondrial (Table S2 in Chapter 5). This matches these sequences' coverage of one order of magnitude higher than the coverage of regions that are clearly nuclear (Zoschke et al., 2007).

Mitochondrial genomes readily integrate DNA from other genomic compartments (Stern and Lonsdale, 1982; Kubo and Newton, 2008; Alverson et al., 2010, 2011). This is thought to be due to the high recombination rate in mitochondrial genomes, driven by small repeats scattered through the genome (Marechal and Brisson, 2010; Woloszynska, 2010; Mileshina et al., 2011) combined with their ability to fuse (Sheahan et al., 2005). In addition, mitochondria have an active mechanism of DNA uptake that is absent from plastids (Koulintchenko et al., 2003). This probably explains the lower proportions of internal gene transfer (IGT) towards plastid genomes (Smith, 2011), with only two reported cases in land plants (Iorizzo et al., 2012a, 2012b; Straub et al., 2013). The first case was inferred by complete sequencing and comparison of the plastid and mitochondrial genomes of *Daucus carota* and involved the transfer of a mitochondrial DNA fragment into the plastome of the most recent common ancestor of *Daucus* and *Cuminum* (Apiaceae), possibly mediated by a retrotransposon (Iorizzo et al., 2012a, 2012b). The second case was inferred by comparing the complete organelle genomes of the Apocynaceae species *Asclepias syriaca* and involved a sequence that was originally inside the plastome, then transferred towards the mitochondrial genome, and subsequently transferred back to the plastome, probably via homologous recombination (Straub et al., 2013). The most parsimonious explanation of plastid-like sequences with mitochondrial flanking regions in Apodanthaceae is therefore the IGT of plastid sequences into the mitochondrial genome. Large-scale transfer of plastid DNA into the mitochondrial genome may be a general characteristic of Cucurbitales, as illustrated by the high proportion of plastid DNA in the mitochondrial genomes of *Citrullus lanatus*, *Cucurbita pepo* and *Cucumis sativus* (Alverson et al., 2010; 2011). The mitochondrial

genome of *C. pepo* contains >11% of plastid DNA, 1.7 to 19 times higher than what is observed among other angiosperms (Alverson et al., 2010).

In addition to integrate DNA from other genomic compartments, the mitochondrial genome can integrate DNA from other species as reviewed in Chapter 6. Horizontal gene transfer (HGT) events are difficult to distinguish from DNA contamination or from methodological artefacts because they often involve short sequences. The increased availability of completely sequenced nuclear and mitochondrial genomes from various angiosperm lineages should help to classify DNA sequences with respect to their phylogenetic affinity. However, the rate of false positives and false negatives may be high in cases where the donor and the receiver have similar genomes. To address this problem, I developed empirical thresholds applicable to Apodanthaceae do detect false positive and negative HGT events. Such approach can be used to develop thresholds for other clades. The BLAST results fit with the phylogenetic results (Table S3 in Chapter 5).
20-27% of *P. aethiopica* and 19% of *P. hamiltonii* plastid-like sequences are more similar to Fabales than to other rosids. This fits with the finding of host DNA in the genome of *Rafflesia cantleyi* and *R. lagascae* (Xi et al., 2012, 2013; Molina et al., 2014). Experimental work on tracking the transfer of macromolecules between hosts and parasites using the exo-parasite *Cuscuta* grown on *Arabidopsis thaliana* or *Solanum lycopersium* has shown that large amounts of macromolecules, such as mRNA, pass through the plasmodesmata that cross chimeric cell walls at the interface of parasite and host (Kim et al., 2014). Because in endoparasites the interface between host and parasite is very large one can expect such chimeric walls to allow the constant traffic of molecules in both directions.

GENERAL CONCLUSIONS

My research on Apodanthaceae sheds light on the evolutionary consequences of one of the most extreme ways of life encountered in land plants, endoparasitism. I resolved the taxonomy and phylogeny of the family and gathered fundamental ecological information on two species, *Pilostyles aethiopica* and *P. haussknechtii.* By placing Apodanthaceae in the general context of the evolution of parasitic angiosperms, I also addressed if there are biological traits that favour, or are favoured by, parasitism. Molecular dating experiments provided information about the way to handle rate heterogeneity in phylogenetic trees, which is often a critical point in obtaining absolute times, for example in studies of the biogeography of non-photosynthetic plants. Finally, the analyses of the organelle genomes of two species (*P. aethiopica* and *P. hamiltonii*) provide a new perspective on the factors shaping plastid genome evolution in parasitic and mycotrophic angiosperms. Specifically, I am putting the hypothesis that endoparasitism is the precondition for the complete loss of a plastome.

Apodanthaceae are a family of 10 species distributed on four continents. Most of species occur in Central and South America, but one occurs in Africa, one in Iran and three occur in Australia. More species may await discovery in under-collected areas, such as Colombia and East-Africa. On the other hand, my assessment of the types and other historic collections representing all published species names led to the synonymization of seven of them (a likely new synonym was created in 2014). The morphological and molecular species boundaries accepted here agree well with geographical and host ranges, suggesting that these factors played a large role in the speciation of Apodanthaceae. The molecular phylogeny of the family shows that the genera *Apodanthes* and *Pilostyles* are mutually monophyletic, and that the African/Iranian and Australian *Pilostyles* clades together form the sister clade of the American *Pilostyles*. The stem lineage of Apodanthaceae is too young for continental drift to explain the group's geographic range. Instead, ancestors must have dispersed overseas, probably inside their hosts as components of floating timber "islands".

Flowers of *P. aethiopica* and *P. haussknechtii* are visited and probably pollinated by flies from the families Calliforidae and Ulidiidae. All Apodanthaceae have unisexual flowers and many appear to be dioecious, although I found a few hosts carrying flowers of both sexes. Dioecy is typical of parasites, as revealed by my review of pollinators and sexual systems in flowering plants. I also found that more parasites are animal-pollinated compared to non-parasitic flowering plants. Biotic pollination may be advantageous for parasites that

are hidden in the foliage of their hosts or under litter. The tendency of parasites to be dioecious may relate to the predominance of a perennial life style among them, which may reduce risks associated with host specialization and having to reach the host every year.

Non-photosynthetic plants (holoparasites and holomycotrophs) are characterized by high nucleotide substitution rates compared to their photosynthetic sister groups, although the data are still limited and mostly come from long stem lineage branches. The uncorrelated lognormal relaxed clock model implemented in the software BEAST (Drummond et al., 2006) is useful when rate changes in a data set are not too strong. But in cases of strong rate changes, perhaps even repeated rate jumps, random local clock models may fit the data better, depending on the prior distribution set for the allowed number of rate jumps. Where these models give highly different ages, as was the case for Apodanthaceae, marginal likelihood estimators may help choose one model over the other, although for Apodanthaceae, Bayes Factor estimators remained indecisive.

With my finding that Apodanthaceae have lost their plastid genome, there are now two endoparasite families for which this is the case, the other being Rafflesiaceae, the study species *Rafflesia lagascae*. The remaining endoparasite families, Cytinaceae and Mitrastemonaceae, have not yet been studied. All studied exo-parasites and mycotrophs retain a minimal plastome consisting of *ycf*2, seven transfer RNAs, 10 ribosomal proteins and the four ribosomal RNAs. Since the plastomes of the exo-holoparasites investigated so far belong to clades that are younger than are Rafflesiaceae and Apodanthaceae (judging from molecular clocks), the investigation of the plastid genomes of older exo-holoparasites would help test whether loss of a plastome is just a matter of time or if endoparasitism is a precondition for plastome loss.

Another finding of my research is that the remnants of the plastid genes of *P. aethiopica* and *P. hamiltonii* are located in their mitochondrial genomes, clearly transferred there by inter-cellular transfer, a common phenomenon in plants. In addition to their own plastid fragments, both species have acquired plastid fragments from their Fabales hosts, in the case of *P. aethiopica,* West African legumes, in the case of *P. hamiltonii,* Australian legumes. None of the fragments is functional but a high rate of horizontal gene transfer matches what has been observed in other angiosperms, including the parasite *Rafflesia cantleyi*.

REFERENCES

Alverson, AJ, X Wei, DW Rice, DB Stern, K Barry, JD Palmer. 2010. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Mol Biol Evol.* 27: 1436–1448.

Alverson, AJ, DW Rice, S Dickinson, K Barry, JD Palmer. 2011. Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell* 23: 2499–513.

Amaral, MM do, G Ceccantini. 2011. The endoparasite *Pilostyles ulei* (Apodanthaceae – Cucurbitales) influences wood structure in three host species of *Mimosa*. *IAWA Journal* 32: 1–13.

Baele, G, P Lemey, T Bedford, A Rambaut, MS Suchard, AV Alekseyenko. 2012. Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol. Bio. Evol.* 29: 2157–2167.

Barbrook, AC, CJ Howe, S Purton. 2006. Why are plastid genomes retained in non-photosynthetic organisms? *Trends Plant Sci.* 11: 101–108.

Barkman, TJ, S-H Lim, KM Salleh, J Nais. 2004. Mitochondrial DNA sequences reveal the photosynthetic relatives of Rafflesia, the world's largest flower. *PNAS* 101: 787–792

Barkman TJ, JR McNeal, SH Lim, G Coat, HB Croom, ND Young, CW dePamphilis. 2007. Mitochondrial DNA suggests at least 11 origins of parasitism in angiosperms and reveals genomic chimerism in parasitic plants. *BMC Evol. Biol.* 7: 248.

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

Barrett, CF, JV Freudenstein, J Li, DR Mayfield-Jones, L Perez, JC Pires, C Santos. 2014. Investigating the path of plastid genome degradation in an early-transitional clade of heterotrophic orchids, and implications for heterotrophic angiosperms. *Mol Biol Evol* doi: 10.1093/molbev/msu252.

Bellot, S, SS Renner. 2013. Pollination and mating systems of Apodanthaceae and the distribution of reproductive traits in parasitic angiosperms. *Am. J. Bot.* 100: 1083–1094.

Bellot, S, SS Renner. 2014a. The systematics of the worldwide endoparasite family Apodanthaceae (Cucurbitales), with a key, a map, and color photos of most species. *Phytokeys* 36: 41–57.

Bellot, S, SS Renner. 2014b. Exploring new dating approaches for parasites: The worldwide Apodanthaceae (Cucurbitales) as an example. *Mol. Phyl. Evol.* 80: 1-10.

Bergthorsson, U, KL Adams, B Thomason, JD Palmer. 2003. Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* 424: 197–201.

Blarer, A, DL Nickrent, PK Endress. 2004. Comparative floral structure and systematics in Apodanthaceae (Rafflesiales). *Plant Syst. Evol.* 245: 119–142.

Brasil, B de Azevedo. 2010. Life cycle, phenology and floral anatomy of *Pilostyles* (Apodanthaceae - Rafflesiaceae s.l.): phylogenetic position allowances for Apodanthaceae family. Abstract of Master Thesis. Instituto de Biociências. Sao Paulo

Bromham, L, PF Cowman, R Lanfear. 2013. Parasitic plants have increased rates of molecular evolution across all three genomes. *BMC Evol. Biol.* 13: 126.

Bouckaert, R, J Heled, D Kühnert, T Vaughan, C-H Wu, D Xie, MA Suchard, A Rambaut, AJ Drummond. 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLOS Comput. Biol.* 10: e1003537.

Bouman, F, W Meijer. 1994. Comparative structure of ovules and seeds in *Rafflesiaceae*. *P1. Syst. Evol.* 193: 187-212.

Coates, AG, LS Collins, MP Aubry, WA Berggren. 2004. The geology of the Darien, Panama, and the late Miocene–Pliocene collision of the Panama arc with northwestern South America. *Geol. Soc. Am. Bull.* 116: 1327–1344.

Conn, JS, TR Wentworth, U Blum. 1980. Patterns of dioecism in the flora of the Carolinas. *American Midland Naturalist* 103: 310-315.

Cui, Z, Z Guo, J Miao, Z Wang, Q Li, X Chai, M Li. 2013. The genus *Cynomorium* in China: An ethnopharmacological and phytochemical review. *Journal of Ethnopharmacology* 147: 1–15.

Crisp, MD, LG Cook. 2009. Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. *Evolution* 63: 2257–2265.

Davis, CC, M Latvis, DL Nickrent, KJ Wurdack, DA Baum. 2007. Floral gigantism in Rafflesiaceae. *Science* 315: 1812.

Delannoy, E, S Fujii, CC des Francs, M Brundrett, I Small. 2011. Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. *Mol Biol Evol* 28: 2077–2086.

Dell, B, AH Burbidge. 1981. Notes on the biology of *Pilostyles* (Rafflesiaceae) in Western Australia. *Western Australian Herbarium Research Notes* 5: 71–79.

Dell, B, J Kuo, AH Burbidge. 1982. Anatomy of *Pilostyles hamiltonii* C. *A.* Gardner (Rafflesiaceae) in Stems of *Daviesia*. *Aust. J. Bot.* 30: 1–9.

Downie, SR, DS Katz-Downie, KH Wolfe, PJ Calie, JD Palmer JD. 1994. Structure and evolution of the largest chloroplast gene (ORF2280): internal plasticity and multiple gene loss during angiosperm evolution. *Curr Genet* 25: 367–378.

Dornburg, A, MC Brandley, MR McGowen, TJ Near. 2012. Relaxed clocks and inferences of heterogeneous patterns of nucleotide substitution and divergence time estimates across whales and dolphins (Mammalia: Cetacea). *Mol. Biol. Evol.* 29: 721–736.

Drummond, AJ, SYW Ho, MJ Phillips, A Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* 4: e88.

Drummond, AJ, MA Suchard. 2010. Bayesian random local clocks, or one rate to rule them all. *BMC Biol.* 8: 114.

Drummond, AJ, MA Suchard, D Xie, A Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. and Evol.* 29: 1969–1973.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17: 368–376.

Filipowicz, N, SS Renner. 2010. The worldwide holoparasitic Apodanthaceae confidently placed in the Cucurbitales by nuclear and mitochondrial gene trees. *BMC Evol. Biol.* 10: 219.

Fuentes, I, S Stegemann, H Golczyk, D Karcher, R Bock. 2014. Horizontal genome transfer as an asexual path to the formation of new species. *Nature* 511: 232–235.

Funk, H, S Berg, K Krupinska, U Maier, K Krause. 2007. Complete DNA sequences of the plastid genomes of two parasitic flowering plant species, *Cuscuta reflexa* and *Cuscuta gronovii*. *BMC Plant Biol* **7**: 45.

Gao C, X Ren, AS Mason, H Liu, M Xiao, J Li, D Fu. 2014. Horizontal gene transfer in plants. *Funct Integr Genomics* 14: 23–9.

Gomes, AL, GW Fernandes. 1994. Influence of Parasitism by *Pilostyles ingae* (Rafflesiaceae) on its Host Plant *Mimosa naguirei* (Leguminosae). *Annals of Botany* 74: 205–208.

González, F, N Pabón-Mora. 2014a. First reports and generic descriptions of the achlorophyllous holoparasites Apodanthaceae (Cucurbitales) of Colombia. *Revista Actualidades Biologicas* 36: 101.

González, F, N Pabón-Mora. 2014b. *Pilostyles boyacensis*, a new species of Apodanthaceae (Cucurbitales) from Colombia with the highest elevation record in the Americas. *Phytotaxa* 178.

Gutiérrez-García, TA, E Vázquez-Domínguez. 2013. Consensus between genes and stones in the biogeographic and evolutionary history of Central America. *Quaternary Res.* 79, 311–324.

Hardy, NB, LG Cook. 2012. Testing for ecological limitation of diversification: a case study using parasitic plants. *American Naturalist* 180: 438–449.

Harris, ME, G Meyer, T Vandergon, V Oberholzer Vandergon. 2013. Loss of the acetyl-CoA Carboxylase (*accD*) gene in Poales. *Plant Molecular Biology Reporter* 31: 21–31.

Herendeen, PS, PR Crane, AN Drinnan. 1995. Fagaceous flowers, fruits, and cupules from the Campanian (Late Cretaceous) or central Georgia, USA. *Int. J. Plant Sci.* 156: 93–116.

Ho, SYW, S Duchene. 2014. Molecular-clock methods for estimating evolutionary rates and timescales. *Molecular Ecology* doi: 10.1111/mec.12953.

Iorizzo, M, D Grzebelus, D Senalik, M Szklarczyk, D Spooner, P Simon. 2012a. Against the traffic: the first evidence for mitochondrial DNA transfer into the plastid genome. *Mob Genet Elements* 2: 261–266.

Iorizzo, M, D Senalik, M Szklarczyk, D Grzebelus, D Spooner, P Simon. 2012b. De novo assembly of the carrot mitochondrial genome using next generation sequencing of whole genomic DNA provides first evidence of DNA transfer into an angiosperm plastid genome. *BMC Plant Biol.* 12: 61.

Isono, K, Y Niwa, K Satoh, H Kobayashi. 1997. Evidence for transcriptional regulation of plastid photosynthesis genes in *Arabidopsis thaliana* roots. *Plant Physiol* 114: 623–630.

Jarvis, P, E López-Juez. 2013. Biogenesis and homeostasis of chloroplasts and other plastids. *Nat Rev Mol Cell Biol* 14: 787–802.

Jensen, PE, D Leister. 2014. Chloroplast evolution, structure and functions. *F1000Prime Rep* 6: 40.

Jokat, W, T Boebel, M König, U Meyer. 2003. Timing and geometry of early Gondwana breakup. *Journal of Geophysical Research: Solid Earth* (1978–2012) 108.

Kass, RE, AE Raftery. 1995. Bayes factors. *Journal of the American Statistical Association* 90: 773–795.

Keeling, PJ. 2013. The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annual Review of Plant Biology* 64: 583–607.

Kikuchi, S, J Bédard, M Hirano, Y Hirabayashi, M Oishi, M Imai, M Takase, T Ide, M Nakai. 2013. Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* 339: 571–574.

Kim, G, ML LeBlanc, EK Wafula, CW dePamphilis, JH Westwood. 2014. Genomic-scale exchange of mRNA between a parasitic plant and its hosts. *Science* 345: 808–811.

Koulintchenko, M, Y Konstantinov, A Dietrich. 2003. Plant mitochondria actively import DNA via permeability transition pore complex. *EMBO J.* 22: 1245–1254.

Krause, K. 2012. Plastid genomes of parasitic plants: A trail of reductions and losses. In CE Bullerwell (ed.), Organelle Genetics Springer-Verlag Berlin Heidelberg pp. 79–103.

Kubo, T, KJ Newton. 2008. Angiosperm mitochondrial genomes and mutations. *Mitochondrion* 8: 5–14.

Lemaire, B, S Huysmans, E Smets, VMerckx. 2011. Rate accelerations in nuclear 18S rDNA of mycoheterotrophic and parasitic angiosperms. *J Plant Res* 124: 561–576.

Li, X, T-C Zhang, Q Qiao, Z Ren, J Zhao, T Yonezawa, M Hasegawa, MJC Crabbe, J Li, Y Zhong. 2013. Complete chloroplast genome sequence of holoparasite *Cistanche deserticola* (Orobanchaceae) reveals gene loss and horizontal gene transfer from its host Haloxylon ammodendron (Chenopodiaceae). *PLOS One* 8: e58747.

Logacheva, MD, MI Schelkunov, AA Penin. 2011. Sequencing and analysis of plastid genome in mycoheterotrophic orchid *Neottia nidus-avis*. *Genome Biol. Evol.* 3: 1296–1303.

Logacheva, MD, MI Schelkunov, MS Nuraliev, TH Samigullin, AA Penin. 2014. The plastid genome of mycoheterotrophic monocot *Petrosavia stellaris* exhibits both gene losses and multiple rearrangements. *Genome Biol. Evol.* 6: 238–246.

Lovisa, A, S Gustafsson, CF Verola, A Antonelli. 2010. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus *Hoffmannseggella* (Orchidaceae: Epidendroideae). *BMC Evol Biol.* 10: 177.

Maia, VH, MA Gitzendanner, P Soltis, GK-S Wong, DE Soltis. 2014. Angiosperm phylogeny based on 18S/26S rDNA sequence data: constructing a large dataset using next-generation sequence data. *Int. J. Plant Sci.* 175: 613–650.

Marechal, A, N Brisson. 2010. Recombination and the maintenance of plant organelle genome stability. *New Phytol.* 186: 299–317.

Mileshina, D, M Koulintchenko, Y Konstantinov, A Dietrich. 2011. Transfection of plant mitochondria and in organello gene integration. *Nucleic Acids Research* 39: e115.

Naumann, J, K Salomo, JP Der, EK Wafula, JF Bolin, E Maass, L Frenzke, M-S Samain, C Neinhuis, CW dePamphilis, S Wanke. 2013. Single-copy nuclear genes place haustorial Hydnoraceae within Piperales and reveal a Cretaceous origin of multiple parasitic angiosperm lineages. *PLoS ONE* 8: e79204.

McNeal, JR, J Kuehl, J Boore, CW dePamphilis. 2007. Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. *BMC Plant Biol* 7: 57.

Martin, W, T Rujan, E Richly, A Hansen, S Cornelsen, T Lins, D Leister, B Stoebe, M Hasegawa, D Penny. 2002. Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *PNAS* 99: 12246–12251.

Molau, U. 1995. Reproductive ecology and biology. *In* M. C. Press and J. D. Graves [eds.], Parasitic Plants, 141–176. Chapman and Hall, London, England.

Molina, J, KM Hazzouri, D Nickrent, M Geisler, RS Meyer, MM Pentony, JM Flowers, P Pelser, J Barcelona, SA Inovejas, I Uy, W Yuan, O Wilkins, CI Michel, S Locklear, GP Concepcion, MD Purugganan. 2014. Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Mol. Biol. Evol.* 31: 793–803.

Montes, C, A Cardona, R McFadden, SEMorón, CA Silva, S Restrepo-Moreno, DA Ramírez, N Hoyos, J Wilson, D Farris, GA Bayona, CA Jaramillo, V Valencia, J Bryan, JA Flores. 2012. Evidence for middle Eocene and younger land emergence in central Panama: implications for isthmus closure. *Geol. Soc. Am. Bull.* 124: 780–799.

Mower, JP, S Stefanović, W Hao, JS Gummow, K Jain, D Ahmed, JD Palmer. 2010. Horizontal acquisition of multiple mitochondrial genes from a parasitic plant followed by gene conversion with host mitochondrial genes. *BMC Biology* 8: 150.

Neuhaus, HE, MJ Emes. 2000. Nonphotosynthetic metabolism in plastids. *Annual Review of Plant Physiology and Plant Molecular Biology* 51: 111–140.

Nickrent, DL, Y Ouyang, RJ Duff, CW dePamphilis. 1997. Do nonasterid holoparasitic flowering plants have plastid genomes? *Plant Molecular Biology* 34: 717–729.

Nickrent, DL, A Blarer, Y-L Qiu, DE Soltis, PS Soltis, M Zanis. 2002. Molecular data place Hydnoraceae with Aristolochiaceae. *Am. J. Bot.* 89: 1809–1817.

Nickrent, DL, A Blarer, Y-L Qiu, R Vidal-Russell, FE Anderson. 2004. Phylogenetic inference in Rafflesiales: the influence of rate heterogeneity and horizontal gene transfer. *BMC Evol. Biol.* 4: 40.

Nickrent, DL, JP Der, FE Anderson. 2005. Discovery of the photosynthetic relatives of the "Maltese mushroom" *Cynomorium*. *BMC Evol. Biol.* 5: 38.

Palmer, JD. 1991. Plastid chromosomes: structure and evolution. In: Cell Culture and Somatic Genetics of Plant, Vol. 7A, Molecular Biology of Plastids, eds. L. Bogorad and I.K. Vasil, Academic Press, San Diego, pp. 5–53.

Parker, C. 2009. Observations on the current status of *Orobanche* and *Striga* problems worldwide. *Pest Management Science* 65: 453–459.

Player, G. 1979. Pollination and wind dispersal of pollen in *Arceuthobium*. *Ecological Monographs* 49: 73–87.

Poiteau, MA. 1824. Description de l'*Apodanthes* nouveau genre de plantes phanérogames parasite. *Annales des Sciences Naturelles* 3: 421–425.

Rannala, B, Z Yang. 2007. Inferring speciation times under an episodic molecular clock. *Syst. Biol.* 56: 453–466.

Renner, SS, RE Ricklefs. 1995. Dioecy and its correlates in the flowering plants. *American Journal of Botany* 82: 596–606.

Renner, SS, S Bellot. 2012. Horizontal gene transfer in eukaryotes, focusing on fungi-to-plant and plant-to-plant transfers. Pp. 223-235 in R. Bock & V. Knoop, eds., Genomics of Chloroplasts and Mitochondria. Springer, Heidelberg.

Rice, DW, AJ Alverson, AO Richardson, GJ Young, MV Sanchez-Puerta, J Munzinger, K Barry, JL Boore, Y Zhang, CW dePamphilis, EB Knox, JD Palmer. 2013. Horizontal Transfer of Entire Genomes via Mitochondrial Fusion in the Angiosperm *Amborella*. *Science* 342: 1468–1473.

Richly, E, D Leister. 2004. NUPTs in sequenced eukaryotes and their genomic organization in relation to NUMTs. *Mol Biol Evol.* 21: 1972–1980.

Rothfels, CJ, E Schuettpelz. 2014. Accelerated rate of molecular evolution for vittarioid ferns is strong and not driven by selection. *Syst. Biol.* 63, 31–54.

Rutherford, RJ. 1970. The Anatomy and Cytology of *Pilostyles thurberi* Gray (Rafflesiaceae). *Aliso* 7: 263–288.

Sarich, VM, AC Wilson. 1967. Immunological time scale for hominid evolution. *Science* 158: 1200–1203.

Schaefer, H, SS Renner. 2011. Phylogenetic relationships in the order Cucurbitales and a new classification of the gourd family (Cucurbitaceae). Taxon 60: 122–138.

Sheahan, MB, DW McCurdy, RJ Rose. 2005. Mitochondria as a connected population: ensuring continuity of the mitochondrial genome during plant cell dedifferentiation through massive mitochondrial fusion. *The Plant Journal* 44: 744–755.

Shih, PM, NJ Matzke. 2013. Primary endosymbiosis events date to the later Proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins. *PNAS* 110: 12355–12360.

Smith, DR. 2011. Extending the limited transfer window hypothesis to inter-organelle DNA migration. *Genome Biol Evol* 3: 743–748.

Smith, DR, RW Lee. 2014. A plastid without a genome: evidence from the nonphotosynthetic green algal genus *Polytomella*. *Plant Physiology* 164: 1812–1819.

Stegemann, S, R Bock. 2009. Exchange of genetic material between cells in plant tissue grafts. *Science* 324: 649–651.

Stegemann, S, M Keuthe, S Greiner, R Bock. 2012. Horizontal transfer of chloroplast genomes between plant species. *Proc. Natl. Acad. Sci. USA* 109: 2434–2438.

Stern, DB, DM Lonsdale. 1982. Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. *Nature* 299: 698–702.

Straub, SCK, RC Cronn, C Edwards, M Fishbein, M Liston. 2013. Horizontal transfer of DNA from the mitochondrial to the plastid genome and its subsequent evolution in milkweeds (Apocynaceae). *Gen. Biol. and Evol*. 10: 1872–1885.

Su, H-J, J-M Hu. 2012. Rate heterogeneity in six protein-coding genes from the holoparasite *Balanophora* (Balanophoraceae) and other taxa of Santalales. *Annals of Botany* 110: 1137–1147.

The Tree of Sex Consortium. 2014. Tree of Sex: A database of sexual systems. Sci. Data 1:140015 doi: 10.1038/sdata.2014.15

Thiele, K. R., S. J. Wylie, L. Maccarone, P. Hollick, J. A. McComb. 2008. *Pilostyles coccoidea* (Apodanthaceae), a new species from Western Australia described from morphological and molecular evidence. *Nuytsia* 18: 273–284.

Van Duzer, C. 2004. Floating islands: a global bibliography. Cantor Press, Los Altos Hills, CA

Van Duzer, C. 2006. Addenda to Floating islands: a global bibliography. Cantor Press, Los Altos Hills, CA.

Vattimo, I de. 1950. *Pilostyles stawiarskii*. *Revista Brasileira de Biologia* 10: 196.

Vattimo, I de. 1955. Notice sur la tribu Apodantheae R.Br. (Rafflesiaceae). *Taxon* 4: 211–212

Vattimo, I de. 1971. Contribuicao ao conhecimento da tribo Apodanthea R. Br. Parte I – Conspecto das especies (Rafflesiaceae). *Rodriguesia* 38: 37–62.

Vattimo, I de. 1973. Notas sobre o genero *Apodanthes* Poit. com descricao de duas novas especies (Rafflesiaceae). *Revista Brasileira de Biologia* 33: 140.

Vidal, N, A Azvolinsky, C Cruaud, S Blair Hedges. 2008. Origin of tropical American burrowing reptiles by transatlantic rafting. *Biol. Lett.* 4: 115–118.

Wicke S, GM Schneeweiss, CW dePamphilis, KF Müller, D Quandt. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol* 76: 273–297.

Wicke, S, KF Müller, dePamphilis CW, D Quandt, NJ Wickett, Y Zhang, SS Renner, GM Schneeweiss. 2013. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *The Plant Cell* 25: 3711–3725.

Wilson, WG, and LD Harder. 2003. Reproductive uncertainty and the relative competitiveness of simultaneous hermaphroditism versus dioecy. *American Naturalist* 162: 220–241.

Wolfe, KH, CW Morden, JD Palmer. 1992. Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc Natl Acad Sci USA* 89: 10648–10652.

Woloszynska, M. 2010. Heteroplasmy and stoichiometric complexity of plant mitochondrial genomes — though this be madness, yet there's method in't. *J Exp Bot* 61: 657–671.

Won, H, SS Renner. 2003. Horizontal gene transfer from flowering plants to *Gnetum*. *PNAS* 100: 10824–10829.

Xi, Z, RK Bradley, KJ Wurdack, KM Wong, M Sugumaran, K Bomblies, JS Rest, CC Davis. 2012. Horizontal transfer of expressed genes in a parasitic flowering plant. *BMC Genomics* 13: 227.

Xi, Z, Y Wang, RK Bradley, M Sugumaran, CJ Marx, JS Rest, CC Davis. 2013. Massive mitochondrial gene transfer in a parasitic flowering plant clade. *PLoS Genet* 9: e1003265.

Yoshida, T, HY Furihata, A Kawabe. 2014. Patterns of genomic integration of nuclear chloroplast DNA fragments in plant species. *DNA Res* 21: 127–140.

Zhang, Z-H, C-Q Li, J Li. 2009. Phylogenetic placement of *Cynomorium* in Rosales inferred from sequences of the inverted repeat region of the chloroplast genome. *J. Syst. Evol.* 47: 297–304.

Zoschke, R, K Liere, T Börner. 2007. From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. *Plant Journal* 50: 710–722.

APPENDIX

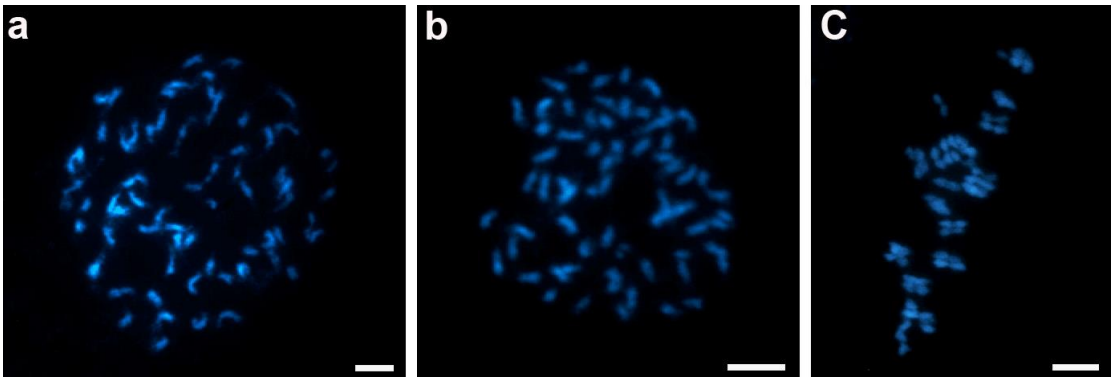

Figure 1. Sporophytic (diploid) and gametophytic (haploid) chromosome numbers of *Pilostyles haussknechtii*. **a.** Mitosis with chromosomes in prometaphase 2n = ca. 66. **b.** Mitosis with chromosomes is metaphase *2n* = 66. **c.** Meiosis II with bivalents starting to divide *n* = ca. 33. Scale bars represent 5 µm.

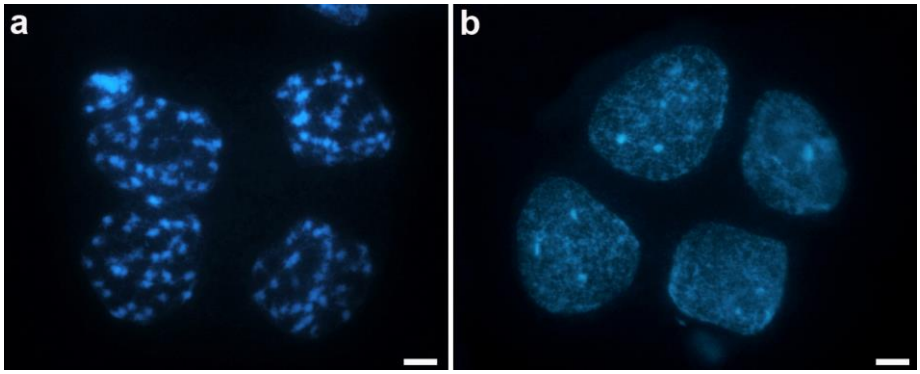

Figure 2. Gametophytic (haploid) chromosome number of *Pilostyles aethiopica*. **a.** Early tetrad *n* = 29. **b.** Late tetrad. Scale bars represent 5 µm.

ACKNOWLEDGEMENTS

I express my warmest thanks to Professor S. S. Renner, for her unequalled availability, her patience, her thoroughness, and her continued trust.

I am also indebted to M. Silber, for her good mood, her devotion in the lab and her help with paperwork. I am also grateful to the people from the institute who were always eager to help whenever I had a technical problem, especially to G. Gerlach for providing field material, teaching me the basics of photography, and sharing his wonderful parasites pictures with me.

My research would not have been possible without the people who shared their house, their time and their knowledge with me in Australia: Annette, K. Dixon, J. McComb, K. Thiele, and S. Wylie; in Iran: Majid, Mariam, Y. Salmaki, S. Zarre, and two unrivalled (rallye) drivers; and in Zimbabwe: M. Coates Palgrave, M. and L. Hyde, Tuwanda, the Mukuvisi woodlands staff, D. Plowes, Nina and Kassirai.

I also thank Dr. A. Sousa dos Santos and E. Temsch, for their help with the chromosome work, the people of the Zoolosische Staatssammlung München for identifying insects, S. Dötterl for his advises regarding fragrance analysis and K. Krause, S. Wanke and S. Wicke for insightful discussions and technical help.

The past and present students and researchers of the institute, some of which are now unforgettable friends, contributed to enlighten my Ph.D. with their sympathy as well as with their precious knowledge: S. Abrahamczyk, J. Bechteler, C. Braüchler, F. Antunes Carvalho, J. Chacón, M. Chartier, G. Chomicki, N. Cusimano, N. Filipowicz, A. Fleischman, M. Gottschling, A. Gröger, N. Holstein, F. Jabbour, D. Maba, A. Martins, L. Nauheimer, O. Pérez, M. Piednoel, A. Rockinger, A. Scheunert, A. Sousa dos Santos, J.-C. Villarreal, C. Zohner…

Finally, my family, my d., and my friends from here and there largely contributed to the good achievement of this Ph.D. by keeping me grounded. Merci, gracias, obrigada, danke, thank you all for your understanding, your support and your affection!

CURRICULUM VITAE

Name: Sidonie Bellot
Date of birth and nationality: 1987-02-17, French
Address: LMU Department für Biologie, Systematische Botanik und Mykologie, Menzinger Straße 67, 80638 München
E-mail: sido.bellot@gmail.com

EDUCATION

2005	Baccalauréat, specializing in Biology, Lycée Majorelle - Toul (France)
2005-2008	Bachelor of Life Sciences**,** specializing in cellular biology and ecology, ethology and botany, University Nancy 1 (France)
2008-2009	First year of Master's degree: "Integrative Plant Biology: Genes, Plant, Agrosystem", University of Rennes 1 (France)
2009-2010	Second year of Master's degree: "Biodiversity, Ecology, Evolution", University of Montpellier 2 (France)
2010-2014	Ph.D: "Natural history, taxonomy, biogeography and genome evolution of the widespread endoparasites Apodanthaceae (Cucurbitales); a contribution to the understanding of the peculiar evolution of holoparasites" Advisor: Prof. S. S. Renner, Ludwig Maximilian University of Munich (Germany)

INTERNSHIPS

August 2007 **Voluntary training**
Topic: Installation of a "Biologic Integral Reserve" in the forest of Parroy (54, France)
Methods, tasks: Plant species inventories and dendrometric measurements
Organisation: National Office of Forestry (ONF Meurthe-et-Moselle) France

April 2008 **Voluntary training**
Topic: Relationship between stomatal density and water-use efficiency
Methods, tasks: Stomates counting from photos of leaves of different European trees
Organisation: "Forest Ecology and Ecophysiology" (Research unit depending on the National Institute of Agronomic Research – INRA –and the University of Nancy 1) – Nancy, France

August 2008 **Voluntary research training**
Topic: Physiological responses of *Eucalyptus globulus*, *E. cladocalyx* and *Acacia mangium* to drought. Relationship between physiological state and carbon isotopic composition of phloem sap. Variations of carbon allocation between saccharose and cyclic hexitols pools.
Methods, tasks: ^{13}C marking, gas fluxes measurements, collecting of phloem and xylem saps. Organisation: "Forest Ecology and Ecophysiology" (Research unit depending on the National Institute of Agronomic Research – INRA –and the university of Nancy 1) – Nancy, France

April-June 2009 **First year of Master's degree's research training**
Topic: Phylogeny of the genus *Ulex* (Fabaceae) inferred from chloroplast DNA polymorphism. Origin of the chloroplast genome of the allohexaploïd invasive European gorse (*U. europaeus europaeus*).
Methods, tasks: PCR, sequences alignments, phylogeny (maximum of parsimony)

Tutor: Abdelkader Aïnouche, "ECOBIO", research unit belonging to the "Centre National de la Recherche Scientifique – CNRS – and the University of Rennes.

January -June 2010 Second year of Master's degree's research training

Topic: Reconstruction from mass sequencing data (454 GS-FLX) of the complete chloroplast genome of the hexaploïd *Spartina maritima* (Chloridoideae, Poaceae). Study of chloroplast genome evolution in Poaceae. Use of chloroplast markers to date speciation events inside the *Spartina* clade.

Methods, tasks: Local and global alignments, PCR, phylogeny by maximum of parcimony and maximum likelihood, dating by bayesian inference

Tutor: Malika Aïnouche, also at ECOBIO.

ORAL PRESENTATIONS

Bellot, S. and Renner, S.S.: The evolutionary retention of plastid genomes in non-photosynthetic plants: A comparative approach centred on the endoparasitc Apodanthaceae. Third symposium on the biology of non-weedy parasitic plants – Namur, Belgium, September 2013

Bellot, S., Ainouche, K., Chelaifa, H., Coudouel, S., Lima, O., Naquin, D. and Ainouche, M.: Reconstruction of the chloroplast genome of *Spartina maritima*, and molecular dating of the divergence between the two genomes reunited in the recent allopolyploid *S. anglica*. International Botanical congress – Melbourne, Australia, July 2011

Bellot, S. and Ainouche, M.: The complete chloroplast genome of *Spartina maritima* (Pocaeae, Chloridoideae) using 454 sequencing. Berlin Biosystematics – 22 February 2011

Bellot, S.: Evolution of chloroplast genome in *Spartina* (Poaceae). Cytology and Polyploïdy meeting – INRA, Versailles, France, 18 May 2010

PUBLICATIONS

Bellot, S., Renner, S.S., 2014. Exploring new dating approaches for parasites: the worldwide Apodanthaceae (Cucurbitales) as an example. Molecular Phylogenetics and Evolution 80: 1-10

Bellot, S., Renner, S.S., 2014. The systematics of the worldwide endoparasite family Apodanthaceae (Cucurbitales), with a key, a map, and color photos of most species Phytokeys 36: 41-57.

Bellot, S., Renner, S.S., 2013. Pollination and mating systems of Apodanthaceae and the distribution of reproductive traits in parasitic angiosperms. Am. J. Bot. 100, 1083–1094.

Renner, S.S., Bellot, S., 2012. Horizontal gene transfer in eukaryotes, focusing on fungi-to-plant and plant-to-plant transfers. Pp. 223-235 in R. Bock & V. Knoop, eds., Genomics of Chloroplasts and Mitochondria. Springer, Heidelberg.

Merchant, A., Wild, B., Richter, A., Bellot, S., Adams, M.A., Dreyer, E., 2011. Compound-specific differences in $_{13}$C of soluble carbohydrates in leaves and phloem of 6-month-old *Eucalyptus globulus* (Labill). Plant, Cell and Environment 34, 1599–1608.

Bellot, S., Rousseau-Gueutin, M., Martin, G., Boutte, J., Chelaifa, H., Lima, O., Michon-Coudouel, S., Naquin, D., Salmon, A., Aïnouche, A., Aïnouche, M., *in prep*. The complete chloroplast genome of *Spartina maritima*: insights from comparative analyses in the Poaceae family and molecular dating.