



Caught in action: fine-scale plastome evolution in the parasitic plants of *Cuscuta* section *Ceratophorae* (Convolvulaceae)

Arjan Banerjee^{1,2} · Saša Stefanović¹

Received: 14 February 2019 / Accepted: 18 May 2019 / Published online: 28 May 2019
© Springer Nature B.V. 2019

Abstract

Key message An exhaustive analysis of a group of closely related parasitic plants shows a predominantly gradual reduction in plastid genome composition and provides the most reduced plastomes in the genus *Cuscuta*.

Abstract Parasitic plants have a diminished to completely absent reliance on photosynthesis, and are characterized by sweeping morphological, physiological, and genomic changes. The plastid genome (plastome) is highly conserved in autotrophic plants but is often reduced in parasites, and provides a useful system for documenting the genomic effects of a loss of photosynthesis. Previous studies have shown a substantial degree of heterogeneity in plastome length and composition across the species of the genus *Cuscuta*. Specifically, species in *Cuscuta* sect. *Ceratophorae* were suspected to exhibit even more dynamic plastome evolution than the rest of the genus. This complex of eight closely related species was exhaustively sampled here, and one accession per species was sequenced via a high-throughput approach. Complete plastid genomes were assembled and annotated for each of these species and were found to be 61–87 kbp in length, representing a 45–60% reduction relative to autotrophic Convolvulaceae. The most reduced plastomes on this spectrum have lost the bulk of their photosynthetic genes and are the first fully holoparasitic plastomes described for *Cuscuta*. The fine-scale nature of the system introduced here allowed us to phylogenetically triangulate the locations of gene loss and pseudogenization events precisely, and to construct a step-by-step model of plastome evolution in these plants. This model reveals an intense burst of gene loss along the branch leading to the most reduced plastomes, and a few idiosyncratic changes elsewhere, allowing us to conclude that the tempo of plastid evolution in sect. *Ceratophorae* is a blend of gradual and punctuated mode.

Keywords *Ceratophorae* · Convolvulaceae · *Cuscuta* (dodders) · Holoparasites · Parasitic plants · Plastome

Introduction

Photosynthesis is the hallmark of green-plant evolution; however, this process has been lost repeatedly, most often observed now in flowering plants. As a consequence, parasitic plants have reduced to completely absent reliance on photosynthesis and are instead dependent on autotrophs for

meeting their carbon budget. True parasitism, which requires the forging of direct connections between the vascular systems of the host and parasite via haustoria, has evolved independently at least 13 times across angiosperm lineages and is associated with sweeping morphological, physiological and genomic changes brought about by a decreased reliance or complete lack of photosynthesis (Barkman et al. 2007; Colwell 1994; Kuijt 1969; Westwood et al. 2010). This so-called ‘parasitic reduction syndrome’ defines a set of convergent evolutionary changes shared amongst distinct groups of heterotrophs (including both haustorial parasites as well as mycotrophic plants) and provides a fertile system to study genomic changes caused by a relaxation of erstwhile strong selective pressures.

The genomic changes observed as a result of a transition to a parasitic lifestyle are expected to be particularly striking in the plastid genomes (plastomes) of heterotrophic plants given the importance of plastids as the site of photosynthesis

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11103-019-00884-0>) contains supplementary material, which is available to authorized users.

✉ Arjan Banerjee
arjan.banerjee@mail.utoronto.ca

¹ Department of Biology, University of Toronto Mississauga, Mississauga, ON L5L 1C6, Canada

² Ecology and Evolutionary Biology, University of Toronto, Toronto, ON M5S 2Z9, Canada

in the cell. Photosynthetic land plants typically have plastomes that are subjected to strong purifying selection and are consequently highly conserved in terms of their size, gene content, and sequence order (Wolfe et al. 1992). They are usually 140 to 160 kbp in length (e.g., *Nicotiana tabacum*: 155,939 bp, *Arabidopsis thaliana*: 154,478 bp), coding for c. 113 genes and contain four major structural components: the large and small single copy regions (LSC and SSC respectively) and two inverted repeat regions (IRs) (Downie and Palmer 1992; Wicke et al. 2011). Most plastome genes code for key portions of the photosynthetic apparatus or for plastid housekeeping functions, although there are a few genes with other or unknown functions (Braukmann et al. 2017; Wicke et al. 2011). Hence, plastomes in heterotrophic plants are often found to have undergone drastic reductions in size and gene composition. A series of recent studies highlight heterotrophic plastomes of severely reduced length and content (Barrett et al. 2018; Bellot and Renner 2015; Braukmann et al. 2017; Naumann et al. 2016), and there is even indication that plants may have lost their plastid genomes entirely (e.g., *Rafflesia lagascae*) (Molina et al. 2014).

Comparative analyses of plastomes along the full trophic spectrum, from autotrophy to mixotrophy to full heterotrophy, would allow for assessment of the degree to which genomic changes take place prior to complete loss of photosynthesis, and the dissection of the evolutionary pressures on plastomes exerted by their other metabolic functions. Over time, as ever greater numbers of parasitic and mycoheterotrophic plant plastomes are being reported, the general pattern of sequence loss linked to heterotrophy is emerging across the angiosperms: *ndh* genes (primarily responsible for mitigating the effects of photo-oxidative stress) are usually the first system of genes to be lost, followed by groups of photosynthesis-related genes (*psa*, *psb*, *pet* etc.), while a core group of non-bioenergetic genes are generally retained even in severely reduced genomes (Barrett and Davis 2012; Graham et al. 2017). However, these initial observations still need broader support from plastomes in heterotrophic lineages which are underrepresented in current research. Furthermore, the currently available plastome data are mainly coarse in the scale of their sampling as far as phylogenetic distance between species is concerned, with notable exceptions being provided by the recent work on *Corallorhiza* (Orchidaceae; Barrett et al., 2018) and on *Aphyllon* (Orobanchaceae; Schneider et al., 2018). Fine-scale comparisons of plastomes amongst closely related species are more likely to fully capture the sequence of evolutionary events as plants change along the gradient from photosynthetic to holoparasitic.

Owing to its species richness, broad distribution, and the fact that it represents one of the 13 independent transitions to haustorial parasitism within the angiosperms, *Cuscuta* (dodders, Convolvulaceae) is one of the most studied lineages of

parasitic plants. While these plants have a nearly world-wide distribution, c. 75% of the generic diversity is in the Americas (Garcia et al. 2014; Yuncker 1932). This lineage is comprised of c. 200 branch parasites characterized by scale-like leaves, twining, pale, slender stems and an absence of roots. Their unusual colour is attributed to absent or reduced accumulation of chlorophylls (van der Kooij et al. 2000), and species in this genus are all unable to meet their food requirements on their own (i.e., they are obligate parasites). Most *Cuscuta* species are thus nearly completely dependent upon their hosts and forge direct xylem and phloem connections with green plants for sourcing water and nutrients (Heide-Jorgensen 2008). Some dodders are, however, capable of limited and localized photosynthesis (Dawson et al. 1994; Hibberd et al. 1998) and have been referred to as ‘cryptically photosynthetic’ (McNeal et al. 2007a). This feature makes *Cuscuta* a heterogeneous group from a trophic point of view (Westwood et al. 2010), containing both hemi- and holoparasitic representatives. Also, phylogenetically, this genus is well-understood and circumscribed in four subgenera (*Grammica*, *Pachystigma*, *Cuscuta*, and *Monogynella*) and 19 sections (Costea et al. 2015).

Despite the intensity of scrutiny on the dodders over the years and the relatively large size of the genus, only four plastomes from *Cuscuta* have been published to date (Funk et al. 2007; McNeal et al. 2007b), two each from the subgenera *Monogynella* and *Grammica* (essentially representing two data points, owing to phylogenetic proximity). The plastomes from subg. *Monogynella* are ~ 120 kbp long (with c. 103 genes) while those from subg. *Grammica* are ~ 85 kbp in length (with c. 92 genes) demonstrating a great degree of variation in plastome length and gene composition in the genus. However, the tempo and mode of plastid genome evolution across the diversity of the group remain largely unknown. Recognizing the potential importance of *Cuscuta* as a system of closely-related plants that seem to exhibit fine-scale variation in a typically conserved and slow-evolving genome, Braukmann et al. (2013) surveyed the genus for the presence or absence of 48 protein-coding plastome genes across a phylogenetically diverse sample of 112 species through a Southern hybridization slot-blot assay. They identified confirmed trends of gene loss across the dodders (e.g., all *ndh* genes, *rpl23*, *rpl32*, *rps16*) but were also able to pinpoint two sections, both within the largest subgenus *Grammica*, that exhibited more varied and extensive reductions as particularly interesting for further studies (Braukmann et al. 2013).

One of these groups, section *Ceratophorae*, is remarkable for the variation in plastome content apparent among the five species sampled initially (Braukmann et al. 2013). This complex seemed to contain members that have ancestral plastomes (i.e., similar to previously reported *Grammica* samples) as well as those that appeared highly modified.

In other words, sect. *Ceratophorae* presents a gradient of plastome degradation in what is phylogenetically a group of very closely related species. This pattern appeared to be congruent with a progressive gradual change, as described by the Evolutionary Transition Series hypothesis (Mayr 1982; Young et al. 1999), and offers an opportunity to study the evolution of parasitic plastomes ‘in action’. The broad survey by Braukmann et al. presented interesting trends but was incomplete in terms of taxonomic sampling and gene coverage, contained no sequence data, and was thus unable to provide information at the most basic, sequence level, including none on elements like ribosomal and transfer RNA genes, introns, promoters, etc. Our current investigation builds on their work and comprehensively examines the evolution of the plastid genome in sect. *Ceratophorae* through the assembly of plastomes from all eight known species in this clade in order to learn how sequence changes accumulate in *Cuscuta*. We also compare these trends to the patterns of plastome evolution observed in other lineages of heterotrophic plants.

Materials and methods

Taxon sampling, DNA extraction, and sequencing

Cuscuta section *Ceratophorae* was essentially exhaustively sampled; all eight species confirmed to belong to this clade (Costea et al. 2015, 2011) were used for our study: *C. boldinghii* Urb., *C. bonafortunae* Costea and I. Garcia, *C. carnosae* Costea and Stefanovic, *C. chapalana* Yunck., *C. costaricensis* Yunck., *C. erosa* Yunck., *C. mexicana* Yunck., and *C. strobilacea* Liebm (voucher information available on NCBI; accession numbers listed in Table 1). A potential ninth species (*C. ortegana*) has been previously placed in *Ceratophorae* but there are doubts as to whether it actually belongs in this section, and therefore has been excluded here. Total genomic DNA was isolated from fresh, herbarium, or silica-dried tissue using the modified

cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1987) and checked for quantity and quality using a Nano Drop 1000 Spectrophotometer (Thermo Fisher Scientific). Extractions were sent for high-throughput sequencing on either an Illumina HiSeq 2500 platform (2 × 126 bp paired-end reads; The Centre for Applied Genomics, Sick-Kids Hospital, Toronto, Ontario) or an Illumina HiSeq 2000 platform (2 × 100 bp paired-end reads; McGill University and Genome Quebec, Montreal, Quebec). Demultiplexing of raw reads and the removal of indexing barcodes were performed at the sequencing facilities.

Plastome assembly, annotation, and correlation with Southern hybridization results

Reads were trimmed using Sickle v1.33 (Joshi and Fass 2011) with minimum post-trim read lengths set at 71 bp for the 2 × 100 bp reads and 99 bp for the 2 × 126 bp reads, and the threshold for quality set at a minimum PHRED score between 25 and 27 at each nucleotide. Several separate assemblies were conducted in the native Geneious (R9 and R10) (Biomatters, Auckland, New Zealand) de novo assembler using 10–15% of total paired-end reads (between 35,110,054 and 81,302,878 reads for each species) with the ‘produce scaffolds’ and ‘don’t merge variants’ boxes unchecked. The derived contigs were aligned and joined, with remaining gaps closed manually on Geneious.

Initial annotations for *Cuscuta costaricensis* were conducted on Geneious using *Cuscuta obtusiflora* (McNeal et al. 2007b) as a reference and then refined and confirmed using BLASTx searches (Altschul et al. 1990). The fully annotated *C. costaricensis* plastome was then used as a reference for the annotation of the remaining seven species in this study. *Cuscuta costaricensis* was chosen for this role because it has been shown to be sister for the rest of section *Ceratophorae* in previously published and well-supported phylogenies (Costea et al. 2011). BLASTx was used to establish all open reading frames and to identify pseudogenization events

Table 1 Plastid genome size and structure information for the eight species assembled from the *Cuscuta* section *Ceratophorae*

	Plastome size (bp)	Genes (protein/ tRNA/rRNA)	SSC (bp %)	IR (bp %)	NCBI accession
<i>C. costaricensis</i>	86,691	59/25/4	7355 (8.48)	14,549 (16.83)	MK881072
<i>C. chapalana</i>	84,607	59/21/4	8216 (9.71)	13,068 (15.45)	MK887214
<i>C. mexicana</i>	83,526	57/23/4	7170 (8.58)	13,242 (15.85)	MK887213
<i>C. bonafortunae</i>	82,346	59/24/4	7916 (9.61)	12,650 (15.36)	MK887215
<i>C. carnosae</i>	81,881	58/25/4	7536 (9.20)	12,891 (15.74)	MK887212
<i>C. strobilacea</i>	63,787	36/25/4	6860 (10.75)	13,102 (20.54)	MK867795
<i>C. boldinghii</i>	62,375	32/25/4	6699 (10.74)	12,793 (20.51)	MK881074
<i>C. erosa</i>	60,959	33/25/4	6275 (10.29)	12,544 (20.58)	MK881073

Detailed voucher information for each sample can be accessed at the corresponding National Center for Biotechnology Information (NCBI) accession numbers

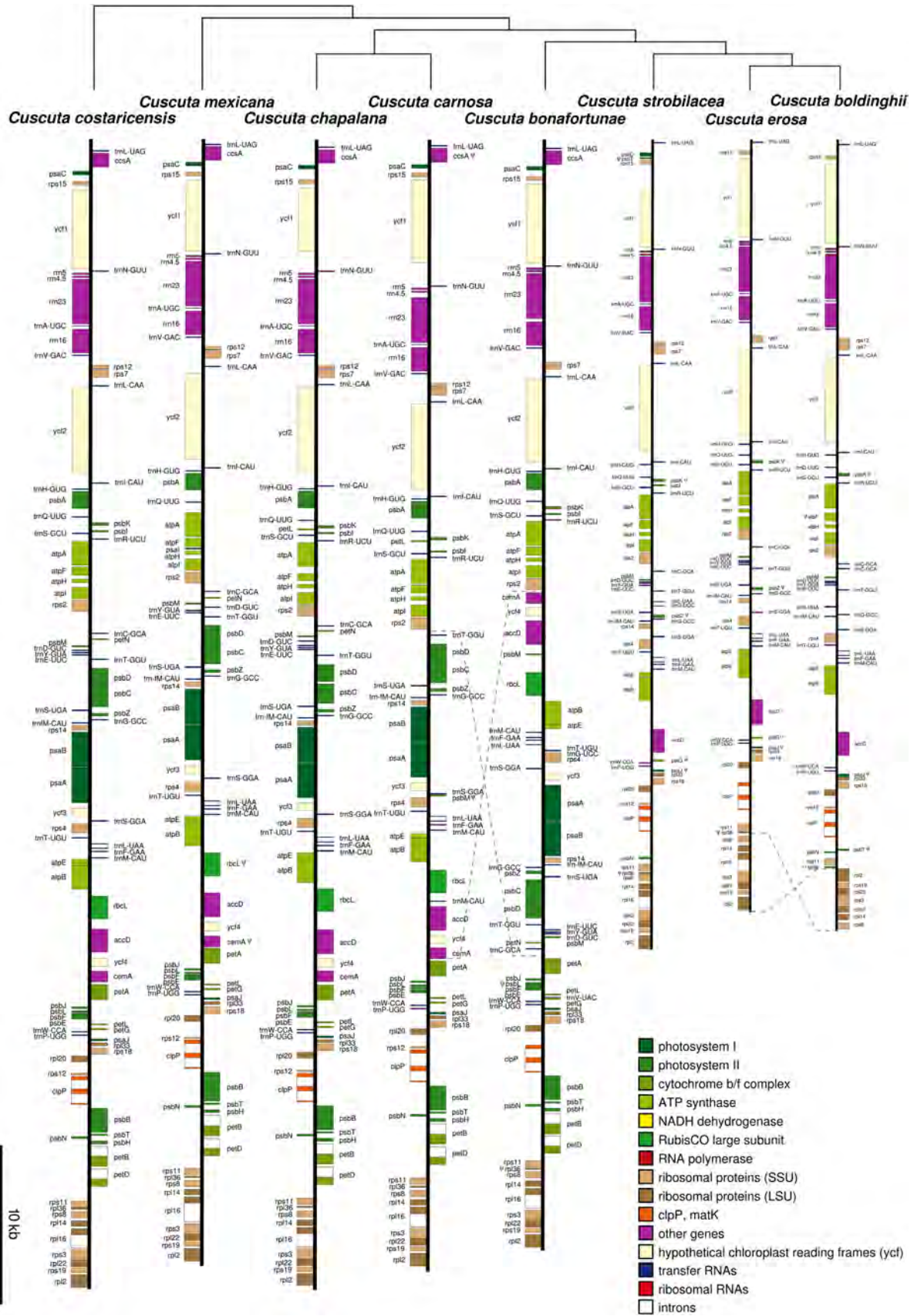


Fig. 1 Annotated plastid genomes from eight species belonging to *Cuscuta* section *Ceratophorae* created using OrganellarGenomeDRAW v1.3.1 (Greiner et al. 2019). For legibility, only one of the two inverted repeat regions are represented. Pseudogenes are represented by the ψ symbol beside the gene labels. Sequence inversions are shown by the gray dashed lines. The phylogenetic relationships are based on the whole-plastome phylogenetic analysis (compare with tree in Fig. 3)

while tRNAscan-SE 2.0 was used to identify and determine the boundaries of tRNAs (Lowe and Chan 2016). The genomes were aligned natively on Geneious for comparative analyses regarding size and composition, and using the progressiveMauve system (Darling et al. 2010) for structural comparisons.

The results generated by the whole plastome assemblies produced in this research were correlated with the results of the Southern hybridization survey conducted for the same species complex in Braukmann et al. 2013. Their survey had included five of the eight species included here. The slot-blot probe results were classified either as full strength ‘positive signal’ or as ‘reduced or no signal’ while the genes from the whole plastome data were denoted ‘present’ as open reading frames and considered functional, ‘pseudogenized’ and considered non-functional, or entirely ‘absent’. A goodness-of-fit (Chi square) analysis of independence (with one degree of freedom) was conducted for each species to compare the results from the two studies.

Phylogenetic analyses and computational methods

All protein coding regions greater than 150 bp in length present across the eight *Ceratophorae* plastomes sampled were aligned and used for the whole-plastome phylogenetic reconstruction. This amounted to 27,587 bp in the dataset. *Cuscuta obtusiflora*, also in subgenus *Grammica* and whose plastome has been previously published (McNeal et al. 2007b), was included as an outgroup. jModelTest 2 (Darrriba et al. 2012; Guindon and Gascuel 2003) was used to infer the best model for DNA substitution rates (identified to be GTR + G + I for these data). Given the moderate number of operational taxonomic units, a branch-and-bound search was conducted for the best tree in PAUP v4.0 (Wilgenbusch and Swofford 2003) ensuring the recovery of all optimal trees. Both maximum likelihood (ML) and parsimony were used as optimality criteria. The software was allowed to estimate all parameters of the model, including the frequency of state changes and the rate of site-to-site variations. To provide statistical support, bootstrapping was conducted under ML, with a full heuristic search, tree-bisection and reconnection branch swapping, and 10,000 replicates.

The dN and dS branch lengths for seven protein coding genes (*accD*, *ycf1*, *ycf2*, *atpB*, *atpH*, *clpP* and *rpl20*), chosen as representatives from the major functional groups of

genes that are retained in the plastomes of all eight species sampled, were estimated based on the same model of DNA evolution and using the fixed topology obtained from the above whole-plastome ML analysis. For dN trees, only bases occupying the first and second codon positions were used to calculate branch lengths because substitutions in these bases generally lead to nonsynonymous amino acid changes. Only third codon bases were used to calculate branch lengths for the dS trees because substitutions involving the wobble position usually cause synonymous change. The type of selection was estimated on a codon-by-codon basis using a fast, unconstrained Bayesian approximation for inferring selection (FUBAR) (Murrell et al. 2013) on the platform HyPhy (Pond et al. 2005) using 5 independent MCMC chains to obtain posterior samples of grid point weights based on 2,000,000 total steps and sampling every 10,000 steps. The phylogenetic relationships used for the FUBAR analysis was inferred using plastome data in this research (as described previously).

Results

Plastome reductions in *Cuscuta* sect. *Ceratophorae*

Closed circular plastomes were assembled for each of the eight species sampled from section *Ceratophorae*, ranging from 86,691 bp to 60,959 bp in length (Table 1, Fig. 1). Three species (*Cuscuta boldinghii*, *C. strobilacea*, and *C. erosa*) are substantially more reduced than the other five ingroup species, and represent a 60–62% sequence reduction compared to the photosynthetic *Ipomoea nil*, also belonging to the morning glory family.

In terms of gene composition, plastomes in this group contain between 32 and 59 protein coding genes, and between 21 and 25 tRNA genes. *Cuscuta costaricensis* is the largest plastome with 59 protein coding and 25 tRNA genes, and shows a 21% composition reduction compared to *Ipomoea nil* (whose plastome contains 78 protein coding and 30 tRNA genes) (Park et al. 2018). The most diminished genome, *Cuscuta boldinghii*, contains 32 protein coding and 25 tRNA genes, a reduction of 46% (Table 1). These results confirm the expectation of substantial variation in terms of genome size and composition within this group of very closely related plants.

Similar to other sequenced *Cuscuta* species (Funk et al. 2007; McNeal et al. 2007b), the whole *ndh* family of genes has been lost from all *Ceratophorae* plastomes along with the *rpo* genes. The maturase-encoding *matK*, photosystem genes *psaD* and *psaI*, ribosomal protein coding *rpl23*, *rpl32* and *rps16* genes, and *ycf15* of unknown function are also missing from the eight plastomes sampled (Fig. 2). In addition to these reductions, a subclade containing *Cuscuta*

Fig. 2 Heat map showing plastome sequence content in the eight species assembled from *Cuscuta* section *Ceratophorae*. Genes represented in black are present as open reading frames and considered functional, those in gray are present as pseudogenes, and those in white are absent. For legibility, the four *rpo* genes have been excluded from this figure because they are all absent from these species. The phylogenetic relationships are based on the whole-plastome phylogenetic analysis (compare with tree in Fig. 3)

strobilacea, *C. boldinghii* and *C. erosa* have also lost most of the cytochrome *b6/f* (*pet*) and photosystem I and II (*psa* and *psb*) genes, pivotal to the functioning of the photosynthetic apparatus (Figs. 2 and 4). In contrast, all ATP synthase (*atp*) genes (except for *atpF* in *C. boldinghii*) and most ribosomal protein coding genes have been retained even in the smallest *Ceratophorae* plastomes.

Table 2 shows presence and absence data for tRNA genes in the genomes assembled. *trnG*-UCC, *trnI*-GAU, *trnK*-UUU, *trnR*-ACG and *trnV*-UAC have been lost completely in sect. *Ceratophorae*, as is the case for *C. campestris* and *C. obtusiflora* in sect. *Cleistogrammica* (24 tRNA genes). However, sect. *Ceratophorae* maintains *trnA*-UGC in seven of its eight species (all except for *C. bonafortunae*), a gene absent from both known sect. *Cleistogrammica* plastomes. *Cuscuta mexicana* is also missing *trnR*-UCU and *trnS*-GCU, while *C. carnosa* has lost *trnC*-GCA, *trnD*-GUC, *trnE*-UUC, and *trnY*-GUA (Table 2).

Correlation with the Southern hybridization survey results

As mentioned before, Braukmann et al. (2013) conducted a comprehensive Southern hybridization survey of protein coding genes, encompassing nearly 60% of the known species diversity in *Cuscuta* (Braukmann et al. 2013) including five sect. *Ceratophorae* species. Table 3 compares the results of that study with the whole-plastome analysis conducted here. In general, there is good correlation between the presence and absence information generated by Braukmann et al. and the next-generation sequencing (NGS) approach used here, with 239 of 275 results in agreement. Pairwise Chi squared analyses comparing the two datasets found the differences between them to be non-significant for each of the five comparable species with *p* values ranging from 0.21 to 0.62 (Table 3).

Pseudogenes

Ten pseudogenization events are recorded in these plastomes (Figs. 2 and 4), based on observations of premature stop codons present early on in the ORFs, high sequence divergence, or missing start codons. *atpF* is pseudogenized in *Cuscuta boldinghii* due to multiple stop codons beginning approximately two-thirds of the way through the

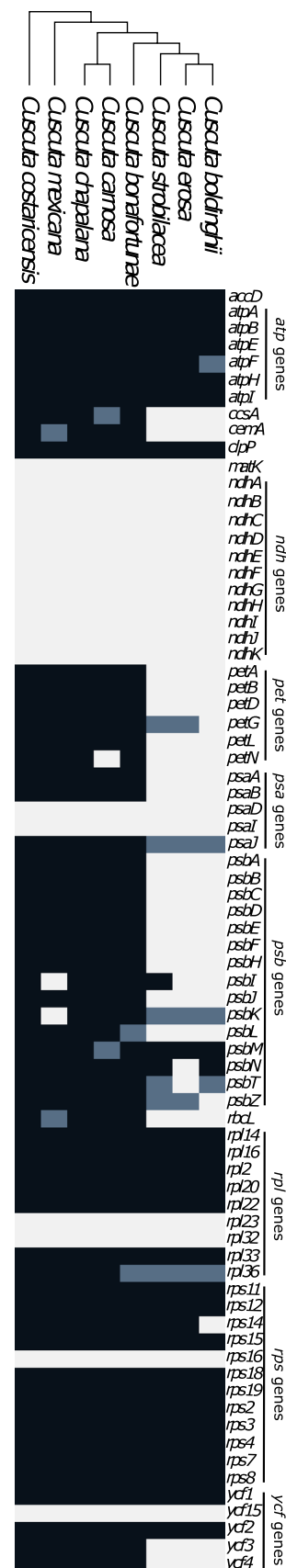


Table 2 tRNA gene presence (+) and absence (–) in the eight species assembled from *Cuscuta* section *Ceratophorae*

	<i>C. costaricensis</i>	<i>C. mexicana</i>	<i>C. carnososa</i>	<i>C. chapalana</i>	<i>C. bonafortunae</i>	<i>C. strobilacea</i>	<i>C. erosa</i>	<i>C. boldinghii</i>
<i>trnA-UGC</i>	+	+	+	+	–	+	+	+
<i>trnC-GCA</i>	+	+	–	+	+	+	+	+
<i>trnD-GUC</i>	+	+	–	+	+	+	+	+
<i>trnE-UUC</i>	+	+	–	+	+	+	+	+
<i>trnF-GAA</i>	+	+	+	+	+	+	+	+
<i>trnM-CAU</i>	+	+	+	+	+	+	+	+
<i>trnG-GCC</i>	+	+	+	+	+	+	+	+
<i>trnG-UCC</i>	–	–	–	–	–	–	–	–
<i>trnH-GUG</i>	+	+	+	+	+	+	+	+
<i>trnI-CAU</i>	+	+	+	+	+	+	+	+
<i>trnI-GAU</i>	–	–	–	–	–	–	–	–
<i>trnK-UUU</i>	–	–	–	–	–	–	–	–
<i>trnL-CAA</i>	+	+	+	+	+	+	+	+
<i>trnL-UAA</i>	+	+	+	+	+	+	+	+
<i>trnL-UAG</i>	+	+	+	+	+	+	+	+
<i>trnM-CAU</i>	+	+	+	+	+	+	+	+
<i>trnN-GUU</i>	+	+	+	+	+	+	+	+
<i>trnP-UGG</i>	+	+	+	+	+	+	+	+
<i>trnQ-UUG</i>	+	+	+	+	+	+	+	+
<i>trnR-ACG</i>	–	–	–	–	–	–	–	–
<i>trnR-UCU</i>	+	–	+	+	+	+	+	+
<i>trnS-GCU</i>	+	–	+	+	+	+	+	+
<i>trnS-GGA</i>	+	+	+	+	+	+	+	+
<i>trnS-UGA</i>	+	+	+	+	+	+	+	+
<i>trnT-GGU</i>	+	+	+	+	+	+	+	+
<i>trnT-UGU</i>	+	+	+	+	+	+	+	+
<i>trnV-GAC</i>	+	+	+	+	+	+	+	+
<i>trnV-UAC</i>	–	–	–	–	–	–	–	–
<i>trnW-CCA</i>	+	+	+	+	+	+	+	+
<i>trnY-GUA</i>	+	+	–	+	+	+	+	+

gene, caused by a frameshift mutation. The same is true of *ccsA* in *C. carnososa* (approximately in the first quarter of the sequence). Single stop codons, likely due to single substitutions, have caused pseudogenization events in *cemA* and *rbcL* in *C. mexicana*, as well in *rpl36* along the branch leading to *C. bonafortunae*, *C. strobilacea*, *C. boldinghii* and *C. erosa*. High sequence divergence has caused the absence of recognizable ORFs at the DNA- as well as the protein-level for *petG*, *psaJ* and *psbT* in the three most reduced species, with *petG* and *psbT* eventually being completely lost in *C. boldinghii* and *C. erosa*, respectively. Finally, missing start codons have led to the losses of ORFs for *psbL* in *C. bonafortunae*, *psbM* in *C. carnososa* and *psbZ* in *C. strobilacea* and *C. erosa*, followed by its total removal from *C. boldinghii*.

Structural changes

Two sequence inversions have been found, as marked in Fig. 1: one between *cemA* and *trnC-GCA* (~ 23,100 bp) in *Cuscuta bonafortunae*, and the other between *rpl2* and *rps8* (~ 3900 bp) in *C. boldinghii*. Aside from these two cases, there are no other major plastome rearrangements or structural mutations in section *Ceratophorae* (as shown in supplementary Figure S1).

The basic four-part structure (IR-SSC-IR-LSC) of the plastome molecule is maintained in the species assembled, with no major sequence ebb-and-flow between these components. The only minor change noted is the movement of tRNA-Ile (CAU) from the IR to the LSC in all *Ceratophorae* plastomes other than *C. costaricensis*, which retains this tRNA gene in the IR like the other *Cuscuta* genomes

Table 3 Correlations between the whole plastome data generated in this study and the results of the Southern hybridization survey conducted for the same species complex by Braukmann et al. 2013

	Present	Pseu- dogenized	Absent	
<i>Cuscuta erosa</i>				
Positive signal	13	0	3	p=0.28
Reduced or no signal	4	0	35	
<i>Cuscuta strobilacea</i>				
Positive signal	13	0	1	p=0.62
Reduced or no signal	4	1	36	
<i>Cuscuta carnosa</i>				
Positive signal	23	0	0	p=0.44
Reduced or no signal	7	1	24	
<i>Cuscuta mexicana</i>				
Positive signal	23	1	0	p=0.27
Reduced or no signal	7	0	24	
<i>Cuscuta costaricensis</i>				
Positive signal	24	0	0	p=0.21
Reduced or no signal	7	0	24	

Chi squared analyses (with a degree of freedom of 1) comparing the datasets species-by-species found the differences to be stochastic and not significant (p values indicated)

published thus far and like photosynthetic members of the *Convolvulaceae*.

Phylogenetic relationships and selection on genes

Branch-and-bound searches using the whole-plastome sequence data produced the same single best topology regardless of which optimality criterion (ML or parsimony) was employed. The phylogeny created using whole-plastome alignments under maximum likelihood is shown in Fig. 3 and has been used throughout this paper and in all figures.

Figure 5 shows the branch lengths of trees generated for the protein coding genes *accD*, *ycf1* and *ycf2* for sites that are assumed to accrue more nonsynonymous substitutions (i.e., the first two codon positions, dN in Fig. 5) vs. those that are expected to accrue more synonymous substitutions (the wobble position, dS). For each of the three genes, the branch lengths, and thus the amount of evolutionary change, accumulated in the dS tree is greater than that in the dN tree indicating synonymous changes are more prevalent than nonsynonymous ones in *Ceratophorae* plastomes. This analysis was also conducted on four additional genes (*atpB*, *atpH*, *clpP*, and *rpl20*) and results, showing the same trends, are depicted in the supplementary Figure S2.

Table 4 shows the number of codons under positive and purifying selection (with greater than/equal to 90% probability) for each of the seven genes mentioned previously based on alignments constrained by the inferred phylogeny

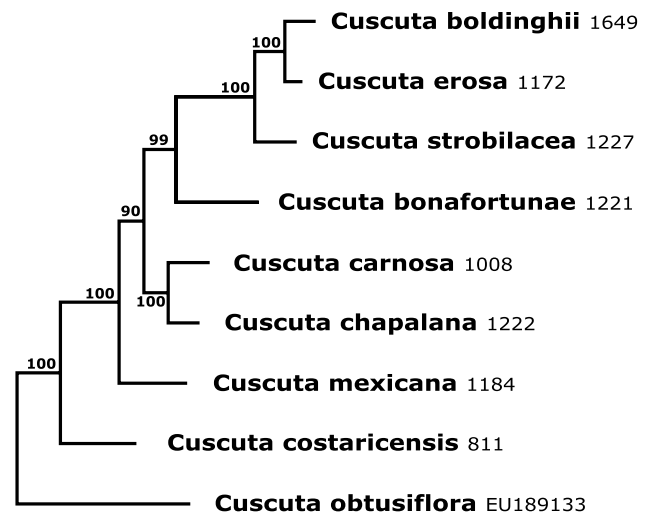


Fig. 3 Maximum likelihood tree obtained through a branch and bound search under the GTR+I+G model generated using 27,587 bp aligned from the whole plastomes dataset assembled in this study. All genes common to the eight *Cuscuta* section *Ceratophorae* species (at least 150 bp in length) were used to populate the dataset. ML bootstrap support values have been provided at nodes. Numbers following species names correspond to internal DNA accessions

in Fig. 3. The number of codons under purifying selection is greater than those under positive selection for every gene. This difference is most stark for *atpB* and *atpH* for which 25.1% and 11.1% of codons are under purifying selection respectively.

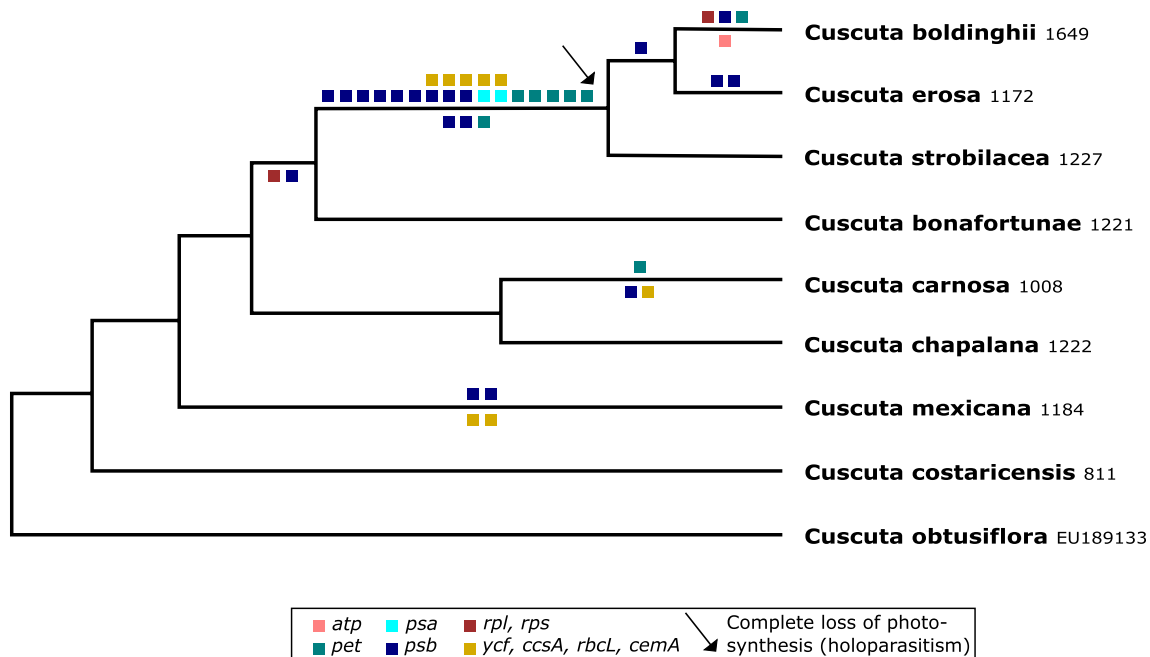
Discussion

Plastid genomes in *Cuscuta* sect. *Ceratophorae*

The results of the whole-plastome assembly research reported here document in-depth the fine-scale differences in plastome size and gene composition within *Cuscuta* sect. *Ceratophorae*, a complex of closely related species previously predicted to harbour a substantial heterogeneity in their plastid genomes (Braukmann et al. 2013). Furthermore, we distinguish between two categories of plastomes in this lineage. In one category we see plastomes that are fully comparable in size and gene content to those of other species in the subgenus *Grammica*; these plastomes are reduced when compared to autotrophs but still contain a large number of photosynthetic genes/ORFs, compatible with their ‘cryptic’ photosynthetic status. The other category contains plastomes with a further strong reduction in size and gene content, including the loss of most genes involved in photosynthesis; these plants are likely not capable of even residual (cryptic) amounts of photosynthesis and are the first truly holoparasitic plastomes described for *Cuscuta* (Table 1, Fig. 1).

Table 4 Analysis of selection on selected plastome genes in *Cuscuta* sect. *Ceratophorae* conducted using FUBAR (a fast, unconstrained Bayesian approximation for inferring selection; Murrell et al. 2013) on HyPhy (Pond et al. 2005)

Gene	Total sites	Codons	Distinct sites	Codons under positive selection ($p > 0.9$) (%)	Codons under purifying selection ($p > 0.9$) (%)
<i>accD</i>	1881	627	456	4 (0.64)	12 (1.91)
<i>atpB</i>	1482	494	170	0 (0.00)	124 (25.10)
<i>atpH</i>	243	81	33	0 (0.00)	9 (11.11)
<i>clpP</i>	594	198	61	0 (0.00)	6 (3.03)
<i>rpl20</i>	384	128	88	4 (3.13)	5 (3.91)
<i>ycf1</i>	5445	1815	1104	1 (0.06)	122 (6.72)
<i>ycf2</i>	5856	1952	486	16 (0.82)	31 (1.59)

**Fig. 4** Modified plastome tree from Fig. 3 mapping functional and actual protein-coding gene loss events in *Cuscuta* section *Ceratophorae*. Genes functionally lost (i.e., pseudogenes) are shown below

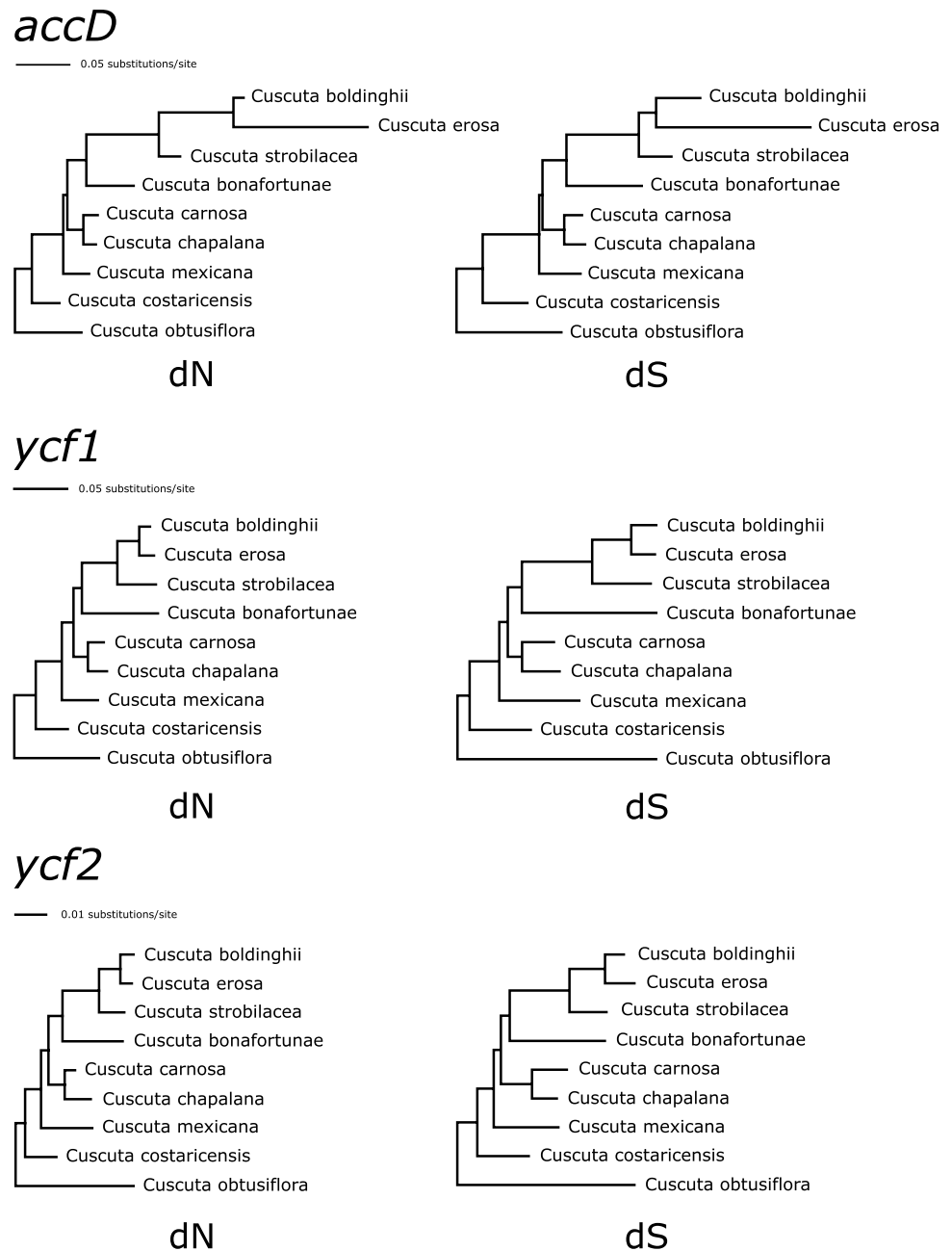
Cuscuta costaricensis has been shown to be sister to the rest of sect. *Ceratophorae* both in previous research (Costea et al. 2011) and currently (Fig. 3). The plastid genome of this species seems to be of the ancestral type for subgenus *Grammica*. It is nearly identical in terms of sequence length and protein-coding gene content to the previously published *Cuscuta obtusiflora* and *C. campestris* genomes (Funk et al. 2007; McNeal et al. 2007b) from the relatively closely related section *Cleistogrammica* (Stefanovic et al. 2007). It also retains *trnI*-CAU in the inverted repeat (IR) region, a structural similarity with *C. obtusiflora* and *C. campestris* that none of the other *Ceratophorae* species share. *Cuscuta chapalana*, likewise, is identical to *C. obtusiflora* when it comes to protein-coding gene composition and has only lost four tRNA genes (Tables 1, 2, Fig. 2). The six other species have accumulated a number of losses

branches and physical gene losses are shown above. Numbers following species names correspond to internal DNA accessions

and/or pseudogenization events in protein-coding genes, as shown in Fig. 4, although most of these changes occur on the branch leading to the clade containing *C. boldinghii*, *C. erosa*, and *C. strobilacea*. It is on this branch that a complete loss of photosynthesis appears to have evolved in *Cuscuta*.

As in four previously sequenced species of *Cuscuta* (Funk et al. 2007; McNeal et al. 2007b), and as suggested by the Southern hybridization survey across the genus (Braukmann et al. 2013), plastomes in sect. *Ceratophorae* do not contain any NADH dehydrogenase (*ndh*) or plastid-encoded polymerase (PEP, *rpo*) genes. The loss of *ndh* genes is expected in heterotrophic plants because they are not essential in normal, non-stress environments (Krause 2011). The products of these genes have been implicated in photo-oxidative stress response through the regulation of electron flow (Pelletier et al. 2016) but have even been lost from autotrophic

Fig. 5 dN versus dS branch lengths for three genes retained in all *Cuscuta* section *Ceratophorae* plastomes generated using maximum likelihood and the fixed topology presented in Fig. 3. dN trees were created using the bases occupying the first and second codon positions while the dS trees were created using just the third codon (wobble) positions



plastomes, as in the case of some photosynthetic orchids (Chang et al. 2006; Wu et al. 2010). In dodders, the absence of *ndh* genes is further expected because of their limited gas exchange and consequent high internal levels of carbon dioxide, which mitigates photo-oxidative stress (Braukmann et al. 2013; Hibberd et al. 1998). Similarly, the absence of *rpo* genes from heterotrophic plastomes has been noted regularly in previous studies (Graham et al. 2017; Wicke et al. 2013) and is considered to be an early event in the transition away from a photosynthetic lifestyle. While the loss of PEP may seem unexpected at first, because plastomes without *rpo* genes are still often intact and functioning, there is evidence

that the responsibility for the expression of plastid genes is subsumed by the nucleus and nuclear-encoded polymerase (NEP) in these cases (Krause et al. 2003).

Mirroring results found previously in *Cuscuta* sect. *Cleistogrammica* (Funk et al. 2007; McNeal et al. 2007b), plastomes in sect. *Ceratophorae* have also all lost the intron maturase gene *matK* (McNeal et al. 2009), which is usually responsible for splicing group IIA introns in the plastid genome. Correspondingly, seven of the eight group IIA introns are also missing in these species, all except for the one in *clpP* (second intron) which is capable of splicing itself (Zoschke et al. 2010). Five group IIB introns are

retained in the five larger *Ceratophorae* plastomes (in *rpl16*, *petD*, *petB*, *clpP* (first intron), and *trnL-UAA*) while the ones in *petD* and *petB* are lost in the three reduced genomes along with the genes that contain them.

The clade containing *Cuscuta boldinghii*, *C. erosa*, and *C. strobilacea* has incurred substantial further losses in genes encoding key portions of the photosynthetic apparatus. Sequences responsible for photosystems I and II (*psa* and *psb*), as well as those coding for components of the cytochrome *b6/f* complex (*pet*) are mostly absent in these three plastomes, and those that remain largely do so as pseudogenes (as described above). The gene *rbcL*, which codes for the large subunit of RUBISCO, has been lost in these species (and pseudogenized in *C. mexicana*, Fig. 2). This is a particularly interesting finding, because *rbcL* has been known to often remain intact even in greatly reduced plastomes, such as those of *Aphyllon* (Orobanchaceae; Schneider et al., 2018) or various lineages of mycoheterotrophs (Graham et al. 2017) and is thought to have a supplementary function in lipid synthesis pathways (Schwender et al. 2004). In addition, two genes encoding membrane proteins involved in photosynthesis, *cemA* and *ccsA*, are also absent from this clade, as are the photosystem I assembly factors *ycf3* and *ycf4* (Fig. 2).

The only genes with a photosynthetic function that are still retained in this clade, otherwise containing species with highly reduced plastomes, are the ATP synthase (*atp*) family of genes of which all six still remain intact and presumably functional (except for *atpF* in *C. boldinghii*, Fig. 2). These genes are frequently found intact in heterotrophs, in particular in the Orobanchaceae and in several mycotrophic orchids (Barrett et al. 2018; Wicke et al. 2013) and hence presumably have a secondary non-photosynthetic function. They are known to have a role in protein transport across thylakoid membranes in autotrophic plants (Kohzuma et al. 2012), but it is not known if they maintain this particular function in heterotrophs (Graham et al. 2017).

As shown in Table 2, most tRNA genes that seem to be ancestral for subg. *Grammica* are preserved in sect. *Ceratophorae*, with five of the eight plastomes in the species complex containing the same 25 genes. The tRNA loss events in the remaining three species are all idiosyncratic and display no discernable phylogenetic pattern. There is no relationship between the number of protein-coding and tRNA sequences lost, with the three members of the clade containing more diminished plastomes possessing the full complement of these genes.

The correlation analysis conducted between the whole-plastome sequences from this research and the Southern hybridization survey conducted by Braukmann et al. (2013) concluded that there was a high degree of agreement in the results of the two studies (Table 3). A Chi squared goodness-of-fit test (with one degree of freedom) was unable to reject

the null hypothesis that there are no significant differences between the two datasets for any of the five species. We can thus conclude that, for these species at least, the findings of the Southern hybridization probing in Braukmann et al. (2013) is accurate and can be used for analyses based on gene presence-and-absence data.

Selection on plastome genes

In order to further assess the state of plastomes in *Cuscuta* sect. *Ceratophorae*, we examined the type of selection acting upon plastid genes that are retained by all eight species. Figure 5 approximates the amount of nonsynonymous and synonymous substitutions (as represented by branch lengths derived from nonsynonymous and synonymous changes mapped on the constrained organismal topology) in three core non-bioenergetic genes (*accD*, *ycf1*, *ycf2*) known to retain vital functions even in strongly reduced plastomes (Bellot and Renner 2015; Graham et al. 2017; Wicke et al. 2013).

Across the board, branch lengths generated by nonsynonymous substitutions (i.e., on the trees derived from the first and second codon positions) were shorter than branch lengths generated by synonymous substitutions (on the trees derived from the third codon position only) indicating that these genes are under purifying selection because they accumulate more synonymous changes than nonsynonymous ones. This trend is particularly strong for the *ycf1* and *ycf2* genes (Fig. 5). These results suggest that these core essential genes seem not to be in danger of impending pseudogenization/loss in any of these species and will probably be retained as functional in sect. *Ceratophorae* in the future. The results of the codon-by-codon analysis of selection conducted on FUBAR (Table 4) further substantiate these conclusions, with all three genes having more codons under purifying selection than under positive selection. For *ycf1* in particular, 122 codons (6.7% of the total number) were found to be under purifying selection compared to just 1 codon (<0.1%) under positive selection. Twelve codons (1.9%) were found to be under purifying selection in *accD* compared to four (0.6%) under positive selection, and 31 codons (1.6%) were found to be under purifying selection in *ycf2* compared to 16 (0.8%) under positive selection.

These analyses were also conducted for four more genes, representative of additional functional gene groups retained by all species (supplementary Figure S2). For *atpB*, *atpH*, and *clpP*, the synonymous branch lengths are much longer than the nonsynonymous ones indicating that they are also under strong purifying selection. There is no clear trend with regards to the branch lengths for *rpl20* suggesting that this gene may be under more neutral selection. Mirroring these results, 124 (25.1%) and 9 (11.1%) codons in *atpB* and *atpH* respectively were found to be under purifying selection

compared to 0 codons under positive selection for both genes (Table 4). The gene *clpP* also has 0 codons under positive selection with 6 (3.0%) under purifying selection. Echoing the conclusions drawn from Figure S2, Table 4 shows that *rpl20* seems to be under more neutral selection than the other genes observed, with 4 codons (3.1%) under positive selection and 5 codons (3.9%) under purifying selection. The gene *clpP*, which codes for a proteolytic subunit of a complex responsible for the breakdown of abnormally folded proteins, is also considered an important non-bioenergetic gene. As discussed earlier, though, the role of *atp* genes in heterotrophic plants is not well understood, and the fact that they appear to be under strong purifying selection in sect. *Ceratophorae* adds to the growing body of evidence indicting this gene family is generally retained even in advanced cases of heterotrophy, strongly supporting the possibility of their products having an additional vital function in the plant cell.

The whole-plastome phylogeny for *Cuscuta* sect. *Ceratophorae*

The maximum likelihood phylogeny presented in Fig. 3 is entirely compatible with the plastid trees for section *Ceratophorae* generated in the past. The previous plastid tree, based solely on *trnL-F*, a short non-coding region (~620 bp aligned sequences), had multiple individuals sampled for a number of species. However, it suffered from a lack of support and resolution, especially along the backbone of the phylogeny, and the absence of plastid sequence data from two species. All of these issues are rectified here: we present a comprehensive sampling, with full resolution, and very strong support across the board (Fig. 3). On the other hand, our plastome-derived topology differs from previous trees produced using nuclear ribosomal sequences (ITS and partial 26S rDNA) in several aspects: the relative placement of *Cuscuta mexicana* and *C. bonafortunae* as well as the relationship between *C. chapalana* and *C. carnosa* (Costea et al. 2011). All of these discrepancies received only weak to moderate support with the ribosomal DNA data (Costea et al. 2011). Therefore, because the research described here focuses on plastid evolution, and has the benefit of more comprehensive taxon sampling and a much larger dataset, the phylogeny reconstructed here can be considered the best current plastome tree for sect. *Ceratophorae*.

Patterns of plastome reduction

The fine-scale system of parasitic plants sampled in the research reported here allows for comparisons of sequence composition between very closely related (and in some cases sister) species. This permits us to be able to phylogenetically triangulate the location of gene loss and pseudogenization

events in sect. *Ceratophorae* in a very precise fashion and present a step-by-step description of plastome evolution. This is particularly interesting because *Cuscuta* is one of two clades of parasitic angiosperms, along with the Orobanchaceae, that are known to span the continuum from hemiparasitic to holoparasitic (Westwood et al. 2010; their Fig. 1). However, *Cuscuta* (and specifically its subgenus *Grammica*) achieves this trophic diversity at a lower phylogenetic level (i.e., species level) compared to Orobanchaceae. This makes the *Cuscuta* model described here more tractable because there is expected to be substantially less interference from the confounding historical factors, usually difficult to avoid at the higher phylogenetic level.

Figure 4 maps the location of changes in protein-coding genes in this species complex and has been generated based on the assumption of irreversibility, i.e., the premise that plastomes can lose function (and corresponding genes), but once lost, the function cannot be regained. The broad sequence of reduction supports the models of plastome degradation summarized earlier in the Introduction (Barrett and Davis 2012; Graham et al. 2017) whereby the *ndh* genes are lost first (in all *Cuscuta*), followed by *rpo* genes (apparently across subg. *Grammica*). The five species forming the basal grade in sect. *Ceratophorae* that exhibit limited plastid genome reduction are missing *ndh* and *rpo* genes but not much more and seem to fit solidly in the category of hemiparasites (i.e., “cryptically” photosynthetic species). The plastomes in the more reduced clade, however, have gone much further on the scale of degradation and have retained only the housekeeping genes, genes with other non-bioenergetic functions, and *atp* genes.

Braukmann et al. (2013), when analyzing the results of their Southern hybridization survey, speculated that plastid genomes in sect. *Ceratophorae* may exemplify gradual evolution according to the Evolutionary Transition Series hypothesis where changes accumulate phylogenetically ‘progressively’ (Mayr 1982; Young et al. 1999). Based on the results of the whole-genome approach in this research, it seems that plastome change in this species complex is somewhere in between gradual and punctuated (whereby short bursts of evolutionary change are separated by long periods of relative stasis). There is indeed an intense burst of gene loss along the branch leading to the more reduced clade in sect. *Ceratophorae*, but three of the other species in the group have undergone idiosyncratic sequence losses in protein coding genes as well (Fig. 4) and therefore cannot be described as simply being ‘in stasis’.

Acknowledgements Special thanks to Eleazer Carranza, Mihai Costea, Ignacio Garcia Ruiz, Marc Johnson, and Tom Van Devender for tissue collections crucial to this project. We are also grateful to two anonymous reviewers for their valuable comments that have improved the quality of the manuscript. This work was supported by the Natural Sciences and Engineering Research Council of Canada (Grant No.

326439); the Canada Foundation for Innovation (Grant No. 12810); and Ontario Research Funds, all to S.S.

Author contributions AB and SS conceived the research conducted here. SS obtained plant tissue and extracted DNA. AB prepared samples for high-throughput sequencing and performed plastome assemblies and annotations. Analyses were conducted by both AB and SS. AB produced the first draft of the manuscript. Both authors have read and approved the final version of the manuscript.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Barkman TJ, McNeal JR, Lim SH, 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. <https://doi.org/10.1186/1471-2148-7-248>
- Barrett CF, Davis JI (2012) The plastid genome of the mycoheterotrophic *Corallorhiza striata* (Orchidaceae) is in the relatively early stages of degradation. *Am J Bot* 99:1513–1523. <https://doi.org/10.3732/ajb.1200256>
- Barrett CF, Wicke S, Sasse C (2018) Dense infraspecific sampling reveals rapid and independent trajectories of plastome degradation in a heterotrophic orchid complex. *New Phytol*. <https://doi.org/10.1111/nph.15072>
- Bellot S, Renner SS (2015) The plastomes of two species in the endoparasite genus *pilostyles* (Apodanthaceae) each retain just five or six possibly functional genes. *Genome Biol Evol* 8:189–201. <https://doi.org/10.1093/gbe/evv251>
- Braukmann T, Kuzmina M, Stefanovic S (2013) Plastid genome evolution across the genus *Cuscuta* (Convolvulaceae): two clades within subgenus *Grammica* exhibit extensive gene loss. *J Exp Bot* 64:977–989. <https://doi.org/10.1093/jxb/ers391>
- Braukmann TWA, Broe MB, Stefanovic S, Freudenstein JV (2017) On the brink: the highly reduced plastomes of nonphotosynthetic Ericaceae. *New Phytol* 216:254–266. <https://doi.org/10.1111/nph.14681>
- Chang CC et al (2006) The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae): comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. *Mol Biol Evol* 23:279–291. <https://doi.org/10.1093/molbev/msj029>
- Colwell AE (1994) Genome evolution in a non-photosynthetic plant, *Conopholis americana*. Washington University, Louis
- Costea M, Garcia Ruiz I, Stefanovic S (2011) Systematics of “horned” dodders: phylogenetic relationships, taxonomy, and two new species within the *Cuscuta chapalana* complex (Convolvulaceae). *Botany* 89:715–730
- Costea M, Garcia MA, Stefanovic S (2015) A phylogenetically based infrageneric classification of the parasitic plant genus *Cuscuta* (dodders, Convolvulaceae). *Syst Bot* 40:269–285
- Darling AE, Mau B, Perna NT (2010) progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* 5:e11147. <https://doi.org/10.1371/journal.pone.0011147>
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9:772. <https://doi.org/10.1038/nmeth.2109>
- Dawson JH, Musselman LJ, Wolswinkel P, Dorr I (1994) Biology and control of *Cuscuta*. *Rev Weed Sci* 6:265–317
- Downie SR, Palmer JD (1992) Restriction site mapping of the chloroplast DNA inverted repeat—a molecular phylogeny of the Asteridae. *Ann Mo Bot Gard* 79:266–283. <https://doi.org/10.2307/2399769>
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15
- Funk HT, Berg S, Krupinska K, Maier UG, Krause K (2007) Complete DNA sequences of the plastid genomes of two parasitic flowering plant species, *Cuscuta reflexa* and *Cuscuta gronovii*. *BMC Plant Biol* 7:45. <https://doi.org/10.1186/1471-2229-7-45>
- Garcia MA, Costea M, Kuzmina M, Stefanovic S (2014) Phylogeny, character evolution, and biogeography of *Cuscuta* (dodders; Convolvulaceae) inferred from coding plastid and nuclear sequences. *Am J Bot* 101:670–690. <https://doi.org/10.3732/ajb.1300449>
- Graham SW, Lam VK, Merckx VS (2017) Plastomes on the edge: the evolutionary breakdown of mycoheterotroph plastid genomes. *New Phytol* 214:48–55. <https://doi.org/10.1111/nph.14398>
- Greiner S, Lehwark P, Bock R (2019) OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res*. <https://doi.org/10.1093/nar/gkz238>
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704
- Heide-Jorgensen H (2008) Parasitic flowering plants. Brill Academic Publishers, Leiden
- Hibberd JM, Bungard RA, Press MC, Jeschke WD, Scholes JD, Quick WP (1998) Localization of photosynthetic metabolism in the parasitic angiosperm *Cuscuta reflexa*. *Planta* 205:506–513. <https://doi.org/10.1007/s004250050349>
- Joshi NA, Fass JN (2011) Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files [Software]. <https://github.com/najoshi/sickle>
- Kohzuma K, Dal Bosco C, Kanazawa A, Dhingra A, Nitschke W, Meurer J, Kramer DM (2012) Thioredoxin-insensitive plastid ATP synthase that performs moonlighting functions. *Proc Natl Acad Sci USA* 109:3293–3298. <https://doi.org/10.1073/pnas.1115728109>
- Krause K (2011) Piecing together the puzzle of parasitic plant plastome evolution. *Planta* 234:647–656. <https://doi.org/10.1007/s00425-011-1494-9>
- Krause K, Berg S, Krupinska K (2003) Plastid transcription in the holoparasitic plant genus *Cuscuta*: parallel loss of the *rrn16* PEP-promoter and of the *rpoA* and *rpoB* genes coding for the plastid-encoded RNA polymerase. *Planta* 216:815–823. <https://doi.org/10.1007/s00425-002-0933-z>
- Kuijt J (1969) The biology of parasitic flowering plants. University of California Press, Berkeley
- Lowe TM, Chan PP (2016) tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res* 44:W54–W57. <https://doi.org/10.1093/nar/gkw413>
- Mayr E (1982) Speciation and macroevolution. *Evolution* 36:1119–1132. <https://doi.org/10.2307/2408147>
- McNeal JR, Arumugunathan K, Kuehl JV, Boore JL, Depamphilis CW (2007a) Systematics and plastid genome evolution of the cryptically photosynthetic parasitic plant genus *Cuscuta* (Convolvulaceae). *BMC Biol* 5:55. <https://doi.org/10.1186/1471-7007-5-55>
- McNeal JR, Kuehl JV, Boore JL, de Pamphilis CW (2007b) Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. *BMC Plant Biol* 7:57. <https://doi.org/10.1186/1471-2229-7-57>
- McNeal JR, Kuehl JV, Boore JL, Leebens-Mack J, dePamphilis CW (2009) Parallel loss of plastid introns and their maturase in the genus *Cuscuta*. *PLoS ONE* 4:e5982. <https://doi.org/10.1371/journal.pone.0005982>

- Molina J et al (2014) Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Mol Biol Evol* 31:793–803. <https://doi.org/10.1093/molbev/msu051>
- Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, Scheffler K (2013) FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol* 30:1196–1205. <https://doi.org/10.1093/molbev/mst030>
- Naumann J et al (2016) Detecting and characterizing the highly divergent plastid genome of the nonphotosynthetic parasitic plant *Hydnora visseri* (Hydnoraceae). *Genome Biol Evol* 8:345–363. <https://doi.org/10.1093/gbe/evv256>
- Park I, Yang S, Kim WJ, Noh P, Lee HO, Moon BC (2018) The complete chloroplast genomes of six ipomoea species and indel marker development for the discrimination of authentic pharbitidis semen (seeds of *I. nil* or *I. purpurea*). *Front Plant Sci* 9:965. <https://doi.org/10.3389/fpls.2018.00965>
- Peltier G, Aro EM, Shikanai T (2016) NDH-1 and NDH-2 plastoquinone reductases in oxygenic photosynthesis. *Annu Rev Plant Biol* 67:55–80. <https://doi.org/10.1146/annurev-arplant-043014-114752>
- Pond SL, Frost SD, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679. <https://doi.org/10.1093/bioinformatics/bti079>
- Schwender J, Goffman F, Ohlrogge JB, Shachar-Hill Y (2004) Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature* 432:779–782. <https://doi.org/10.1038/nature03145>
- Stefanovic S, Kuzmina M, Costea M (2007) Delimitation of major lineages within *Cuscuta* subgenus *Grammica* (Convolvulaceae) using plastid and nuclear DNA sequences. *Am J Bot* 94:568–589. <https://doi.org/10.3732/ajb.94.4.568>
- van der Kooij TA, Krause K, Dorr I, Krupinska K (2000) Molecular, functional and ultrastructural characterisation of plastids from six species of the parasitic flowering plant genus *Cuscuta*. *Planta* 210:701–707. <https://doi.org/10.1007/s004250050670>
- Westwood JH, Yoder JI, Timko MP, dePamphilis CW (2010) The evolution of parasitism in plants. *Trends Plant Sci* 15:227–235. <https://doi.org/10.1016/j.tplants.2010.01.004>
- Wicke S, Schneeweiss GM, dePamphilis CW, Muller KF, Quandt D (2011) The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol* 76:273–297. <https://doi.org/10.1007/s11103-011-9762-4>
- Wicke S et al (2013) Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *Plant Cell* 25:3711–3725. <https://doi.org/10.1105/tpc.113.113373>
- Wilgenbusch JC, Swofford D (2003) Inferring evolutionary trees with PAUP*, Chapter 6. *Curr Protoc Bioinform*. <https://doi.org/10.1002/0471250953.bi0604s00>
- Wolfe KH, Morden CW, Ems SC, Palmer JD (1992) Rapid evolution of the plastid translational apparatus in a nonphotosynthetic plant—loss or accelerated sequence evolution of transfer-rna and ribosomal-protein genes. *J Mol Evol* 35:304–317. <https://doi.org/10.1007/Bf00161168>
- Wu FH et al (2010) Complete chloroplast genome of *Oncidium Gower Ramsey* and evaluation of molecular markers for identification and breeding in *Oncidiinae*. *BMC Plant Biol* 10:68. <https://doi.org/10.1186/1471-2229-10-68>
- Young ND, Steiner KE, dePamphilis CW (1999) The evolution of parasitism in Scrophulariaceae/Orobanchaceae: plastid gene sequences refute an evolutionary transition series. *Ann Mo Bot Gard* 86:876–893. <https://doi.org/10.2307/2666173>
- Yuncker TG (1932) The genus *Cuscuta*. *Mem Torrey Bot Club* 18:113–331
- Zoschke R, Nakamura M, Liere K, Sugiura M, Borner T, Schmitz-Linneweber C (2010) An organellar maturase associates with multiple group II introns. *Proc Natl Acad Sci USA* 107:3245–3250. <https://doi.org/10.1073/pnas.0909400107>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.