# Disproportional Plastome-Wide Increase of Substitution Rates and Relaxed Purifying Selection in Genes of Carnivorous Lentibulariaceae

Susann Wicke,*[†,1] Bastian Schäferhoff,[†,1] Claude W. dePamphilis,[2] and Kai F. Müller[1]

[1]Institute for Evolution and Biodiversity, University of Muenster, Muenster, Germany

[2]Department of Biology and Institute of Molecular Evolutionary Genetics, Pennsylvania State University

[†]These authors contributed equally to this work.

***Corresponding author:** E-mail: susann.wicke@uni-muenster.de.

**Associate editor:** Charles Delwiche

## Abstract

**Carnivorous Lentibulariaceae exhibit the most sophisticated implementation of the carnivorous syndrome in plants. Their unusual lifestyle coincides with distinct genomic peculiarities such as the smallest angiosperm nuclear genomes and extremely high nucleotide substitution rates across all genomic compartments. Here, we report the complete plastid genomes from each of the three genera *Pinguicula*, *Utricularia*, and *Genlisea*, and investigate plastome-wide changes in their molecular evolution as the carnivorous syndrome unfolds. We observe a size reduction by up to 9% mostly due to the independent loss of genes for the plastid NAD(P)H dehydrogenase and altered proportions of plastid repeat DNA, as well as a significant plastome-wide increase of substitution rates and microstructural changes. Protein-coding genes across all gene classes show a disproportional elevation of nonsynonymous substitutions, particularly in *Utricularia* and *Genlisea*. Significant relaxation of purifying selection relative to noncarnivores occurs in the plastid-encoded fraction of the photosynthesis ATP synthase complex, the photosystem I, and in several other photosynthesis and metabolic genes. Shifts in selective regimes also affect housekeeping genes including the plastid-encoded polymerase, for which evidence for relaxed purifying selection was found once during the transition to carnivory, and a second time during the diversification of the family. Lentibulariaceae significantly exhibit enhanced rates of nucleotide substitution in most of the 130 noncoding regions. Various factors may underlie the observed patterns of relaxation of purifying selection and substitution rate increases, such as reduced net photosynthesis rates, alternative paths of nutrient uptake (including organic carbon), and impaired DNA repair mechanisms.**

***Key words:*** **carnivorous plants, chloroplast genome, substitution rates, purifying selection, Lentibulariaceae, noncoding DNA.**

## Introduction

The carnivorous syndrome in plants comprises the ability to attract, retain, trap, kill, digest prey, and finally absorb and make use of the nutrients resulting from digestion (Juniper 1986). Carnivory evolved several times among angiosperms with at least seven independent origins in five different orders of flowering plants (Albert et al. 1992; Soltis et al. 2005; Ellison and Gotelli 2009; Pereira et al. 2012), with the order Lamiales being the hotspot of carnivorous plant diversity. Lamiales possess the greatest trap type diversity and a minimum of three independently evolved carnivorous lineages (Byblidaceae, *Philcoxia* from Plantaginaceae, Lentibulariaceae) and several subcarnivorous taxa (Plachno et al. 2006; Schäferhoff et al. 2010), which are capable to trap or kill prey such as insects but cannot digest and absorb the released nutrients.

With well more than 350 species, the cosmopolitan Lentibulariaceae are by far the most diverse carnivorous family, exhibiting extreme physiological adaptations and what has been described as the most complex leaf structures known in the entire plant kingdom (Juniper et al. 1989). *Pinguicula*, sister to the so-called bladderwort lineage that comprises *Utricularia* and *Genlisea*, has the simplest body plan, capturing prey with the help of (often) pale-green leaves borne in a rosette that function as flypaper traps. *Genlisea*, the corkscrew plant, possesses Y-shaped and helically twisted underground rhizophylls (eel traps) used to passively capture and digest protozoa (Barthlott et al. 1998). In contrast, the bladder traps of *Utricularia* work by means of low pressure, capturing small organisms from the surrounding water or soil, and have been characterized as having the most complex leaf modifications (Juniper et al. 1989). In contrast to *Pinguicula*, bladderworts and corkscrew plants lack a root system, and thus solely rely on the uptake of nutrients through prey and by means of foliar absorption (Adamec 1997, 2013).

Lentibulariaceae also exhibit remarkable traits at the molecular level. For instance, the smallest genomes reported from angiosperms belong to species of *Genlisea* and *Utricularia* (Greilhuber et al. 2006). DNA substitutional rates in several loci representing all three genomic compartments were found to be significantly higher in members of the







Fast Track Article

**Table 1.** Overview of the Physical Properties of Lentibulariaceae Plastid Genomes.

| | Size (bp) | %GC | No. of Unique Intact Genes (Protein-Coding/tRNA/rRNA) | Pseudogenes or Deleted Genes (*) |
|---|---|---|---|---|
| *Pinguicula* | 147,147 | 38,193 | 75/30/4 | *ndhC*, ndhD, ndhF*, ndhK* |
| *Utricularia* | 153,228 | 37,041 | 79/30/4 | — |
| *Genlisea* | 141,255 | 38,342 | 69/30/4 | *ndhA*, ndhC*, *ndhD*, ndhE*, ndhF*, ndhG*, ndhH*, ndhI, ndhJ*, ndhK** |

bladderwort lineage than in other angiosperms (Jobson and Albert 2002; Müller et al. 2004; Ibarra-Laclette et al. 2011). The finding that plastid genes of Lentibulariaceae evolve at elevated substitution rates was unexpected as plastid genomes, and particularly their coding regions, usually evolve in a remarkably conservative manner across land plants (Palmer 1985; Raubeson and Jansen 2005; Wicke et al. 2011), with the majority of the ~110–120 unique genes being indispensable for photosynthesis and a few other metabolic and genetic pathways. Relatively few autotrophic lineages demonstrate unusual modes of plastid genome and plastid gene evolution (Palmer et al. 1988; Cosner et al. 2004; Guisinger et al. 2010, 2011; Fajardo et al. 2012; Sloan et al. 2012), including large-scale reconfiguration of the typically compact quadripartite structure that consists of a large and small single-copy region (LSC, SSC) and two large inverted repeats (IRs).

The most dramatic changes in plastid gene content have been reported for plants with a heterotrophic mode of nutrition (Wolfe et al. 1992; Nickrent et al. 1997; Funk et al. 2007; McNeal, Kuehl, et al. 2007; Delannoy et al. 2011; Braukmann and Stefanović 2012). Like bladderworts, many (obligate) parasitic plants exhibit highly accelerated rates, which result (at least in part) from reduced purifying selection in photosynthesis and photosynthesis-related plastid genes (e.g., dePamphilis et al. 1997; Young and dePamphilis 2005; McNeal, Kuehl, et al. 2007). Although no case of C-heterotrophy in carnivorous plants has been documented, available data indicate regular carbon uptake via carnivory, particularly under challenging environmental conditions (Adamec 1997; Pavlovič et al. 2009; Adamec 2011). Previous studies (reviewed in Adamec 2011) suggest a trend toward high rates of molecular evolution in lineages where the normal uptake of inorganic nutrients via roots is accompanied or replaced by some other mode of nutrient acquisition. Naturally, this involves diversely modified biochemical processes and, thus, may dramatically affect the molecular evolution in a variety of genetic loci. It is therefore comparable to some extent with the situation in parasites.

To investigate whether the nutritional benefit of carnivory might impact the evolution of the photosynthetic machinery, we sequenced the plastid genomes of representatives of all three main lineages of Lentibulariaceae. Using a comparative approach and a series of sophisticated analyses of rate variation under maximum likelihood, we discriminate between the role of positive selection, relaxed purifying selection, and background mutation rates. Here, we present evidence for changes in the degree of positive or negative selection acting on the functionally diverse array of plastid genes and discuss it in the light of the degree of dependency on prey in the family. A fine-scale analysis of rate shifts in nonprotein-coding regions and across all plastid genes assesses the degree of selectional changes on plastid protein-coding regions. Above that, we examine the evolution of microstructural changes across the plastid chromosomes of the carnivores with special regard to the strength of purifying selection and functional losses. The results of this study promote our understanding of genomic footprints of carnivory and reveal genome-level convergences between carnivores and parasites.

## Results

### Structure of Plastid Genomes in Lentibulariaceae

We completely sequenced the plastid genomes (plastomes) of *Pinguicula ehlersiae*, *Utricularia macrorhiza*, and *Genlisea margaretae* as representatives of the three genera of the carnivorous Lentibulariaceae. In terms of structure, their plastomes resemble those of the vast majority of angiosperms. With a total length of 153,228 bp, *Utricularia* possesses the largest plastome of the three sequenced species (table 1). The plastid genomes of *Pinguicula* and *Genlisea* are 147,147 bp and 141,255 bp in lengths, respectively, and colinear with that of *Utricularia* (fig. 1). None of the three species show structural reconfigurations such as inversions or gene relocations into other plastid segments. Relative to *Utricularia* and other lamiids, disruptions of gene synteny occur, however, in the LSC and SSC of *Genlisea*, and in the SSC of *Pinguicula* due to the deletion of genes encoding subunits of the NAD(P)H dehydrogenase complex. Considering phylogenetic relationships, *ndh* gene loss of *Genlisea* and *Pinguicula* must have occurred two times independently within Lentibulariaceae. In *Genlisea*, the genes *ndhC, D, F, G, H, J*, and *K* were lost from the plastome; the genes *ndhA, E*, and *I* reside as truncated pseudogenes. *NdhB* is retained with an intact reading frame in both *Genlisea* and *Pinguicula*. The latter also retains reading frames for *ndhA, D, E, G, H, I, J*, and *K* although these show notably long indels and/or frameshifts, implying that some of these may be pseudogenes. We could not detect *ndhC* and *F* in *Pinguicula*.

A genome-wide alignment between the Lentibulariaceae plastomes and that of the closely related sesame (*Sesamum indicum*, Pedaliaceae) reveals several mutational hotspots (fig. 1). This includes the entire SSC as well as the region around the *ndh* genes in the LSC (fig. 1). Further hotspots of sequence variation accumulate near tRNAs and short (<200 bp) protein-coding genes, which becomes especially evident in the *rps16-atpA* and the *rpoB-psbD* fragment.

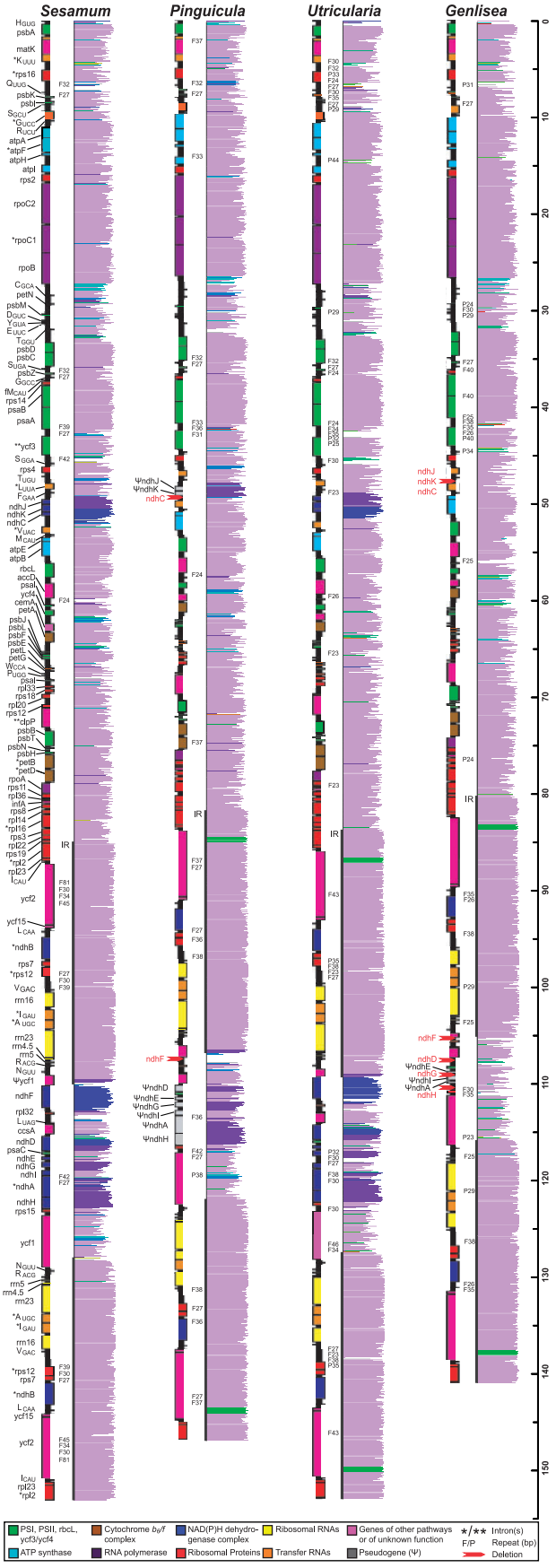


**Fig. 1.** Physical maps and sequence similarity between Lentibulariaceae and closely related noncarnivorous species. The organization of the plastid genome is depicted as a physical map for each species, showing the relative strandedness of genes, which are colored according to their functional classes. In some instances, tRNA genes appear as black bars because of their minimal length. Gene names are provided on the left, and asterisks indicate the number of introns in a specified gene. Pseudogenes are indicated by Ψ, whereas a red arrow marks the deletion of a specified gene region. Based on a plastome alignment of the four species with progressiveMauve, similarity plots to the right of each plastome map indicate the hotspots of divergence. Regions (represented by thin bars) colored mauve represent segments conserved among all genomes, while differently colored ones mark those present only in a subset of taxa. The height of a bar shows the degree of conservation, and gaps in a plot indicate that this particular region could not be aligned to any of the other genomes, suggesting this fragment to be species specific. The approximate localizations of larger repeats (>12 nt) are illustrated below the similarity plot, with a two-digit number indicating the length and orientation of the repeat (F, forward; P, palindromic; IR, large IR).

Relative to the sesame plastome, the number of insertions and deletions is slightly lower in *Pinguicula* (cRIN: 206, cRDN: 213) than in the bladderwort lineage (*Utricularia*–cRIN: 211, cRDN: 238; *Genlisea*–cRIN: 226; cRDN: 257; fig. 2*a*). Although the number of simple sequence repeats (SSRs) is not evidently different between carnivores and noncarnivores, we observed a slight elevation of AT-rich di- and tetranucleotide motifs repeated more than five or three times, respectively, in the bladderwort lineage (fig. 2*b*).

Similar to most flowering plant plastomes, Lentibulariaceae have few repeats of more than 12 nt in length (fig. 2*c*). Lentibulariaceae exhibit a slight increase particularly in reverse-complement repeats (palindromic) compared with sesame, and we detected a large number of direct and palindromic repeats especially in *Utricularia*. Unlike *Sesamum*, all Lentibulariaceae also possess a small number of repeats larger than 50 bp but none larger than 100 bp (except for the large IR). The distributions of larger repeats in the plastomes are very similar between *Sesamum* and *Pinguicula* (fig. 1), occurring not only in nontranscribed spacers (e.g., *trnQ-rps16*, *psaA-ycf3*, 3′*rps12-trnV*) but also in the *ycf2* gene region. Several more repeats accumulate in these regions of the *Utricularia* and *Genlisea* plastomes, which also harbor repeat hotspots in the SSC, in the *rpoB-psbD* fragment, and several more dispersed repeats in the IRs (fig. 1).

## Synonymous and Nonsynonymous Substitution Rates

Across all plastid protein-coding genes, we observe a general elevation of nucleotide substitution rates in Lentibulariaceae compared with noncarnivorous lamiids (fig. 3*a*). Rates particularly increase in the bladderwort lineage ($P_{MWU} < 0.001$ for dS and dN; table 2). *Pinguicula*, in contrast, shows significant acceleration only in its synonymous rates ($P_{MWU} = 0.039$). Rate differences are already detectable along the branches leading to Lentibulariaceae and the *Utricularia*/*Genlisea* subclade (fig. 3*b*), implying shared ancient rate accelerations.

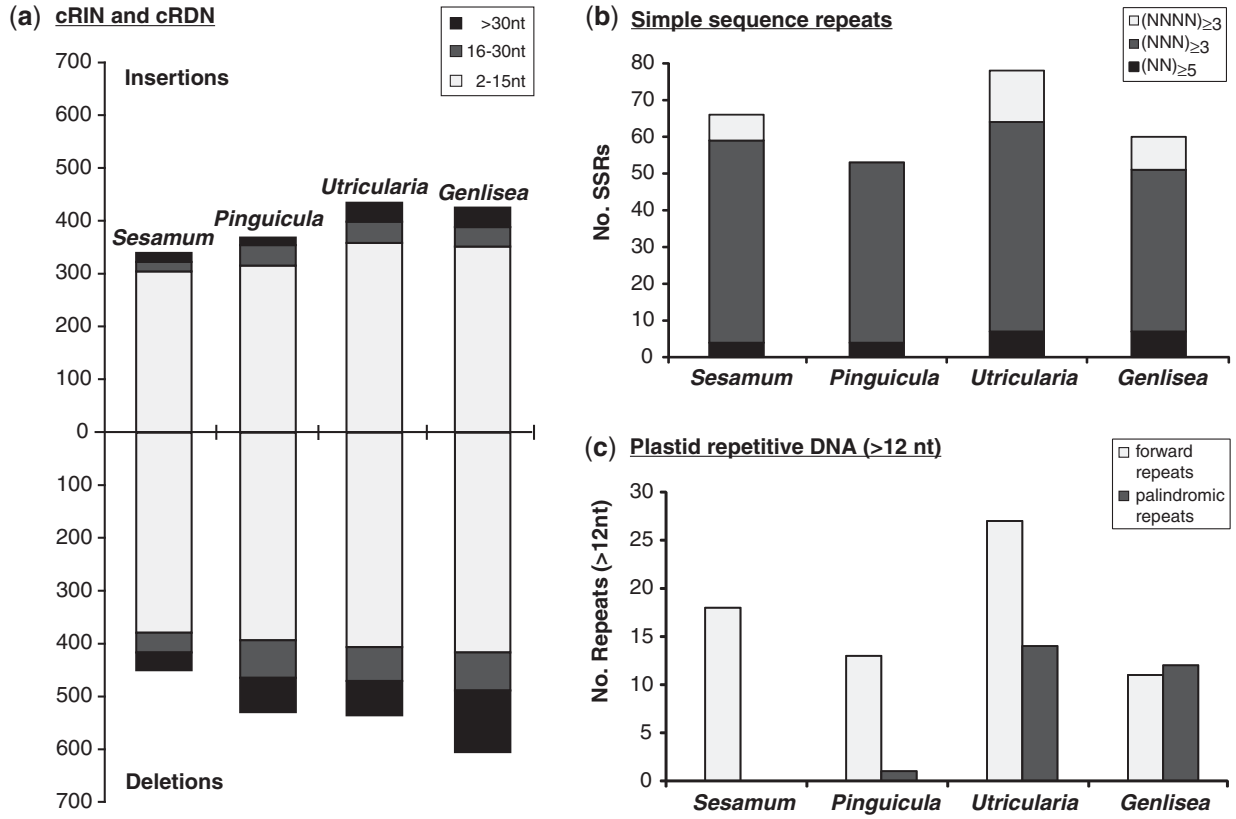


FIG. 2. Repeat DNAs in Lentibulariaceae plastid genomes. Number and ratio of (*a*) 1 nt-corrected relative insertions (cRIN) and deletions (cRDN) of three different length categories, (*b*) di-, tri-, and tetra-SSRs of at least five, three, or three repetitions, respectively, and (*c*) forward repeats and palindromic repeats of more than 12 bp length in *Sesamum* and Lentibulariaceae.

Relative to noncarnivores, 23 protein-coding genes in the *Pinguicula* plastome show notable rate deviations, with 20 genes being elevated in nonsynonymous rates (fig. 3*c*; supplementary table S1, Supplementary Material online). Of these, 13 genes are likely to have experienced rate elevation after the split from *Genlisea*/*Utricularia*, rather than in the Lentibulariaceae ancestor. In *Utricularia* and *Genlisea*, 42 and 52 genes, respectively, evolve significantly faster with, again, more genes having elevated dN than dS (fig. 3*c*; supplementary tables S1 and S2, Supplementary Material online).

Rate acceleration apparently affects all genes of the Lentibulariaceae plastid genomes similarly. Higher rates (dN and dS; fig. 3) can be observed in genes for both the photosynthesis light and dark reactions (photosystems, electron transport, RuBisCO) as well as photosynthesis-associated processes ($CO_2$ uptake, heme attachment), housekeeping genes (ribosomal proteins, polymerase), and genes for light-independent processes (lipid synthesis).

## Changes of Selection in Plastid Genes

Significantly relaxed purifying selection in Lentibulariaceae relative to noncarnivorous lamiids is inferred for the plastid-encoded subunits of the ATP synthase complex (*atp* genes), photosystem I (*psa* genes), and the plastid-encoded polymerase (PEP, *rpo* genes). However, the precise modes of selectional change differ between these complexes (fig. 3*c* and table 3). Although $\omega$ in PEP genes appears to be significantly different in the noncarnivorous clade, the carnivorous ancestor, and the carnivore clade (i.e., three selectional regimes, likelihood ratio test (LRT) from the best model: $P < 0.001$), relaxation of purifying selection apparently has affected the ATP synthase complex predominantly during the transition to the carnivorous lifestyle (LRT, best model: $P = 0.003$). Purifying selection in genes of the photosystem I is relaxed only in the Lentibulariaceae crown group (LRT, best model: $P < 0.001$).

Not all genes with elevated rates of nonsynonymous changes show relaxation of purifying selection. Genewise tests reveal that not all genes of a functional complex display similar patterns of purifying selection. Based on Akaike information criterion (AIC) model selection, a single global $\omega$ may be assumed for four of the six plastid-encoded *atp* genes and three of five *psa* genes (supplementary table S2, Supplementary Material online). Purifying selection is relaxed in *atpB* (LRT, best model $P = 0.003$), *atpI* (LRT, best model $P = 0.029$), *psaB* (LRT, best model $P = 0.006$), and *psaC* (LRT, best model $P < 0.001$) of Lentibulariaceae compared with noncarnivores. *RpoB* as well as *rpoC1* also evolve under relaxed purifying selection in *Genlisea* compared with noncarnivores (*rpoB*–LRT, best model $P = 0.004$; *rpoC1*–LRT, best model $P = 0.003$). *RpoA* is the only subunit of the polymerase complex to exhibit a significant increase of $\omega$ in the branch leading to Lentibulariaceae (LRT, best model $P = 0.002$; supplementary table S2, Supplementary Material online).

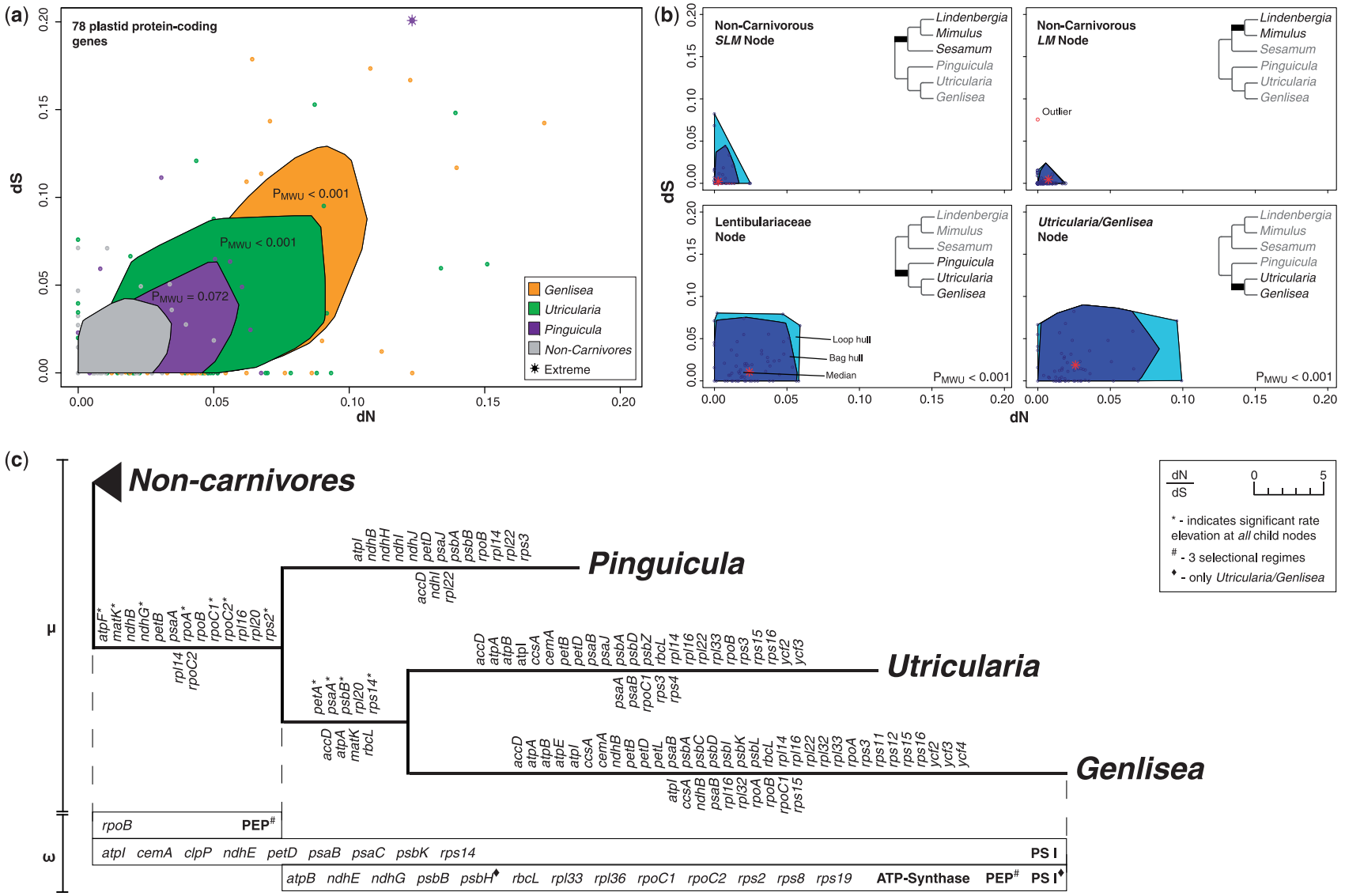


**Fig. 3.** Graphical summary of changes in nucleotide substitutions rates and selective regimes in plastid genes of Lentibulariaceae. Nonsynonymous (dN) and synonymous rates (dS) from 78 plastid protein-coding genes in Lentibulariaceae and noncarnivorous close relatives are represented as bivariate boxplots. (*a*) Boxplot bag hulls for dN and dS in *Pinguicula*, *Utricularia*, and *Genlisea*. *P* values shown are from Wilcoxon tests comparing the range of dN and dS for each of the three carnivores, respectively, with all noncarnivores. (*b*) dN and dS for internal branches. (*c*) Graphical summary showing the genes with elevated dN (above branches) or dS (below branches). Branch length is proportional to the number of genes that show relaxed purifying selection at the specific node. An asterisk indicates genes with accelerated rates at all child nodes. The number and distribution of significantly different selective regimes for a specified gene (left) or functional complex (right, boldface) is depicted below the tree, where position and length of a box indicate the position of a significant different omega either along the branch leading to the carnivores or only in the carnivore clade or both. For example, two selective regimes are inferred for *rpoB*, with omega on the branch leading to Lentibulariaceae (T) being larger than omega in the rest of the tree, i.e., the Lentibulariaceae "crown group" subtree (A) and the noncarnivores (B). A hash key indicates three selective regimes, and a black rhombus marks the gene with a higher omega only in the bladderwort lineage (including the branch leading to the bladderwort lineages) compared to *Pinguicula* plus noncarnivores (see table 3, supplementary tables S2 and S3). PSI, photosystem I genes, PEP, plastid-encoded polymerase genes.

**Table 2.** Results Evaluating the Differences of Relative Nucleotide Substitution Rates across All Coding (dN, dS, and dN + dS) and Noncoding Regions from Lentibulariaceae versus Noncarnivores.

| | Noncarnivores (*Sesamum* + *Lindenbergia* + *Mimulus*) | | | | | Noncarnivores (*Sesamum* + *Lindenbergia*) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 78 Protein-Coding Genes[a] | | | | | 130[b] Noncoding Regions | | | | |
| | dN + dS | dN (min-max) | $P_{Wilcoxon}$ -Value (dN) | dS (min-max) | $P_{Wilcoxon}$ -Value (dS) | SC + IR | $\mu$ (min-max, SC) | SC | $\mu$ (min-max, IR) | IR |
| *Pinguicula* | 0.072 | 0–0.123 | 0.746 | 0–0.201 | 0.039 | <<0.001 | 0–0.545 | <<0.001 | 0–0.105 | 0.2286 |
| *Utricularia* | <<0.001 | 0–0.151 | <<0.001 | 0–0.233 | <<0.001 | <<0.001 | 0–0.661 | <<0.001 | 0–0.432 | 0.0809 |
| *Genlisea* | <<0.001 | 0–0.172 | <<0.001 | 0–0.212 | <<0.001 | <<0.001 | 0–0.845 | <<0.001 | 0–0.274 | 0.04931 |

[a]Without *ycf1*.

[b]Ten regions of *Genlisea* were excluded because of gene loss.

Remaining intact genes of the *NDH* complex in *Utricularia* and *Pinguicula* differ significantly in their respective $\omega$ when compared with noncarnivores. Although *ndhE* (LRT, best model: $P = 0.031$) and *ndhG* (LRT, best model: $P = 0.046$) show relaxed purifying selection, *ndhH* (LRT, best model: $P = 0.069$) and *ndhJ* (LRT, best model: $P = 0.065$) exhibit only marginally different $\omega$. *NdhB*, which is retained also in *Genlisea*, shows no significant change of $\omega$.

Significant relaxation of purifying selection occurs in another 19 plastid genes across all functional complexes (graphically summarized in fig. 3). These changes apply to the entire Lentibulariacae, except for the *psbH* gene, in which the change of $\omega$ is confined solely to the bladderwort lineage (supplementary tables S2 and S3, Supplementary Material online).

There is a tendency of genes in close proximity to one another to exhibit similar changes of selection (*atpI/rps2*, *ndhE/G/psaC*, *clpP/psbB*, *psbH/petB/D*, *rps7/12*, *atpB/rbcL*, *rpl33/rps18*, *psaB/rps14*, *rpl36/rps8/14*, *rpl22/rps19/rpl2*; figs. 1 and 3). Most of those genes are organized as neighboring genes within the same plastid operon-like transcription unit (hereafter: operon). In several cases (*petB*, *rpl2*, *rpl14*, *rpl22*, *rps7*, *rps12*), relaxation of purifying selection in neighboring genes has set in independently, that is, at a different time during the evolution of Lentibulariaceae (fig. 3c; supplementary tables S2 and S3). *PsbK* and *cemA* represent the only cases where we find relaxation of purifying selection (*psbK*–LRT, best model $P = 0.020$, *cemA*–LRT, best model $P = 0.046$), but not in their neighboring genes. However, the *cemA* flanking genes *petA* and *ycf4* show significant elevation of nonsynonymous changes in both species of the bladderwort lineage or only in *Genlisea* (fig. 3c; supplementary table S1, Supplementary Material online), respectively. *Genlisea* shows acceleration of dN also in *psbI*, which is localized adjacent to *psbK* (fig. 3c; supplementary table S1, Supplementary Material online).

## Evolution of Plastid-Noncoding DNA

Relative to noncarnivorous lamiids, all Lentibulariaceae analyzed here exhibit notably higher rates of nucleotide substitutions across the vast majority of the 130 plastid noncoding regions (ncDNA; fig. 4 and table 2). Plastid ncDNA of single-copy regions shows evident sequence divergences in both introns and intergenic spacers ($P_{MWU} < 0.001$, all species), whereas ncDNA in the IRs evolves more conservatively and similar to noncarnivores (fig. 4 and table 2). Notable divergences in spacers and introns are seen most clearly in the IR of *Genlisea* ($P_{MWU} = 0.049$), even though an analogous trend exists also in *Utricularia* albeit only marginally significant ($P_{MWU} = 0.081$). Overall, *Genlisea* accumulates more nucleotide changes in ncDNAs than its sister lineage *Utricularia* ($P_{MWU} < 0.001$), which evolves much faster than *Pinguicula* ($P_{MWU} < 0.001$). No significant differences exist in noncoding regions of the IR between pairs of carnivorous lineages (*Pinguicula* vs. *Utricularia*: $P_{MWU} = 0.691$; *Pinguicula* vs. *Genlisea*: $P_{MWU} = 0.377$, *Utricularia* vs. *Genlisea*: $P_{MWU} = 0.729$; table 2). Notably, rates are elevated in the SSC of *Pinguicula* and *Genlisea*, coinciding with the loss of *ndh* genes. Patterns of ncDNA evolution in the SSC region of the *Utricularia* plastome are, however, convergent with those of noncarnivores. The divergence of ncDNA is most extreme in the intergenic spacers *trnH-psbA*, *trnK-rps16*, *rps16-trnQ*, and *trnG-R*, whereas short or internally transcribed spacers (e.g., *rpoC1-B*) show the greatest sequence conservation (figs. 1 and 4).

We detected a strong correlation between high sequence divergence and low GC contents ($P_\rho < 0.001$, $\rho = -0.685$; fig. 4; supplementary fig. S1, Supplementary Material online). The percentage of GC in nontranscribed spacers tends to be considerably lower than elsewhere in the plastome, even though a correlation is insignificant.

Rate acceleration in plastid spacers and introns matches the distribution of genes with accelerated dN and/or dS ($P_{Sign} = 0.016$) or relaxed purifying selection in *Utricularia* and *Genlisea* (($P_{Sign} = 0.019$), respectively. This trend is also obvious in *Pinguicula*, albeit here with marginal significance only ($P_{Sign} = 0.052$), indicating that *Pinguicula* exhibits fewer rate changes in protein-coding regions than in its noncoding regions.

## Evolution of Plastome Microstructure

Lentibulariaceae plastomes accumulate much more insertions and deletions (indels) than *Sesamum* with 108 indels in *Pinguicula*, 180 in *Utricularia*, and 242 in *Genlisea*. Totaling up to 1,030 indels in *Genlisea* versus 789 in *Sesamum* (fig. 2), the ratio of insertions to deletions is similar in all taxa.

**Table 3.** Omega Values and Results of Maximum Likelihood Tests Evaluating Different Models of Selectional Changes in Eight Functional Classes of Plastid Genes between Lentibulariaceae and Lamiid Relatives.

| | Single ω Model | Two Selective Paths | | | | | | | | | Three Selective Paths | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Branch (T) versus Clade A/B | | | Clade A/Branch versus Clade B | | | Clade A versus Clade B /Branch | | | Clade A, Clade B, Branch | | | |
| Gene Class | $\omega_{global}$ | $\omega_{A+B}$ | $\omega_T$ | *P* Value[a] | $\omega_{A+T}$ | $\omega_B$ | *P* Value[a] | $\omega_A$ | $\omega_{B+T}$ | *P* Value[a] | $\omega_A$ | $\omega_B$ | $\omega_T$ | *P* Value[a] |
| *Lentibulariaceae node: Pinguicula/Utricularia/Genlisea (A) versus noncarnivores (B) and the branch leading to Lentibulariaceae (T)* | | | | | | | | | | | | | | |
| *atp* | 0.082< | 0.080< | 0.076< | 0.341 | *0.097<* | *0.059<* | *0.0034* | 0.096< | 0.065< | 0.014 | 0.095< | 0.059< | 0.075< | 0.014 |
| | 0.094< | 0.092< | 0.121< | | *0.115<* | *0.073<* | | 0.115< | 0.079< | | 0.115< | 0.073< | 0.120< | |
| | 0.107 | 0.105 | 0.180 | | *0.135* | *0.089* | | 0.137 | 0.094 | | 0.136 | 0.089 | 0.179 | |
| *pet* | 0.057< | 0.055< | 0.034< | 0.681 | 0.060< | 0.043< | 0.234 | 0.059< | 0.046< | 0.315 | 0.058< | 0.043< | 0.034< | 0.499 |
| | 0.070< | 0.069< | 0.084< | | 0.079< | 0.060< | | 0.079< | 0.062< | | 0.079< | 0.060< | 0.083< | |
| | 0.085 | 0.084 | 0.165 | | 0.102 | 0.081 | | 0.104 | 0.082 | | 0.103 | 0.081 | 0.163 | |
| *psa* | 0.032< | 0.032< | 0.008< | 0.346 | 0.041< | 0.017< | 0.002 | *0.044<* | *0.018<* | *<0.001* | 0.044< | 0.017< | 0.008< | 0.002 |
| | 0.039< | 0.040< | 0.025< | | 0.053< | 0.026< | | *0.056<* | *0.026<* | | 0.058< | 0.026< | 0.025< | |
| | 0.048 | 0.050 | 0.057 | | 0.067 | 0.036 | | *0.074* | *0.035* | | 0.074 | 0.036 | 0.057 | |
| *psb* | 0.034< | 0.032< | 0.029< | 0.2886 | 0.036< | 0.026< | 0.1233 | 0.035< | 0.028< | 0.298 | 0.034< | 0.026< | 0.029< | 0.237 |
| | 0.040< | 0.039< | 0.060< | | 0.046< | 0.034< | | 0.045< | 0.037< | | 0.045< | 0.034< | 0.059< | |
| | 0.048 | 0.047 | 0.104 | | 0.057 | 0.044 | | 0.057 | 0.046 | | 0.056 | 0.044 | 0.104 | |
| *rpl* | 0.283< | 0.282< | 0.191< | 0.934 | 0.294< | 0.241< | 0.303 | 0.296< | 0.296< | 0.280 | 0.296< | 0.241< | 0.188< | 0.551 |
| | 0.316< | 0.317< | 0.308< | | 0.340< | 0.287< | | 0.346< | 0.3462< | | 0.346< | 0.287< | 0.304< | |
| | 0.352 | 0.354 | 0.462 | | 0.391 | 0.339 | | 0.402 | 0.402 | | 0.401 | 0.339 | 0.457 | |
| *rpo* | 0.190< | 0.182< | 0.257< | 0.002 | 0.220< | 0.149< | <0.001 | 0.211< | 0.165< | 0.003 | *0.209<* | *0.149<* | *0.254<* | *<0.001* |
| | 0.202< | 0.194< | 0.316< | | 0.238< | 0.164< | | 0.229< | 0.180< | | *0.228<* | *0.164<* | *0.313<* | |
| | 0.214 | 0.207 | 0.383 | | 0.257 | 0.180 | | 0.249 | 0.196 | | *0.247* | *0.180* | *0.379* | |
| *rps* | 0.742< | 0.730< | 0.707< | 0.481 | 0.796< | 0.638< | 0.145 | 0.782< | 0.665< | 0.255 | 0.780< | 0.637< | 0.703< | 0.323 |
| | 0.803< | 0.792< | 0.989< | | 0.884< | 0.718< | | 0.87420< | 0.743< | | 0.872< | 0.718< | 0.985< | |
| | 0.866 | 0.857 | 1.334 | | 0.977 | 0.805 | | 0.974 | 0.826 | | 0.971 | 0.804 | 1.329 | |
| *ndh*[a] | 0.828< | 0.855< | 0.410< | 0.091 | 0.783< | 0.828< | 0.69 | 0.807< | 0.807< | 0.779 | 0.856< | 0.829< | 0.394< | 0.210 |
| | 0.873< | 0.905< | 0.58< | | 0.852< | 0.888< | | 0.892< | 0.864< | | 0.945< | 0.888< | 0.560< | |
| | 0.920 | 0.956 | 0.778 | | 0.925 | 0.950 | | 0.983 | 0.923 | | 1.038 | 0.951 | 0.753 | |
| *Bladderwort lineage: Utricularia/Genlisea (A) versus noncarnivores plus Pinguicula (B) and the branch leading to the bladderwort lineage (T)* | | | | | | | | | | | | | | |
| *atp* | 0.082< | 0.081< | 0.061< | 0.609 | 0.086< | 0.074< | 0.202 | 0.085< | 0.076< | 0.267 | 0.084< | 0.074< | 0.060< | 0.440 |
| | 0.094< | 0.0933< | 0.112< | | 0.107< | 0.087< | | 0.107< | 0.089< | | 0.106< | 0.087< | 0.110< | |
| | 0.106 | 0.106 | 0.185 | | 0.130 | 0.102 | | 0.133 | 0.103 | | 0.132 | 0.102 | 0.182 | |
| *pet* | 0.057< | 0.054< | 0.063< | 0.145 | 0.053< | 0.051< | 0.672 | 0.046< | 0.055< | 0.793 | 0.045< | 0.051< | 0.063< | 0.346 |
| | 0.070< | 0.067< | 0.139< | | 0.074< | 0.067< | | 0.067< | 0.071< | | 0.066< | 0.067< | 0.139< | |
| | 0.085 | 0.082 | 0.260 | | 0.101 | 0.086 | | 0.094 | 0.090 | | 0.093 | 0.086 | 0.260 | |
| *psa* | 0.032< | 0.030< | 0.036< | 0.106 | *0.039<* | *0.023<* | *0.021* | 0.035< | 0.026< | 0.118 | 0.035< | 0.023< | 0.035< | 0.047 |
| | 0.040< | 0.037< | 0.075< | | *0.053<* | *0.031<* | | 0.050< | 0.034< | | 0.049< | 0.031< | 0.074< | |
| | 0.048 | 0.046 | 0.135 | | *0.069* | *0.041* | | 0.067 | 0.044 | | 0.067 | 0.041 | 0.133 | |
| *psb* | 0.034< | 0.032< | 0.032< | 0.228 | 0.035< | 0.029< | 0.247 | 0.032< | 0.031< | 0.566 | 0.032< | 0.029< | 0.032< | 0.354 |
| | 0.040< | 0.039< | 0.061< | | 0.046< | 0.037< | | 0.044< | 0.039< | | 0.043< | 0.037< | 0.061< | |
| | 0.048 | 0.047 | 0.105 | | 0.060 | 0.046 | | 0.058 | 0.048 | | 0.058 | 0.046 | 0.104 | |

(continued)

**Table 3.** Continued

| Gene Class | Single ω Model | Two Selective Paths | | | | | | | | | Three Selective Paths | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Branch (T) versus Clade A/B | | | Clade A/Branch versus Clade B | | | Clade A versus Clade B /Branch | | | Clade A, Clade B, Branch | | | | |
| | $\omega_{global}$ | $\omega_{A+B}$ | $\omega_T$ | *P* Value[a] | $\omega_{A+T}$ | $\omega_B$ | *P* Value[a] | $\omega_A$ | $\omega_{B+T}$ | *P* Value[a] | $\omega_A$ | $\omega_B$ | $\omega_T$ | *P* Value[a] |
| *rpl* | 0.283 < | 0.282 < | 0.178 < | 0.937 | 0.314 < | 0.244 < | 0.102 | 0.322 < | 0.248 < | 0.080 | 0.322 < | 0.245 < | 0.172 < | 0.213 |
| | 0.316 < | 0.317 < | 0.307 < | | 0.373 < | 0.283 < | | 0.389 < | 0.285 < | | 0.388 < | 0.284 < | 0.299 < | |
| | 0.352 | 0.354 | 0.480 | | 0.438 | 0.326 | | 0.463 | 0.326 | | 0.463 | 0.327 | 0.469 | |
| *rpo* | 0.190 < | 0.190 < | 0.157 < | 0.922 | 0.217 < | 0.167 < | 0.001 | 0.223 < | 0.169 < | *<0.001* | 0.222 < | 0.168 < | 0.154 < | 0.003 |
| | 0.202 < | 0.202 < | 0.206 < | | 0.238 < | 0.181 < | | *0.247* < | 0.183 < | | 0.246 < | 0.182 < | 0.198 < | |
| | 0.214 | 0.215 | 0.262 | | 0.260 | 0.196 | | *0.272* | 0.197 | | 0.271 | 0.196 | 0.253 | |
| *rps* | 0.742 < | 0.727 < | 0.735 < | 0.421 | 0.742 < | 0.704 < | 0.619 | 0.711 < | 0.722 < | 0.876 | 0.705 < | 0.703 < | 0.7304 < | 0.704 |
| | 0.803 < | 0.789 < | 1.039 < | | 0.838 < | 0.779 < | | 0.815 < | 0.796 < | | 0.808 < | 0.779 < | 1.0328 < | |
| | 0.866 | 0.854 | 1.409 | | 0.943 | 0.860 | | 0.929 | 0.875 | | 0.920 | 0.860 | 1.4007 | |
| *ndh*[a] | 0.828 < | 0.825 < | 0.678 < | 0.902 | 0.752 < | 0.829 < | 0.714 | 0.596 < | 0.837 < | 0.600 | 0.488 < | 0.828 < | 0.914 < | 0.725 |
| | 0.873 < | 0.876 < | 0.852 < | | 0.844 < | 0.884 < | | 0.769 < | 0.888 < | | 0.638 < | 0.883 < | 1.123 < | |
| | 0.920 | 0.929 | 1.038 | | 0.942 | 0.941 | | 0.958 | 0.942 | | 0.801 | 0.940 | 1.347 | |

NOTE.—ω values are given with lower and upper confidence levels; AIC scores were employed to discriminate between the different models, and the best fit is indicated in italics.

[a]P values are derived from LRTs versus the single ω model (see Materials and Methods for details).

Of these, 256 indels are found in plastid protein-coding regions; an exception are the plastid encoded photosystem I genes (*psa*), which lack any length variation. The highest numbers of indels accumulate in *ycf2* (38 indels), *rpoc2* (30), *accD* (21), *rpl22* (18), *matK* (17), and *ndh* genes, with *ndhF* (19) in particular. Across all genes, the longest insertions or deletions (>9 nt) occur in *accD*, *ycf2*, *ndh*, and PEP genes.

Accumulation of microstructural changes in the different protein-coding genes does not generally coincide with elevated nucleotide substitution rates in Lentibulariaceae ($P_\rho = 0.257$; table 4). A correlation can be found, however, for PEP ($P_\rho = 0.031$) and, with marginal significance, in 50S ribosomal protein genes ($P_\rho = 0.054$; fig. 5).

In plastid ncDNA, indel evolution is evidently related to nucleotide substitutional patterns (figs. 4 and 5*b*). Spacers and introns with high sequence divergence significantly accumulate more indels than moderately evolving ncDNA regions ($P_\rho < 0.001$; supplementary fig. S1, Supplementary Material online). Especially, GC-poor regions accumulate insertions and deletions ($P_\rho < 0.001$), perhaps reflecting the observation of proliferating SSRs in the bladderwort-lineage relative to *Pinguicula* and noncarnivorous lamiids.

Autapomorphic indels account for the vast majority of microstructural changes in Lentibulariaceae. Visual inspection suggests that only very few are family-specific insertions or deletions. Absence/presence counts revealed that only five indels are apparently unique to Lentibulariaceae in protein-coding regions as opposed to 146 in noncoding regions. These numbers should, however, be regarded with some caution, because an in-depth analyses of indel history requires an improved taxon sampling.

Notably low rates of microstructural changes occur apparently every 2–3 kb (figs. 1 and 4). In contrast to the LSC and IR, this pattern is less prominent in the SSC, likely because of the observed gene losses in the large *ndh* operon. Distance-related alternation of GC and substitution rates is much less pronounced (fig. 5).

## Discussion

### Architecture of Plastid Genomes in Lentibulariaceae

Complete plastome sequencing and comparative genome analyses revealed distinct paths of plastid genome evolution in the three carnivorous genera of Lentibulariaceae. The structure of Lentibulariaceae plastomes is mostly similar to that of closely related noncarnivorous lamiids (fig. 1), with the *Utricularia* lineage evolving most conservatively in terms of gene synteny. Similar results were also reported for the plastome of *U. gibba* based on a draft assembly of the organellar genomes (Ibarra-Laclette et al. 2013), which was published during the final stages of writing this manuscript and could therefore not be considered in our analyses. Local disruptions of gene order occur in *P. ehlersiae* and *G. margaretae* because of the loss of some *ndh* genes. Phylogenetic relationships and the kind and number of missing genes or pseudogenes strongly suggest that the loss of the *ndh* genes has occurred independently in the two lineages. As 8 out of 11 genes were deleted in *G. margaretae* and four gene losses are observed in

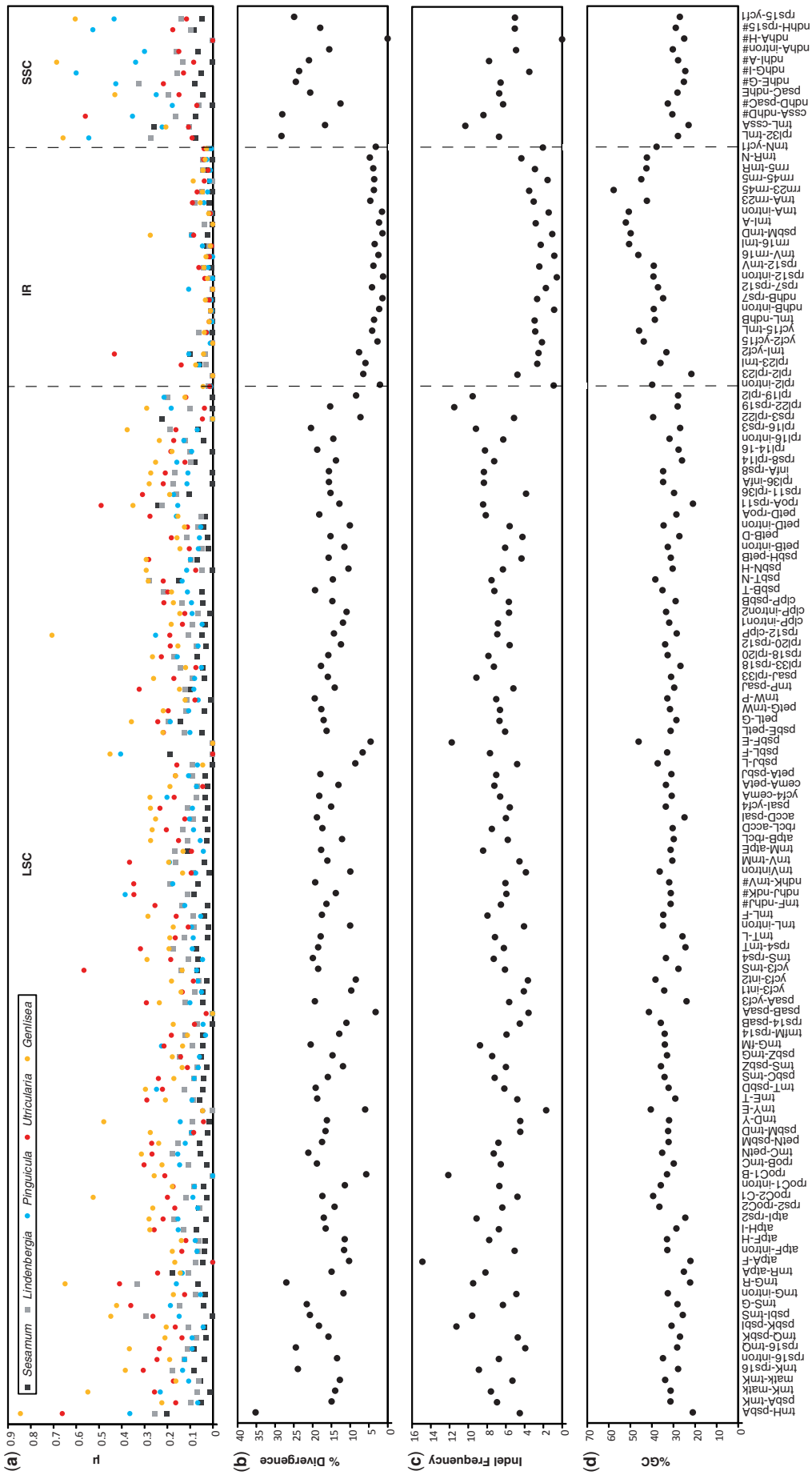


FIG. 4. Nucleotide substitution rates and indel frequencies in plastid noncoding DNA. (*a*) Substitution rates for all plastid noncoding regions are depicted for all Lentibulariaceae and two noncarnivores in the upper diagram, where a row represents a plastid-noncoding regions (as indicated in the bottom of the figure) and the taxon identity is marked by colored dots or squares. The degree of (*b*) divergence, (*c*) the frequency of insertion and deletions event, and (*d*) the GC content is illustrated for all plastid introns and spacers from Lentibulariaceae and the noncarnivores based on the aligned data sets. Dashed lines mark the junctions of single copy regions and the IR. "#" marks regions with lacking data due to gene loss in *Genlisea*.

**Table 4.** Results of Spearman Tests Evaluating the Correlation of Nucleotide Substitution Rates and Indel Events in Plastid Coding and Noncoding Regions.

| | 78 Genes[a] | *atp* | *ndh* | *pet* | *psa* | *psb* | *rpo* | *rpl* | *rps* | 130[b] nc Regions | Introns | Spacer |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 11 | 11 | 11 | 11 | n.a. | 11 | 11 | 11 | 11 | 9 | 9 | 9 |
| $r_s$ | 0.373 | 0.391 | 0.414 | 0.194 | n.a. | 0.593 | 0.650 | 0.297 | 0.444 | 0.970 | 0.950 | 0.964 |
| t | 1.21 | 1.27 | 1.36 | 0.59 | n.a. | 2.21 | 2.56 | 0.93 | 1.49 | 10.47 | 8.05 | 9.59 |
| df | 9 | 9 | 9 | 9 | n.a. | 9 | 9 | 9 | 9 | 7 | 7 | 7 |
| $p_{1\text{-tailed}}$ | 0.129 | 0.118 | 0.103 | 0.285 | n.a. | 0.027 | 0.015 | 0.1881 | 0.085 | <0.001 | <0.001 | <0.001 |
| $p_{2\text{-tailed}}$ | 0.257 | 0.236 | 0.207 | 0.570 | n.a. | 0.054 | 0.031 | 0.377 | 0.170 | <0.001 | <0.001 | <0.001 |

NOTE.—n.a., not tested because of the complete lack of indels.
[a]Without *ycf1*.
[b]Ten regions of *Genlisea* were excluded because of gene loss.

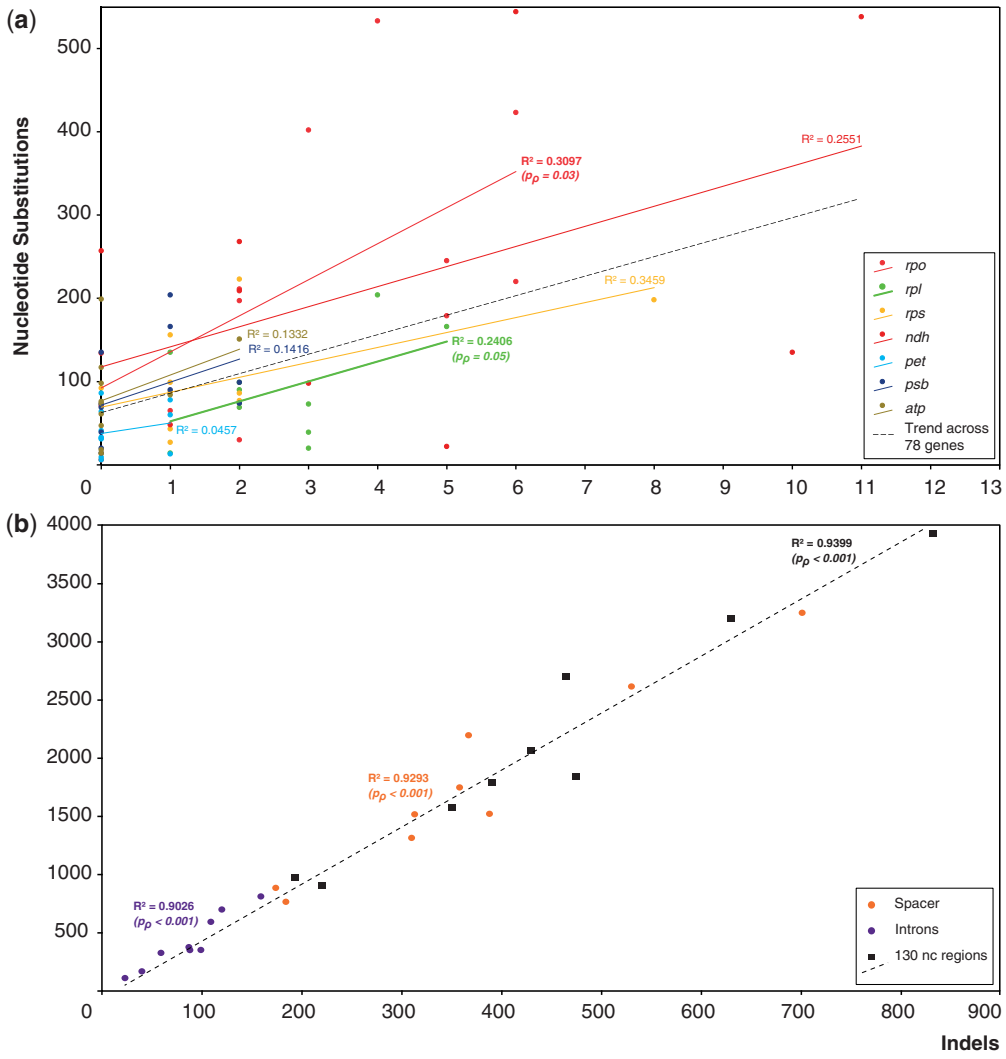


**FIG. 5.** Correlation of microstructural changes and substitution rates in different functional classes of plastid genes. The correlation of indels and substitution rates (total) is illustrated (*a*) for all plastid gene classes and across all plastid protein genes, and (*b*) for spacers, introns, and across all plastid-noncoding regions. The goodness of fit for each of the tested region is provided by $R^2$ (color coded). Where present, *P* values from Spearman tests indicate a significant or marginally significant correlation between the number of indels and substitutions.

*P. ehlersiae*, we believe that these events coincide with the functional loss of the plastid NAD(P)H dehydrogenase complex. In the light of the extreme rare occurrence of functional transfer of genes from the plastid to the nuclear genome in angiosperms (e.g., Bock and Timmis 2008), we appraise the replacement of these genes by nuclear copies as considerably low.

During photosynthesis, the NAD(P)H dehydrogenase complex located in the chloroplast thylakoid membrane mediates electron cycling around the photosystem I and thereby

fine-adjusts the ratio of ATP and the electron carrier NAD(P)H. The latter is used to synthesize monosaccharides by fixating $CO_2$ in the Calvin cycle (Peltier and Cournac 2002; Rumeau et al. 2007). The NAD(P)H dehydrogenase appears to play an especially important role under light-stress conditions or low $CO_2$ concentration, under both of which *ndh* mutants showed significant impairments of photosynthesis (Horvath et al. 2000); no such effects were seen under nonstress conditions and high $CO_2$ concentration. Deletion of *ndh* genes from the plastome was reported earlier for some autotrophic plants (Geraniaceae: Chumley et al. 2006, Blazier et al. 2011; Guisinger et al. 2011; conifers and Gnetales: Brauckmann et al. 2009; Wu et al. 2009). In their natural habitats, carnivorous Lentibulariaceae are exposed to very much the same stress factors as most other plants, which includes the occasional occurrences of light stress. However, the loss of the NAD(P)H complex might not affect the fitness of *Pinguicula* and *Genlisea* or other lineages lacking plastid *ndh* genes. It is possible that these plants regulate gas exchange such that $CO_2$ concentrations remain sufficiently high in the cells or that they make use of organic carbon obtained through prey to overcome periods of environmental stress (e.g., Juniper et al. 1989)—resembling the situation in some parasitic plants to some extent (Krause 2011).

Elevated amounts of plastid repetitive DNA (including SSRs) and low numbers of repeats longer than 50 bp characterize the plastomes of the bladderwort lineage (fig. 2). Some of these repeats colocalize with mutational hotspots in the plastomes of Lentibulariaceae (fig. 1). However, the overall abundance of longer repeats in Lentibulariaceae plastomes is still too low to formally explore a correlation between mutational hotspots and longer repeats, although these are known to drive molecular evolution (e.g., McDonald et al. 2011). Larger repeats are generally rare among angiosperm plastomes and considered to promote illegitimate recombination (Raubeson et al. 2007). Proliferation of long repeats, otherwise, has only been reported for a number of plants with highly reconfigured plastomes (Cosner et al. 2004; Cai et al. 2008; Guisinger et al. 2011; Sloan et al. 2012) or for parasitic plants undergoing relaxed selection of the photosynthesis apparatus (Wicke 2013). Impaired function of DNA repair enzymes or relaxed selection (or both in concert) have been suggested to contribute to the proliferation of recombinogenic factors in plastomes (Guisinger et al. 2008, 2011; Wicke et al. 2013) and might likely result in significant acceleration of both nonsynonymous and synonymous changes and the accumulation of indels.

## Evolution of Substitution Rates and Selective Regimes in Plastid Genes

Across the vast majority of plastid coding and single-copy noncoding regions, *Pinguicula*, *Utricularia*, and *Genlisea* exhibit elevated substitution rates compared with closely related noncarnivorous lamiids (table 2 and fig. 3). Accelerated rates occur in all gene classes of the Lentibulariaceae plastomes, affecting the normally highly conserved plastid-encoded fractions of photosystems, photosynthetic electron transport chains, the thylakoid ATP synthase complex, and genes for transcription and translation.

Plastid genes of Lentibulariaceae show disproportionally (i.e., more frequently) an elevation of nonsynonymous rates than of synonymous rates. This trend is especially pronounced in the bladderwort lineage, where more than half of the protein-coding genes are sped up significantly in both dN and dS. Although the number and type of faster evolving genes differ notably among the Lentibulariaceae species, rate acceleration is not totally random. Accelerated genes shared by all lineages suggest changes in the molecular evolution of some housekeeping genes (e.g., *matK*, *rpo* genes) and electron transport genes that have taken place already in the ancestor of Lentibulariaceae.

In contrast to the general evolutionary patterns seen across angiosperms (Jansen et al. 2007), Lentibulariaceae do not show a correlation of indels and substitution rates across all plastid-coding regions, and most gene classes (except for PEP, and *rpl* genes to a smaller extent) have accumulated more nucleotide substitution than indels. Even though we cannot rule out sampling artifacts contributing to this result, this could imply that the accumulation of nucleotide substitutions in plastid protein-coding regions of Lentibulariaceae proceeds faster than the accumulation of microstructural changes.

Changes of selective constraints either affect the entire clade of Lentibulariaceae (photosystem I and other photosynthesis genes, *clpP*, *rps14*; fig. 3*c*) or are confined to the bladderwort lineage (ATP synthase, several photosynthesis and housekeeping genes; fig. 3*c*). We also detected relaxed purifying selection in plastid-encoded genes for transcription (PEP), transcript maturation (*matK*), translation (ribosomal protein genes), and protein turnover (*clpP*). The polymerase complex is inferred to have experienced a two-step relaxation of purifying selection in Lentibulariaceae: once during the transition to carnivory (i.e., on the branch leading to Lentibulariaceae) and later during the formation of the bladderwort lineage, i.e., along the branch uniting *Utricularia* and *Genlisea* (fig. 3*c*; supplementary tables S2 and S3, Supplementary Material online). Unlike in the other protein–gene classes, the evolution of indels and nucleotide substitution also correlates significantly in the *rpo* genes. This deviating trend of microstructural changes together with the detected relaxation of purifying selection probably indicates that PEP genes have become pseudogenes or are about to become pseudogenes in Lentibulariaceae. This hypothesis would invoke a shift away from PEP to the nuclear-encoded polymerase in plastid transcription, for which there is no data available as of writing this article. On the other hand, the evolution of the *rpo* genes themselves may be the result of mutations in the replication and or repair machinery.

Interestingly, some subunits of the different plastid functional complexes evolve under more relaxed constraints than others. Future studies, which will also need to consider protein-domain interactions, are required to shed light on how much variation is permitted in protein complexes to retain protein function.

In a considerable number of cases, rate elevation and relaxation of purifying selection affect genes that are encoded in close proximity to one another. Higher substitution rates in a gene appear to influence its neighboring genes, even though they might be of a different function. Interestingly, genes in the highly reconfigured plastome of *Pelargonium* (Guisinger et al. 2008) show a comparable, yet less prominent, pattern of higher dN/dS values in neighboring genes. "Localized hypermutation" has also been described in Fabaceae, particularly around the *accD-ycf4* region. A plastome-wide analysis of substitution rate elevation in *Lathyrus* also revealed such a neighboring effect for a few other regions, although to a much smaller extent (e.g., *rpl16-rpoA*, *ndhH-ycf1*; Magee et al. 2010). Potential reciprocal effects of a similar kind have been observed with regard to the series of functional and physical gene losses in nonphotosynthetic parasitic plants (Wicke 2013; Wicke et al. 2013), where DNA deletions seem to be governed by protection from deletion by essential genes. More data are currently being compiled to test this pattern in the framework of spatial autocorrelation and potentially underlying molecular mechanisms.

## Evolution of Noncoding DNA

Nearly all plastid ncDNAs in Lentibulariaceae evolve with accelerated substitution rates compared with noncarnivores. Many of the extremely accelerated spacers and introns are located directly adjacent to or are enclosed by protein-coding genes with elevated dN, dS, or relaxed purifying selection. This pattern is particularly prominent in the bladderwort lineage and suggests the existence of various mutational hotspots in the plastomes of Lentibulariaceae.

A strong correlation of indels and substitution rate is evident in plastid noncoding regions of all Lentibulariaceae (fig. 4). The extent of divergence in spacer regions is obviously strongly related to its GC content. AT richness contributes to potentially deleterious recombination. Long homopolymer stretches and microsatellite regions favor replication errors (Levinson and Gutman 1987; Wagner et al. 1990; Wolfson et al. 1991; Müller et al. 1999), which may explain the proliferation of microstructural changes in the plastomes of Lentibulariaceae (fig. 2; supplementary fig. S1, Supplementary Material online). Due to the current lack of comparable data, it remains elucidated whether the evolutionary patterns of plastid ncDNAs are characteristic for Lentibulariceae or if these can also be observed in other angiosperm plastomes (Müller and Wicke 2013).

## Implications of Relaxed Purifying Selection in Plastid Genes of Lentibulariaceae

Several hypothesized biological phenomena underlying rate variation in angiosperms have already been excluded as potential explanations for Lentibulariaceae (e.g., generation time effect: Jobson and Albert 2002; Müller et al. 2004; speciation rate hypothesis: Müller et al. 2006), while differences in metabolic rates (Martin and Palumbi 1993) and unequal efficiency of DNA repair (Britten 1986) are difficult to assess and still await experimental testing in this family. Taking into account the extraordinary physiological differentiations of Lentibulariaceae, Albert et al. (2010) formulated an alternative mechanistic hypothesis, according to which the genomic peculiarities seen in *Utricularia* may be a consequence of high levels of mutagenic reactive oxygen species (ROS) as the result of a short circuit in the mitochondrial respiratory chain. Due to a mutation in cytochrome c oxidase (Jobson et al. 2004; Laakkonen et al. 2006), this shunt possibly provides energy from sequestering protons more rapidly for the active prey capture mechanism ("firing") of *Utricularia* traps (Albert et al. 2010). Although several lines of evidence are in support of this hypothesis (e.g., Ibarra-Laclette et al. 2011), a few details in the observed rate patterns in plastid genomes are difficult to explain when interpreted exclusively in the light of the ROS hypothesis (as it is currently phrased). Rate acceleration as well as relaxation of purifying selection also occur in *Pinguicula*, which does not share the *COX* mutation with the bladderwort lineage. Substitutional and microstructural changes are even higher in *Genlisea*, which captures its prey exclusively passively, other than in *Utricularia*. Likewise, ultrastructural data from *Utricularia* traps provide evidence for the presence of passive traps in basal lineages of the genus (Müller et al. 2004; Reifenrath et al. 2006), resulting in a more likely evolutionary scenario of a most recent common ancestor of the bladderwort lineage that had not captured prey actively and thus may have benefited less from the short circuit in the mitochondrial respiratory chain. Even though full genome sequencing showed that the organellar genomes evolve more conservatively and apparently uncoupled from the evolutionary dynamics of the nuclear genome in *U. gibba* (Ibarra-Laclette et al. 2013), further comparative genomic studies based on a dense taxonomic sampling will be very helpful in evaluating the role of ROS in shaping genome architectural traits in Lentibulariaceae. Most DNA damage is likely to occur at the immediate source of ROS, i.e., in mitochondria (Friedberg et al. 2006), so particularly comparative mitochondrial genome analyses are required to specifically test the outreach of the ROS hypothesis.

As mentioned above, the patterns of rate elevation and relaxation of selection as well as the loss of *ndh* genes in plastid genomes of Lentibulariaceae are also interesting because of its similarities to obligate photosynthetic parasitic plants like *Cuscuta* (McNeal, Arumugunathan, et al. 2007; McNeal, Kuehl, et al. 2007) and Orobanchaceae (Wicke 2013; Wicke et al. 2013). Given these convergences with photosynthetic parasites regarding molecular evolutionary patterns, it stands to reason that alternative paths of acquiring nutrients may likely promote the rapid evolution of plastid genes in Lentibulariaceae. Even though all Lentibulariaceae are photosynthetic and capable of carbon fixation, carnivorous plants have been demonstrated to counterbalance negative photosynthesis benefits due to unfavorable environmental conditions by the uptake of nitrogen, phosphorous, and organic carbon from their prey (Juniper et al. 1989; Adamec 1997). Besides metabolized prey, a close interaction with microbial communities in or near traps of *Utricularia* has been speculated to contribute to alternative paths of nutrient uptakes (Sirová et al. 2010), which therefore might change

selection pressures in normally highly conserved genomic fractions.

Although causation remains to be elucidated, relaxed purifying selection in photosynthesis and photosynthesis-related plastome regions appears to be mirrored to some extent in average net photosynthesis rates of carnivorous plants. Ellison (2006) states that both the photosynthetic nutrient use efficiency as well as the photosynthesis rate per leaf mass is considerably lower in aquatic and terrestrial carnivorous plants than in noncarnivores. Estimates range between 1 and 5% lowered photosynthesis rates in carnivores (Mendez and Karlsson 1999; Adamec 2011). Subarctic *Pinguicula* species were even reported to show a reduction of both leaf and mass-based photosynthesis rates by 30% compared with geographically similarly distributed noncarnivorous plants, including hemiparasitic plants (Mendez and Karlsson 1999). These results are, however, confounded by reports of very high photosynthesis rates in aquatic *Utricularia* species relative to other aquatic plants (Adamec 2009; Sirová et al. 2010). Interestingly, these high photosynthesis rates appear to directly benefit the prey capturing and cultivation of microbial communities in the traps of the investigated bladderworts (Sirová et al. 2010).A denser sampling of genome-scale data will be needed to thoroughly evaluate causes and consequences of genomic evolutionary patterns in relation to nutritional and physiological peculiarities in carnivorous plants in general, and in Lentibulariaceae in particular.

## Materials and Methods

### Plant Material and Shotgun Sequencing

*Pinguicula ehlersiae* (cultivated at the Bonn Botanical Gardens, Germany, voucher: BONN, 3252-1998, B. Schäferhoff 60, 2009-01-05), *U. macrorhiza* (collected in the Black Moshannon Lake, PA, voucher: PAC, leg./det.: B. Schäferhoff and S. Wicke, 2009-07-22), and *G. margaretae* (cultivated and kindly provided by Dr A. Fleischmann, University of Munich, Germany; voucher: M, A. Fleischmann Z29, 2006-09-27), representing the three genera of Lentibulariaceae, were selected for whole genome shotgun-sequencing. Total genomic DNA was extracted from fresh plant material using a modified cetyl trimethylammonium bromide (CTAB) method, which included adjustment of CTAB extraction buffer to pH 9.8 to maximize DNA yield. All samples were treated with RNAse (Qiagen) before purification by PEG-8000 precipitation (McNeal et al. 2006). High-molecular weight DNA (between 3.5 and 5 µg) was subjected to whole genome shotgun sequencing with 454 FLX Titanium at the Genomics Core Facility of the Pennsylvania State University, University Park, PA.

### Plastid Genome Reconstruction and Sequence Finishing

Pyrosequencing data were quality trimmed and clipped off of the linker sequences plus the adjacent 15 nucleotides. The processed data sets contained 555,679 reads (159 Mbp) for *P. ehlersiae*, 521,827 reads (121 Mbp) for *U. macrorhiza*, and 356,598 reads (96 Mbp) for *G. margaretae*. Reads were assembled using MIRA v3 (Chevreux et al. 1999; Chevreux 2011) under the accurate, de novo, genomic assembly mode.

Contigs were sorted according to their similarity to known plastid protein-coding genes using the local NCBI Blast package and a custom plastid gene database. Plastid-like contigs were extracted from the contig pool (BlastX evalue: 5e−30, BlastN evalue: 5e−50) and automatically postassembled in CAP3 (Huang and Madan 1999) using a minimum identity of 97%, a minimal overlap of 100 nt, and a gap length of max. 5 nt. This approach produced two or three supercontigs in *Genlisea* and *Pinguicula*, respectively, and six supercontigs in *Utricularia*, all that were joined manually. All manual overlaps and potential conflicts in contig ends as well as junctions between plastome single-copy and inverted-repeat regions were verified by PCR and Sanger resequencing. Assembly breakpoints included the IR-SSC junction (around the *ycf1* region) and *trnH-psbA* in all species, the *trnS-trnG* region in *Utricularia* and *Pinguicula*, and *rps16-trnQ*, *psaA-ycf3*, and *ndhG-ndhI* in *Utricularia*. Except for the *ycf1* region, which contained very long homopolymer stretches, all contig breakpoints appeared to have been caused by alignment issues in short repeat- or microsatellite-rich regions during the automatic assembly rather than because of the presence of divergent copies from the mitochondrial or the nuclear genome.

We mapped all 454 reads to the reconstructed draft plastome sequence in Geneious Pro v5.6 to further verify the correct assembly of the plastid chromosome; read mapping results are shown in supplementary figure S2 (Supplementary Material online). The coverage of the final complete sequences were 37.7× in *Pinguicula* (plastid read abundance: 2.87%), 58.8× in *Utricularia* (plastid read abundance: 7.59%), and 84.8× in *Genlisea* (plastid read abundance: 11.94%).

### Annotation and Structural Analysis

Plastome sequences were annotated with DOGMA (Dual Organellar GenoMe Annotator; identity cutoff for protein-coding genes: 35%, tRNA identity cutoff: 90%, e value: 5; Wyman et al. 2004) with manual refinements regarding gene start, gene stop, and exon–intron boundaries. The completely annotated sequences have been submitted to the European, American and Japanese (EMBL/GenBank/DDBJ) nucleotide database collaboration.

Similarity between Lentibulariaceae and close relatives was assessed using progressiveMauve 3.1 (Darling et al. 2010), assuming colinear genomes with an initial seed weight of 17 and a gap open penalty of −200 to account for small gaps caused by localized deletions.

Because of the good results (assessed by eye) and fewer computational resources and time required for the multiple sequence alignment of complete plastomes, we used MAFFT v6 (Katoh and Toh 2008) under the FFT-NS-I mode and the 1PAM/$\kappa$ = 2 scoring matrix. Based upon those alignments, we employed the Indel Analysis tool v1.0.0 of the Galaxy suite (Blankenberg et al. 2010) to obtain genome-wide counts of insertions and deletions relative to a reference genome

(*Sesamum*). The results were grouped into three size classes (2–15 nt, 16–30 nt, and >30 nt) with the number of 1 nt indels being counted separately, because this type is often an artifact from sequencing and/or read assembly (e.g., Moore et al. 2006). A corrected relative insertion number and deletion number (cRIN, cRDN), respectively, was therefore calculated as the difference between the total number of indels and the number of 1 nt indels. For the analyses of SSRs and other repeated elements, we removed one IR of each of the analyzed species of Lentibulariaceae, because the IR is already defined as a large repeat and subrepeats may therefore be counted twice. Di-, tri-, and tetrarepeat units occurring at least five, three, and three times, respectively, were analyzed with the help of the SSR extractor tool, excluding cases of SSRs comprising repeated subunits (e.g., TATA, TGTG). REPuter (Kurtz et al. 2001) was employed to search for all forward and palindromic repeats of min. 12 nt length (Hamming distance: 3, e value: 5e−5).

We created separate alignments for coding and noncoding regions for analyses of microstructural changes, for which we preferred the use of automatic alignments without subsequent manual modifications to ascertain both reproducibility and accurate data processing by batch scripting. The concatenated and annotated data set comprised 78 unique protein-coding genes of all Lentibulariaceae plastomes and four related lamiid taxa: *Mimulus guttatus* (Mimulus Genome Project, DoE Joint Genome Institute), *Lindenbergia philippensis* (HG530133), *S. indicum* (NC_016433), and *Nicotiana tabacum* (NC_001879). The gene *ycf1* was excluded because of unresolvable sequencing/assembly errors in homopolymer stretches. Data sets were aligned with PRANK using the "align translated codons" mode (Löytynoja and Goldman 2005), because this approach outperformed MAFFT v6 for single gene alignments in terms of quality (assessed by eye).

*Mimulus* was excluded from the analyses of plastid-noncoding DNA (ncDNA), as sequence data of its intergenic spacers was unavailable as of this writing. The ncDNA data set comprised 130 unique regions from *Nicotiana*, *Lindenbergia*, *Sesamum*, and Lentibulariaceae. Because of gene loss, homology assessment was critical for some spacers in the *Genlisea* plastome, and we therefore excluded its *trnF-ndhK*, *ndhJ-K*, *ndhC-trnV*, as well as its SSC ncDNA. Despite pseudogenization of *infA* in *Nicotiana* (Shinozaki et al. 1986; Millen et al. 2001), we included both the *rpl36-ΨinfA* and *ΨinfA-rps8* spacers due to the overall high conservation of the region. Alignment of spacers and introns was carried out in PRANK using the structural model "genomic" for the alignment; a ts/tv ratio of 2, a gap rate of 0.05, and a mean gap length distribution of 5 (for all regions).

Rather than counting indels relative to a reference as done to obtain a general overview across the entire plastome (see above), we inferred length mutational variation in coding regions and ncDNA in more detail with *SeqState 1.4.1* (Müller 2005) under the simple-indel coding option. Correlated evolution of microstructural changes and substitution rates was tested by parsimony-based optimization of indels and total substitution rates on a specified tree as described in Jansen et al. (2007).

## Substitution Rate Analyses

Genewise and gene-class-specific nonsynonymous (dN) and synonymous (dS) rate changes were analyzed by LRTs in Hyphy 2.11 (Kosakovsky-Pond et al. 2004) based on alignments of *Pinguicula*, *Utricularia*, and *Genlisea* with the four noncarnivorous species *Sesamum*, *Lindenbergia*, *Mimulus*, and *Nicotiana*. Besides powerful hypothesis testing using maximum likelihood approaches and several other useful sequence data analysis tools, the Hyphy software offers more flexibility than other currently available programs because of its own scripting (batch) language (HBL), which allows users to design customized tests. A specified lamiid topology that reflects generally established phylogenetic relationships of these taxa (cf. fig. 3) was the basis for all tests. In brief, an unconstrained likelihood function (LF) uses a full model (with local parameters for both dN and dS), and the data were optimized. The resulting maximum likelihood estimates for the parameters dN and dS were extracted for subsequent graphical representation. To test for the significance of differences across species, a constrained LF that assumes dN or dS to be equal in a selected pair of taxa was optimized, and the difference of likelihood scores from constrained and unconstrained functions (LRT statistic) was evaluated by $\chi^2$ distribution.

Using custom R scripts, unpaired Wilcoxon tests with sequential alpha error (Bonferroni) correction were employed to evaluate cross-species differences of dN, dS, and $\omega$ among Lentibulariaceae and noncarnivores. Results of pairwise analyses were visualized as bivariate boxplots (Rousseeuw et al. 1999).

Changes in $\omega$ across clades and branches were evaluated using the HBL scripts "SelectionLRT" and "TestBranchDNDS" (Frost et al. 2005; Kosakovsky-Pond et al. 2007). SelectionLRT tests whether differences in selection exist between two clades, A and B, and the branch T leading to A, using LRTs against a global-$\omega$ model. AIC is then used to select from models, assuming either two different $\omega$ (one for A + B and one for T; one for A + T and one for B; one for A and one for B + T) or three different $\omega$ (each one for A, B, and T). The test requires two distinct clades consisting of two distinct branches with at least two terminal taxa each. SelectionLRT was run genewise and for concatenated data sets, the latter representing distinct functional complexes with roles in either photosynthesis or housekeeping. As we aimed at unraveling selective changes coinciding with the transition to carnivory and with further differentiation of the bladderwort lineage within Lentibulariaceae, two sets of analyses were conducted: 1) treating the Lentibulariaceae crown group as clade A, the branch leading to Lentibulariaceae as T, and the noncarnivorous rest of the tree as B, and 2) treating the terminal branches *Utricularia* + *Genlisea* as A, the branch leading to this clade as T, and *Pinguicula* plus the noncarnivores as B. Taxa lacking a certain gene were excluded from the respective tests (e.g., *infA* without *Nicotiana*). As both *Pinguicula* and *Genlisea* have lost *ndhC*, *ndhD*, *ndhF*, and *ndhK*, the Lentibulariaceae clade collapsed (with respect to the requirements for the SelectionLRT script), so tests for selectional

changes in these genes between *Utricularia* and noncarnivores were conducted with TestBranchDNDS, which allows testing differences in $\omega$ between a single branch and the rest of a given tree. All tests were run under the MG94xREV model, allowing for equivalent amino acid changes and assuming two independent gamma/beta distributions for across-site variation in both dN and dS and three different rate classes.

A custom HBL script was employed to perform pairwise nucleotide substitution rate tests under the GTR + G + I model in ncDNA regions between each of the Lentibulariaceae taxa, *Sesamum*, and *Lindenbergia*; *Nicotiana* was used as outgroup. LRTs were used to test for significance of rate changes in single spacers/introns as described above; across-species differences were evaluated by sequential Bonferroni-corrected unpaired Wilcoxon tests. Nonparametric Spearman tests were used to test for correlations between different genomic traits (GC content, indel frequency, nucleotide divergence). A Sign test was used to evaluate whether the distribution of rate elevation or relaxation of purifying selection in protein-coding genes match rate elevation in the adjacent ncDNA regions; *ndh* genes and adjacent ncDNA regions were excluded from the test.

## Supplementary Material

Supplementary figures S1 and S2 and tables S1–S4 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

## Acknowledgments

We thank Andreas Fleischmann (LMU Munich) and the Bonn Botanical Gardens for fresh plant material, and Lena Landherr and Paula Ralph (all Penn State University) for excellent technical advice. Thanks are also due to the PIs of the *Mimulus Genome Sequencing Project* for giving us early access to plastome data, and Yan Zhang (Penn State University) for the *Mimulus* plastome annotation. We also thank three anonymous reviewers and the Associate Editor for helpful comments and suggestions on an earlier version of this manuscript. This study was funded by the Germany Science Foundation, Deutsche Forschungsgemeinschaft (DFG) grant MU2875/2 to K.F.M.

## References

Adamec L. 1997. Mineral nutrition of carnivorous plants: a review. *Bot Rev.* 63:273–299.

Adamec L. 2009. Photosynthetic $CO_2$ affinity of the aquatic carnivorous plant *Utricularia australis* (Lentibulariaceae) and its investment in carnivory. *Ecol Res.* 24:327–333.

Adamec L. 2011. Ecophysiological look at plant carnivory: why are plants carnivorous? In: Seckbach J, Dubinski Z, editors. All flesh is grass.. Plant–animal interrelationships. Cellular origin, life in extreme habitats and astrobiology, Vol. 16. New York: Springer Science + Business Media B. V. p. 455–489.

Adamec L. 2013. Foliar mineral nutrient uptake in carnivorous plants: what do we know and what should we know? *Front Plant Sci.* 4:10.

Albert VA, Jobson RW, Michael TP, Taylor DJ. 2010. The carnivorous bladderwort (*Utricularia*, Lentibulariaceae): a system inflates. *J Exp Bot.* 61:5–9.

Albert VA, Williams SE, Chase MW. 1992. Carnivorous plants: phylogeny and structural evolution. *Science* 257:1491–1495.

Barthlott W, Porembski S, Fischer E, Gemmel B. 1998. First protozoa-trapping plant found. *Nature* 392:447.

Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko A, Taylor J. 2010. Galaxy: a web-based genome analysis tool for experimentalists. *Curr Protoc Mol Biol.* 89: 19.10.1–19.10.21.

Blazier JC, Guisinger-Bellian MM, Jansen RK. 2011. Recent loss of plastid-encoded *ndh* genes within *Erodium* (Geraniaceae). *Plant Mol Biol.* 76: 1–10.

Bock R, Timmis JN. 2008. Reconstructing evolution: Gene transfer from plastids to the nucleus. *Bioessays* 30:556–566.

Brauckmann T, Kuzmina M, Stefanović S. 2009. Loss of all plastid *ndh* genes in Gnetales and conifers: extent and evolutionary significance for the seed plant phylogeny. *Curr Genet.* 55: 323–337.

Braukmann T, Stefanović S. 2012. Plastid genome evolution in mycoheterotrophic Ericaceae. *Plant Mol Biol.* 79:5–20.

Britten RJ. 1986. Rates of DNA sequence evolution differ between taxonomic groups. *Science* 231:1393–1398.

Cai Z, Guisinger MM, Kim H-G, Ruck E, Blazier J, McMurtry V, Kuehl J, Boore J, Jansen RK. 2008. Extensive reorganization of the plastid genome of *Trifolium subterraneum* (Fabaceae) is associated with numerous repeated sequences and novel DNA insertions. *J Mol Evol.* 67:696–704.

Chevreux B. 2011. Sequence assembly with MIRA3—the definitive guide. [cited 2012 Nov]. Available from: http://mira-assembler.sourceforge.net/docs/DefinitiveGuideToMIRA.pdf.

Chevreux B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information. *Comput Sci Biol (Proc Germ Conf Bioinforma).* 99:45–56.

Chumley TW, Ferraris JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK. 2006. The complete chloroplast genome sequence of *Pelargonium* x *hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Mol Biol Evol.* 23:2175–2190.

Cosner ME, Raubeson LA, Jansen RK. 2004. Chloroplast DNA rearrangements in Campanulaceae: phylogenetic utility of highly rearranged genomes. *BMC Evol Biol.* 4:27.

Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss, and rearrangment. *PLoS One* 5: e11147.

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

dePamphilis CW, Young ND, Wolfe AD. 1997. Evolution of plastid gene *rps2* in a lineage of hemiparasitic and holoparasitic plants: many losses of photosynthesis and complex patterns of rate variation. *Proc Natl Acad Sci U S A.* 94:7367–7372.

Ellison AM. 2006. Nutrient limitation and stoichiometry of carnivorous plants. *Plant Biol.* 8:740–747.

Ellison AM, Gotelli NJ. 2009. Energetics and the evolution of carnivorous plants—Darwin's "most wonderful plants in the world". *J Exp Bot.* 60:19–42.

Fajardo D, Senalik D, Ames M, Zhu H, Steffan SA, Harbut R, Polashock J, Vorsa N, Gillespie E, Kron K, et al. 2012. Complete plastid genome sequence of *Vaccinium macrocarpon*: structure, gene content, and rearrangements revealed by next generation sequencing. *Tree Genet Genomes* 9:489–498.

Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T. 2006. DNA repair and mutagenesis. 2nd ed. Washington (DC): ASM Press.

Frost SDW, Liu Y, Pond SLK, Chappey C, Wrin T, Petropoulos CJ, Little SJ, Richman DD. 2005. Characterization of human immunodeficiency virus type 1 (HIV-1) envelope variation and neutralizing antibody responses during transmission of HIV-1 subtype B. *J Virol.* 79: 6523–6527.

Funk H, Berg S, Krupinska K, Maier U, 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.

Greilhuber J, Borsch T, Müller K, Worberg A, Porembski S, Barthlott W. 2006. Smallest angiosperm genomes found in Lentibulariaceae, with chromosomes of bacterial size. *Plant Biol.* 8:770–777.

Guisinger MM, Chumley TW, Kuehl JV, Boore JL, Jansen RK. 2010. Implications of the plastid genome sequence of *Typha* (Typhaceae, Poales) for understanding genome evolution in Poaceae. *J Mol Evol.* 70:149–166.

Guisinger MM, Kuehl JV, Boore JL, Jansen RK. 2008. Genome-wide analyses of Geraniaceae plastid DNA reveal unprecedented patterns of increased nucleotide substitutions. *Proc Natl Acad Sci U S A.* 105: 18424–18429.

Guisinger MM, Kuehl JV, Boore JL, Jansen RK. 2011. Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage. *Mol Biol Evol.* 28:583–600.

Horvath EM, Peter SO, Joet T, Rumeau D, Cournac L, Horváth GV, Kavanagh TA, Schäfer C, Peltier G, Medgyesy P. 2000. Targeted inactivation of the plastid *ndhB* gene in tobacco results in an enhanced sensitivity of photosynthesis to moderate stomatal closure. *Plant Physiol.* 123:1337–1350.

Huang X, Madan A. 1999. CAP3: a DNA sequence assembly program. *Genome Res.* 9:868–877.

Ibarra-Laclette E, Albert V, Perez-Torres C, Zamudio-Hernandez F, Ortega-Estrada M, Herrera-Estrella A, Herrera-Estrella L. 2011. Transcriptomics and molecular evolutionary rate analysis of the Bladderwort (*Utricularia*), a carnivorous plant with a minimal genome. *BMC Plant Biol.* 11:101.

Ibarra-Laclette E, Lyons E, Hernandez-Guzman G, Pérez-Torres CA, Carretero-Paulet L, Chang TH, Lan T, Welch AJ, Juárez MJ, Simpson J, et al. 2013. Architecture and evolution of a minute plant genome. *Nature* 498:94–98.

Jansen RK, Cai Z, Raubeson LA, Daniell H, Depamphilis CW, Leebens-Mack J, Müller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, et al. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc Natl Acad Sci U S A.* 104:19369–19374.

Jobson RW, Albert VA. 2002. Molecular rates parallel diversification contrasts between carnivorous plant lineages. *Cladistics* 18:127–136.

Jobson RW, Nielsen R, Laakkonen L, Wikström M, Albert VA. 2004. Adaptive evolution of cytochrome c oxidase: infrastructure for a carnivorous plant radiation. *Proc Natl Acad Sci.* 101: 18064–18068.

Juniper BE. 1986. The path to plant carnivory. In: Mori SA, editor. Insects and the plant surface. London: Edward Arnold Publishers.

Juniper BE, Robins RJ, Joel DM. 1989. The carnivorous plants. London: Academic Press.

Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. *Brief Bioinform.* 9:286–298.

Kosakovsky-Pond SL, Frost SDW, Muse SV. 2004. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679.

Kosakovsky-Pond SL, Poon AFY, Frost SDW. 2007. Estimating selection pressures on alignments of coding sequences: analyses using HyPhy. In: Salemi M, Vandamme A-M, editors. The phylogenetic handbook: a practical approach to DNA and protein phylogeny. Cambridge University Press. p. 419–450.

Krause K. 2011. Piecing together the puzzle of parasitic plant plastome evolution. *Planta* 234:647–656.

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29:4633–4642.

Laakkonen L, Jobson RW, Albert VA. 2006. A new model for the evolution of carnivory in the bladderwort plant (*Utricularia*): adaptive changes in cytochrome c oxidase (COX) provide respiratory power. *Plant Biol.* 8:758–764.

Levinson G, Gutman GA. 1987. Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol Biol Evol.* 4:203–221.

Löytynoja A, Goldman N. 2005. An algorithm for progressive multiple alignment of sequences with insertions. *Proc Natl Acad Sci U S A.* 102:10557–10562.

Magee AM, Aspinall S, Rice DW, Cusack BP, Sémon M, Perry AS, Stefanović S, Milbourne D, Barth S, Palmer JD, et al. 2010. Localized hypermutation and associated gene losses in legume chloroplast genomes. *Genome Res.* 20:1700–1710.

Martin AP, Palumbi SR. 1993. Body size, metabolic rate, generation time, and the molecular clock. *Proc Natl Acad Sci U S A.* 90: 4087–4091.

McDonald MJ, Wang W-C, Huang H-D, Leu J-Y. 2011. Clusters of nucleotide substitutions and insertion/deletion mutations are associated with repeat sequences. *PLoS Biol.* 9:e1000622.

McNeal JR, Arumugunathan K, Kuehl J, Boore J, dePamphilis C. 2007. Systematics and plastid genome evolution of the cryptically photosynthetic parasitic plant genus *Cuscuta* (Convolvulaceae). *BMC Biol.* 5:55.

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

McNeal JR, Leebens-Mack JH, Arumuganathan K, Kuehl JV, Boore JL, dePamphilis CW. 2006. Using partial genomic fosmid libraries for sequencing complete organellar genomes. *Biotechniques* 41: 69–73.

Mendez M, Karlsson PS. 1999. Costs and benefits of carnivory in plants: insights from the photosynthetic performance of four carnivorous plants in a subarctic environment. *Oikos* 86:105–112.

Millen RS, Olmstead RG, Adams KL, Palmer JD, Lao NT, Heggie L, Kavanagh TA, Hibberd JM, Gray JC, Morden CW, et al. 2001. Many parallel losses of *infA* from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. *Plant Cell* 13:645–658.

Moore M, Dhingra A, Soltis P, Shaw R, Farmerie W, Folta K, Soltis D. 2006. Rapid and accurate pyrosequencing of angiosperm plastid genomes. *BMC Plant Biol.* 6:17.

Müller A., Kamisugi Y, Grüneberg R, Niedenhof I, Hörold R., Meyer P. 1999. Palindromic sequences and A + T-rich DNA elements promote illegitimate recombination in *Nicotiana tabacum*. *J Mol Biol.* 291:29–46.

Müller K, Borsch T, Legendre L, Porembski S, Theisen I, Barthlott W. 2004. Evolution of carnivory in Lentibulariaceae and the Lamiales. *Plant Biol Stuttg.* 6:477–490.

Müller KF. 2005. SeqState: Primer design and sequence statistics for phylogenetic DNA datasets. *Appl Bioinformatics.* 4:65–69.

Müller KF, Borsch T, Legendre L, Porembski S, Barthlott W. 2006. Recent progress in understanding the evolution of carnivorous Lentibulariaceae (Lamiales). *Plant Biol.* 8:748–757.

Müller KF, Wicke S. 2013. Evolutionary patterns of nucleotide substitutions and microstructural changes in non-coding DNA of angiosperm plastid genomes with an emphasis on extremophytes. Botany 2013—celebrating diversity! July 27–31, New Orleans, LA, USA. Abstract ID: 130. [cited 2014 Jan 2]. Available at: http://www.2013.botanyconference.org.

Nickrent DL, Ouyang Y, Duff RJ, dePamphilis CW. 1997. Do nonasterid holoparasitic flowering plants have plastid genomes? *Plant Mol Biol.* 34:717–729.

Palmer JD. 1985. Comparative organization of chloroplast genomes. *Ann Rev Genet.* 19:325–354.

Palmer JD, Osorio B, Thompson WF. 1988. Evolutionary significance of inversions in legume chloroplast DNAs. *Curr Genet.* 14:65–74.

Pavlovič A, Singerová L, Demko V, Hudák J. 2009. Feeding enhances photosynthetic efficiency in the carnivorous pitcher plant *Nepenthes talangensis*. *Ann Bot.* 104:307–314.

Peltier G, Cournac L. 2002. Chlororespiration. *Annu Rev Plant Biol.* 53: 523–550.

Pereira CG, Almenara DP, Winter CE, Fritsch PW, Lambers H, Oliveira RS. 2012. Underground leaves of *Philcoxia* trap and digest nematodes. *Proc Natl Acad Sci. U S A.* 109:1154–1158.

Plachno BJ, Adamec L, Lichtscheidl IK, Peroutka M, Adlassnig W, Vrba J. 2006. Fluorescence labelling of phosphatase activity in digestive glands of carnivorous plants. *Plant Biol.* 8:813–820.

Raubeson LA, Jansen RK. 2005. Chloroplast genomes of plants. In: Henry RJ, editors. Diversity and evolution of plant-genotypic and phenotypic variation in higher plants. Wallingford, Oxfordshire (United Kingdom): CABI Publ. p. 45–68.

Raubeson LA, Peery R, Chumley TW, Dziubek C, Fourcade HM, Boore JL, Jansen RK. 2007. Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics* 8:174.

Reifenrath K, Theisen I, Schnitzler J, Porembski S, Barthlott W. 2006. Trap architecture in carnivorous *Utricularia* (Lentibulariaceae). *Flora* 201: 597–605.

Rousseeuw PJ, Ruts I, Tukey JW. 1999. The bagplot: a bivariate boxplot. *Am Stat.* 53:382–387.

Rumeau D, Peltier G, Cournac L. 2007. Chlororespiration and cyclic electron flow around PSI during photosynthesis and plant stress response. *Plant Cell Environ.* 30:1041–1051.

Schäferhoff B, Fleischmann A, Fischer E, Albach D, Borsch T, Heubl G, Müller KF. 2010. Towards resolving Lamiales relationships: insights from rapidly evolving chloroplast sequences. *BMC Evol Biol.* 10:352.

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, et al. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J.* 5:2043–2049.

Sirová D, Borovec J, Šantrůčková H, Šantrůček J, Vrba J, Adamec L. 2010. *Utricularia* carnivory revisited: plants supply photosynthetic carbon to traps. *J Exp Bot.* 61:99–103.

Sloan DB, Alverson AJ, Wu M, Palmer JD, Taylor DR. 2012. Recent acceleration of plastid sequence and structural evolution coincides with extreme mitochondrial divergence in the angiosperm genus *Silene*. *Genome Biol Evol.* 4:294–306.

Soltis DE, Soltis PS, Endress PK, Chase MW. 2005. Phylogeny and evolution of angiosperms. Sunderland (MA): Sinauer Associates.

Wagner LA, Weiss RB, Driscoll R, Dunn DS, Gesteland RF. 1990. Transcriptional slippage occurs during elongation at runs of adenine or thymine in *Escherichia coli*. *Nucleic Acids Res.* 18: 3529–3535.

Wicke S. 2013. Genomic evolution in Orobanchaceae. In: Joel DM, Gressel J, Musselman LJ, editors. Parasitic Orobanchaceae—parasitic mechanisms and control strategies. Berlin: Springer. p. 267–286.

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

Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol Biol.* 76: 273–297.

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

Wolfson R, Higgins KG, Sears BB. 1991. Evidence for replication slippage in the evolution of *Oenothera* chloroplast DNA. *Mol Biol Evol.* 8: 709–720.

Wu C-S, Lai Y-T, Lin C-P, Wang Y-N, Chaw S-M. 2009. Evolution of reduced and compact chloroplast genomes (cpDNAs) in gnetophytes: selection toward a lower-cost strategy. *Mol Phylogenet.* 52: 115–124.

Wyman SK, Boore JL, Jansen RK. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20:3252–3255.

Young ND, dePamphilis CW. 2005. Rate variation in parasitic plants: correlated and uncorrelated patterns among plastid genes of different function. *BMC Evol Biol.* 5:16.