Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants

Susann Wicke[a,b,1], Kai F. Müller[b], Claude W. dePamphilis[c], Dietmar Quandt[d], Sidonie Bellot[e], and Gerald M. Schneeweiss[a]

[a]Department of Botany and Biodiversity Research, University of Vienna, A-1030 Vienna, Austria; [b]Institute for Evolution and Biodiversity, University of Muenster, 48149 Muenster, Germany; [c]Department of Biology, Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, PA 16802; [d]Nees Institute for Biodiversity of Plants, University of Bonn, 53115 Bonn, Germany; and [e]Department of Plant Biodiversity, Technical University of Munich, 85354 Freising, Germany

Edited by M. T. Clegg, College of Natural and Agricultural Sciences, Irvine, CA, and approved May 31, 2016 (received for review May 12, 2016)

Because novel environmental conditions alter the selection pressure on genes or entire subgenomes, adaptive and nonadaptive changes will leave a measurable signature in the genomes, shaping their molecular evolution. We present herein a model of the trajectory of plastid genome evolution under progressively relaxed functional constraints during the transition from autotrophy to a nonphotosynthetic parasitic lifestyle. We show that relaxed purifying selection in all plastid genes is linked to obligate parasitism, characterized by the parasite's dependence on a host to fulfill its life cycle, rather than the loss of photosynthesis. Evolutionary rates and selection pressure coevolve with macrostructural and microstructural changes, the extent of functional reduction, and the establishment of the obligate parasitic lifestyle. Inferred bursts of gene losses coincide with periods of relaxed selection, which are followed by phases of intensified selection and rate deceleration in the retained functional complexes. Our findings suggest that the transition to obligate parasitism relaxes functional constraints on plastid genes in a stepwise manner. During the functional reduction process, the elevation of evolutionary rates reaches several new rate equilibria, possibly relating to the modified protein turnover rates in heterotrophic plastids.

parasitism | relaxed selection | evolutionary rates | plastid genomes | Orobanchaceae

Lineages change over time as they adapt to new environments. Novel conditions determine the selection in genes or cellular genomes and shape their functional and structural evolution. A system well suited to study the evolution of genomic traits in the context of altered selective regimes that is also tractable technically (due to its small size and high copy number) is the plastid genome (plastome). The prime function of plastids is photosynthesis, but this essential plant organelle also produces starch, lipids, amino acids, sulfur compounds, and pigments. As a result of the strong selective pressure on plastid gene function, plastid genomes have a conserved gene content (1; but see ref. 2) and their genes functioning in photosynthesis (*atp*, *ndh*, *pet*, *psa*, *psb*, *ccsA*, *cemA*, *ycf3*/*4*, *rbcL*), transcription, transcript maturation or translation (*rpo*, *matK*, *rpl*, *rps*, *infA*), and other pathways (*accD*, *clpP*, *ycf1*, and *ycf2*) evolve at lower evolutionary rates than nuclear genes (3). However, in eukaryotic lineages such as Apicomplexan pathogens and nongreen plants that independently made the transition from an autotrophic to a parasitic way of life, plastomes have experienced convergent reductions and accelerations of evolutionary rates (4). Although there is a general understanding of the association of the nonphotosynthetic lifestyle with plastome degradation and rate acceleration, the precise trajectory of plastome evolution under progressively reduced function along the way from being a full autotroph to an obligate nonphotosynthetic parasite remains unknown.

Parasitic plants are an excellent system for studying genome evolution under altered selective constraints because of lifestyle changes such as the transition from an autotrophic to a parasitic way of life (5). These plants directly connect to their host plants through a specialized organ to steal water and nutrients. The large majority of the 4,000–4,500 parasitic flowering plant species are photosynthetic parasites (hemiparasites), whereas only 10% are nonphotosynthetic parasites (holoparasites). Whereas some hemiparasites can complete their life cycle without ever connecting to a host plant (facultative parasites), most hemiparasites and all holoparasites require a host at least during certain life stages to fulfill their life cycle (obligate parasites). The transition from autotrophy to parasitism coincides with the loss of both photosynthetic and housekeeping plastid genes (6, 7). Whether a gene is retained or lost in parasites mainly depends on its function, its size (8), and its physical and/or transcriptional association to essential genes (5). Most of the retained genes continue to evolve under purifying selection despite uncorrelated shifts of codon use (5, 6) and changes in evolutionary rates (4). However, severe reconfigurations of the plastid chromosomal architecture such as increases in the amount of recombinogenic DNA sequences in obligate hemiparasites suggest that already the shift from a free-living to an obligate parasitic lifestyle alters the evolution of the plastome in general (5, 9).

Here, we assess the course of reductive plastome evolution and its underlying causes under progressively relaxed functional constraints. Specifically, we examine mutation rate variation, encompassing nucleotide substitutions and microstructural changes across different functional gene classes, and we test for correlations of mutational rates with lifestyle and genomic features, taking into account potential effects of life history. We use complete plastid genome sequences of 17 parasitic and nonparasitic species across

Significance

Parasitism is a proven way of life that brings about extraordinary phenotypic and genetic modifications. Obtaining organic carbon from a host rather than synthesizing it, nonphotosynthetic plants lose unneeded genes for photosynthesis from their plastid genomes, while essential genes in the same subgenome may evolve rapidly. We show that long before the nonphotosynthetic lifestyle is established, losses of functional complexes repeatedly trigger the disruption of evolutionary stasis, resulting in "roller-coaster rate variation" along the transition to full parasitism. Our model of the molecular evolutionary principles of plastid genome degradation under modified selective constraints makes a significant contribution to our understanding of the complexity of genetic switches in relation to lifestyle changes.

Author contributions: S.W. and G.M.S. designed research; K.F.M. and C.W.d. contributed to the conceptual layout; S.W. performed research; K.F.M., C.W.d., D.Q., and S.B. contributed new reagents/analytic tools; S.W. analyzed data; and S.W. and G.M.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. HG514460, HG515538, HG514459, M81884, HG530133, HG515539, HG515537, KT387722, HG803179, KT387724, KU212370, KU212371, HG515536, HG803180, HG738866, NC016433, KU212372, and KU212369) and in Dryad Digital Repository, datadryad.org (10.5061/dryad.t2m75).

[1]To whom correspondence should be addressed. Email: susann.wicke@uni-muenster.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1607576113/-/DCSupplemental.

all different trophic specializations of the broomrape family (Orobanchaceae) and two closely related autotrophs. Orobanchaceae represent an ideal group for this type of study, because its phylogeny is well understood and it spans the entire range from autotrophy to full parasitism (10). Our data show that the shift to parasitism and the loss of functional gene groups trigger the disruption of evolutionary stasis, resulting in phases of accelerated evolution alternating with phases of deceleration. Our findings provide the basis for a molecular evolutionary model of plastome degradation along the transition to a non-photosynthetic way of life.

Results

Plastid Genomes in Parasitic Orobanchaceae. Complete sequencing of 17 species of Orobanchaceae revealed that genes for only 16 proteins, 4 ribosomal RNAs, and 14 transfer RNAs of the 113 unique genes found in the nonparasitic *Lindenbergia philippensis* (Orobanchaceae), *Erythranthe guttata* (Phrymaceae), and *Sesamum indicum* (Pedaliaceae) are commonly present in the plastid genomes of all of the hemiparasitic and holoparasitic plants (Fig. 1). Retaining between 42 and 71 intact genes (Fig. S1), including intact photosynthesis genes (*atp* genes), the holoparasites are particularly diverse with respect to both gene content and genome structure. Large-scale structural reconfigurations such as inversions, modifications of the large inverted repeat (IR) regions, or their loss characterize the genomes of several parasites, including the obligate hemiparasites *Schwalbea americana* and *Striga hermonthica*, as well as the holoparasites *Conopholis americana*, *Orobanche gracilis*, *Orobanche crenata*, and all *Phelipanche* species (Fig. S2).

Nucleotide Substitution Rates. We compared nonsynonymous (dN) and synonymous (dS) nucleotide substitution rates of all Orobanchaceae with closely related nonparasites (Fig. 1), building on phylogenetic relationships established earlier (10, 11). In gene-by-gene likelihood ratio tests (LRTs), the facultatively hemiparasitic *Triphysaria versicolor* shows hardly any significant rate shifts in plastid genes compared with nonparasitic plants (Fig. 1). In contrast, multiple genes evolve at elevated substitution rates in the obligate hemiparasites *S. americana* and *S. hermonthica*, mostly with significantly higher dN and dS in the majority of photosynthesis and housekeeping genes. Among holoparasites, dN and dS are highest in the *Epifagus virginiana*/*C. americana*/*Cistanche phelypaea*-clade (Fig. 1). Fewer genes evolve at elevated molecular evolutionary rates in *Myzorrhiza californica* and most of the *Orobanche* and *Phelipanche* species, which all retain genes for the plastid ATP synthase (*atp* genes) (Fig. 1). However, dN and/or dS of the retained *atp* genes are mostly accelerated in these holoparasites compared with those of nonparasites (Fig. 1). Despite some disproportional rate accelerations, both dN and dS are highly correlated [Mantel tests (12), all $P < 0.001$] without apparent lags (Fig. 2*A* and Fig. S3).

Changes of Selection. We assessed the direction and the strength of changes of selection across functional gene complexes by ω (ratio of dN to dS) and the selection strength parameter k (according to ref. 13) through branch-site random effects models (13, 14). Alternative branch partitioning and Akaike weights were used to infer the relative contribution of each major shift of lifestyle (i.e., nonparasitism to parasitism to obligate parasitism to holoparasitism and complete loss of photosynthesis) to changes of selection. Plastid genes in general show significant shifts of selection, especially in obligate parasitic Orobanchaceae compared with nonparasitic species and *T. versicolor* (Fig. 2*B*). In photosynthesis genes, selectional strength is significantly or—in case of *psb* genes if analyzed separately—marginally significantly reduced in obligate parasites (Fig. 2*C* and Table S1). Exceptions are *rbcL* and *pet* genes, which all show no significant change in selection ($k = 0.64$–0.99, LRT P value 0.460–0.879). Ribosomal genes for the small and large ribosome subunit (*rps*, *rpl*) show a relaxation of purifying selection in all parasites (i.e., including *T. versicolor*) compared with nonparasitic species. In contrast, genes for the RNA polymerase (*rpo*) evolve under lower selectional strength only in

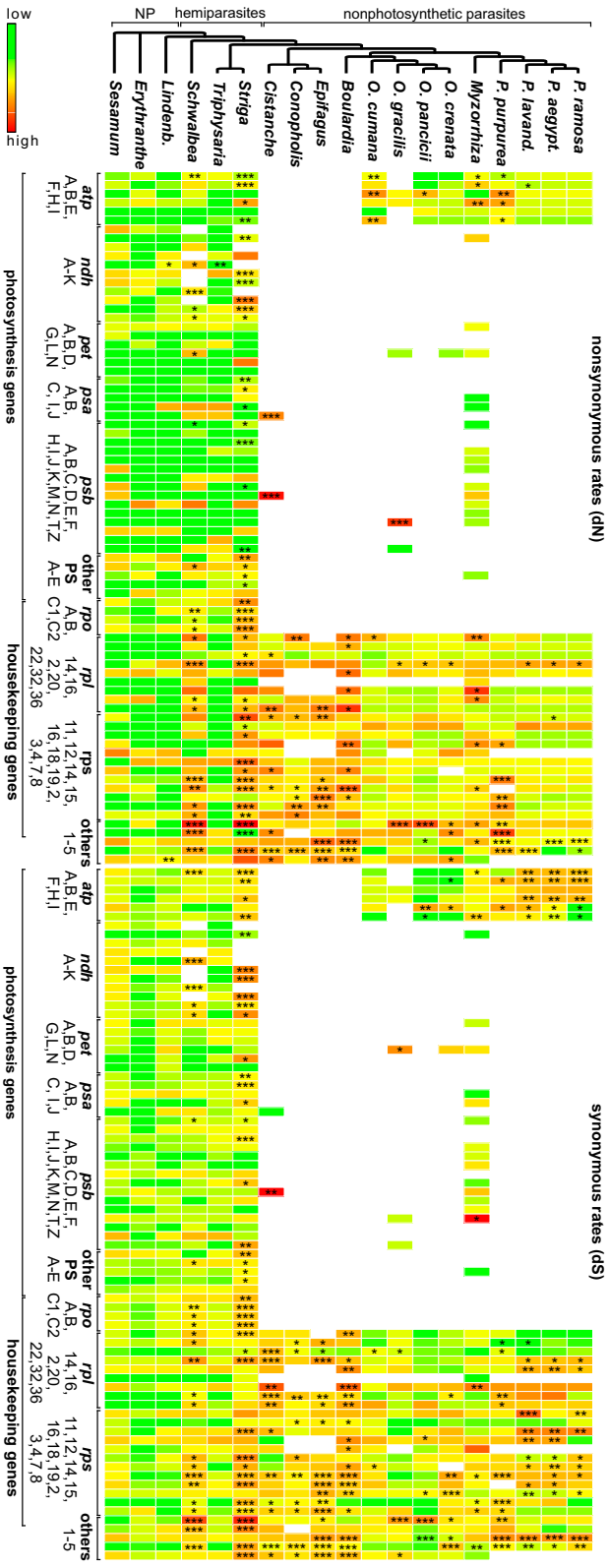

Fig. 1. Rate variation in Orobanchaceae. Heatmaps illustrate differences in dN and dS for each plastid protein gene (as named from top to bottom per gene class), with low rates shown in green and high rates in red. A phylogenetic tree on top indicates species relationships and the trophic specialization. Asterisks indicate the significance of LRTs of a focal taxon against the non-Orobanchaceae taxa (significance levels: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$). other PS, other photosynthesis genes (A, *ccsA*; B, *cemA*; C, *rbcL*; D, *ycf3*; E, *ycf4*), other HK, housekeeping and metabolic genes (1, *matK*; 2, *infA*; 3, *ycf2*; 4, *clpP*; 5, *accD*).

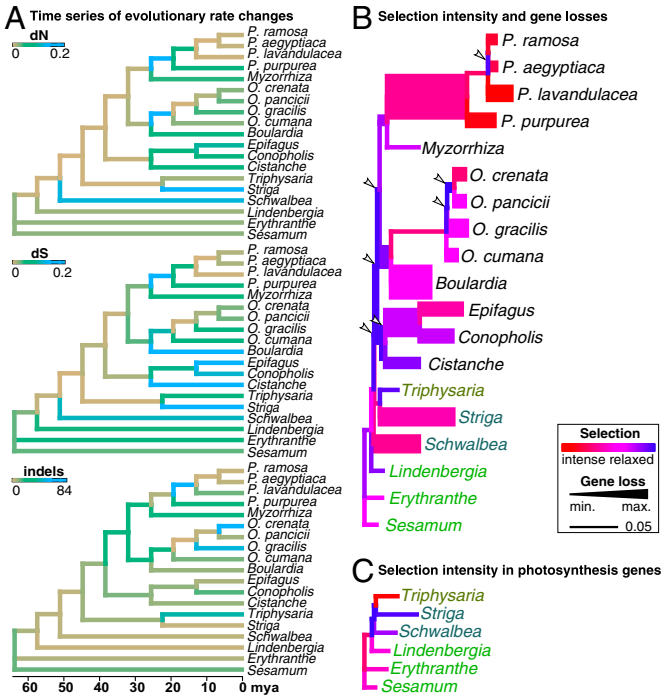

Fig. 2. Evolution of selectional strength. (*A*) Evolution of dN, dS, and indels shown on the dated phylogeny. Low rates of nucleotide substitutions and indels are shown in brown and high ones in blue. (*B*) Selectional changes per branch across all universal protein genes (as in Fig. 1) are color-coded according to the selection strength parameter *k*, inferred under the general descriptive RELAX model (13). Low *k* (<1, blue) indicates a relaxation of purifying selection (i.e., a weaker deviation from neutral evolution), whereas high *k* (>1, red) suggests selection intensification. Branch widths are proportional to the number of gene losses per branch. Arrowheads mark phases of selection relaxations occurring before inferred gene losses. (*C*) Selectional changes per branch across all photosynthesis genes (as in Fig. 1); colors as in *B*.

obligate parasites (Table 1). All other plastid genes (*accD*, *clpP*, *ycf2*) that are involved in other housekeeping or metabolic processes other than photosynthesis show a slight intensification of purifying selection compared with nonparasites (ω = 0.545 vs. 1.05, LRT P <0.001). Across all universal genes, selectional strength is intensified in obligate parasites (Fig. 2 *B* and *C* and Table 1), albeit not evenly. Whereas selection is more relaxed along the backbone (i.e., the selection parameter *k* is low), it is intensified toward terminal lineages (e.g., *S. hermonthica*, *E. virginiana*, *Phelipanche*) with intermittent phases of again relaxed selection within *Orobanche* and *Phelipanche* (Fig. 2 *B* and *C*). In housekeeping genes and a few photosynthesis genes (e.g., *ccsA*, *cemA*, *ycf3*, *ycf4*, but not *rbcL*), we found evidence for adaptive evolution in a small proportion of sites (Fig. S4).

The probabilities of functional complexes to have retained their function along the transition to holoparasitism, which we obtained by averaging over the probabilities of their component genes to be intact as estimated using maximum likelihood reconstructions in *BayesTraits* 2, have been subjected to phylogenetic principal component analysis (*SI Materials and Methods*). The first two principal components (PC1, PC2) account for 94.7% of the overall variance (Fig. S5). The plastid photosynthesis gene classes *ndh*, *pet*, *psa*, *psb*, other photosynthesis-associated genes (*ccsA*, *cemA*, *ycf3/4*), and the *rpo* genes for the plastid-encoded polymerase contribute strongly (>0.95) to the first component. This result indicates that PC1 is a measure of the putative functioning of complexes associated with light harvesting and electron transport, as well as their transcription and assembly. *RbcL* contributes to PC1 to a lesser extent (0.84), indicating that this gene covaries with photosynthesis function. *Atp* genes contribute mainly to PC2 (loading: 0.81), whereas *rpl*, *rps*, *rbcL*, and *accD* load with less than 0.33. We used the Bayesian information criterion (BIC) and Schwarz weights (SW), the analog to Akaike weights, to evaluate the evidence for an array of alternative phylogenetic regression models. We found that PC1 and PC2 both represent significant factors to explain selection pressures in Orobanchaceae. A model that considers multiway interactions between these two variables (SW 0.72) outperforms a model that incorporates PC1 and PC2 additively (SW 0.26) and models that considered only one component (SW < 0.1). These results suggest that the selection shifts in parasites are shaped mainly by the nonfunctionalization of photosynthesis complexes and factors closely associated with them.

Microstructural Changes. The occurrence of length mutations in plastid genes caused by short insertions or deletions (indels) varies strongly between the different gene classes, with intact photosynthesis genes usually accumulating fewer indels (<1 per gene) than housekeeping genes (Fig. S6). Indel rates in genes that are universally present in Orobanchaceae do not differ among nonparasites, photosynthetic parasites, and the nonphotosynthetic parasites *M. californica*, *Orobanche* spp., and *Phelipanche* spp., whereas the remaining holoparasites (*Boulardia latisquama*, *C. phelypaea*, *C. americana*, and *E. virginiana*) show more length mutations. *Phelipanche* spp. (1.5–1.67 indels per gene) show slightly more length mutations than *Orobanche* spp. (0.83–1 indel per gene) and *M. californica* (0.83 indel per gene), which has indel rates similar to those of photosynthetic plants (0.83 indel per gene). Retained photosynthesis genes of holoparasites often show unique indels. For example, *psbZ* in *O. gracilis*, *psaJ* and *psbJ* in *C. phelypaea*, and *ndhB*, *petA*, *psaC/I*, and *psbA/D/K* in *M. californica* show length mutations, but none of the photosynthetic taxa have indels in any of these genes. Indels are rare in *atp* genes of nonphotosynthetic parasites. Mapping substitution rates and microstructural changes onto the dated Orobanchaceae phylogeny suggests that indel accumulation precedes substitution rate changes in some holoparasites in both housekeeping genes and *atp* genes (Fig. 2*A*), whereas in other photosynthesis genes (lacking in holoparasites), this effect is less obvious (Fig. S3).

Lifestyle-Dependent Changes of Evolutionary Rates. The dependency of evolutionary rates (dN and dS separately and jointly) on lifestyle was tested by using models that fuse sequence and trait evolution (15), thus evaluating whether the transition to obligate parasitism contributes significantly to evolutionary rate changes. The overall best models for the total substitution rate, dN, and dS distinguish between nonparasites plus the facultative hemiparasite *T. versicolor* versus obligate parasites irrespective of photosynthetic capacity (LRTs against the respective null models that assumed no influence of lifestyle changes, all P < 0.001). Parametric trait bootstrapping (15), which tests whether the observed rate variation is associated with an analyzed trait significantly more often than with uncorrelated traits that evolve in a similar manner, supports that shifts in dN and dS, alone and jointly, are significantly associated with changes in lifestyle (all P < 0.001) (Fig. S7). These results indicate that changes of nucleotide substitution rates coincide with the establishment of obligate parasitism rather than with the loss of photosynthesis.

Genetic Factors Underlying the Substitution Process. We studied the association of evolutionary rates (dN, dS, and the total rate μ) and ω with genetic traits (genome rearrangements, plastome size, gene content, GC content, indels), lifestyle, and life history by using uniresponse and multiresponse generalized linear mixed models using Markov chain Monte Carlo methods (MCMC-GLMM) (16). Substitution rates (dN, dS) and ω are each highly correlated with genetic traits. Uniresponse MCMC-GLMM indicates that dN relates more strongly to genetic traits than dS. The best models (according to SW) suggest additive effects of indels, the number of gene losses, the plastome size, and the lifestyle to predict both dN and μ, whereas dS is affected by the number of rearrangements (rather than indels per se), gene losses, plastome size, and the lifestyle (Table 2). Life history is present in none of the top-ranked models. Unlike in the evolutionary rates models, the best ω model

Table 1. Selectional strength

Gene set	ω_{mean}	*k*	LR	*P* value
PS	R: 0.150, T: 0.225	0.72	27.9	<0.001
HK	R: 0.592, T: 0.559	6.71	75.7	<0.001
Others	R: 1.020, T: 0.961	1.69	23.0	<0.001
UG	R: 0.318, T: 0.678	5.63	66.5	<0.001

HK, housekeeping; *k*, selection intensity parameter (13); LR, likelihood ratio; Others, *accD*, *clpP*, *ycf2*; PS, photosynthesis; R, nonparasites + *T. versicolor*; T, obligate parasites; UG, universal genes.

requires only the indel rate as factor (SW 0.62), suggesting that indels tend to accumulate in regions with high ω in Orobanchaceae. This univariate model outperforms a bivariate model with additive effects of indels and lifestyle (SW 0.25).

Discussion

Changes in gene content and shifts of nucleotide substitution rates have been commonly thought to relate to the relaxation of selective constraints and the loss of photosynthesis in plastid-bearing lineages that have secondarily acquired a parasitic lifestyle. Here, we showed that obligate parasitism, characterized by a parasite's need for a host plant, and the loss of photosynthesis strongly affects the functional reduction and rate accelerations in plastid genomes of Orobanchaceae, whereas parasitism per se, that is the ability to tap into the vascular tissue of another plant, are of subordinate importance (Fig. 1 and Tables 1 and 2). Both plastid photosynthesis and housekeeping genes evolve at significantly elevated evolutionary rates (including elevated indel rates) not only in holoparasites but already in the obligate hemiparasitic *S. americana* and *S. hermonthica.* Their plastomes also show more genomic rearrangements than autotrophs and the facultative hemiparasites *T. versicolor* and *Bartsia inaequalis* (17). These results are further corroborated by gene expression data from aboveground tissue of *S. hermonthica*, which expresses nuclear genes for light harvesting and photosystems with lower abundance than *T. versicolor* (18). The relaxation of purifying selection in both photosynthesis and housekeeping genes (e.g., *rpl*, *rps*) (Fig. 2) accompanying the transition to an obligate parasitic lifestyle also occurs in other parasitic lineages such as the sandalwood order (Santalales), a large group of flowering plants that has evolved root and stem parasites. Here, the plastomes of facultative hemiparasitic *Ximenia americana* and *Osyris alba* are highly similar to the nonparasite *Heisteria concinna* regarding their gene contents and evolutionary rates, whereas the obligate parasites *Phoradendron leucarpum* and several *Viscum* species show gene losses and relaxed selection in photosynthesis genes (19).

Accelerated evolutionary rates in plastomes have often been associated with changes in life history, which is the shift from long to short generation times (20). We found no consistent effect of generation time on rate variation (Table 2), which might be due to a high variability in life span, because in parasitic plants, generation time may be determined by host quality (only on annual hosts the parasite must be annual itself) rather than by intrinsic features. The strong coevolution of dN and dS (Table 2) suggests a lineage effect via the actual process of mutation (neutral mutation rate hypothesis), which, among others, depends on species-specific differences in DNA replication and repair efficiencies (21). As the selective pressures on plastid function gradually decrease in parasites, the proteins for DNA processing and DNA maintenance may experience relaxations of selection, just like the plastid genome they replicate or repair.

Changes in selection do not occur in a monotonic fashion. Instead, phases of rate acceleration and relaxed selection that coincide with inferred bursts of gene loss (Fig. 2 *A* and *B*) are followed by phases of rate deceleration and intensified selection in the retained functional complexes (Fig. 2*B*), suggesting that the plastomes of parasites have evolved toward a new rate equilibrium; we propose that this is due to transcript and protein turnover rate-dependent substitution rate shifts. In photosynthetic lineages, the high demand for the photosynthesis-related machinery selects for low nucleotide and amino acid substitution rates to maximize rapid translation through optimized codon use maintained by low dN and dS, and to minimize the risk of unfavorable protein misfolding (20). Therefore, codon use and substitution rates differ notably between the different gene classes in plastomes of flowering plants (22), whereas in parasites, this distinctness diminishes (5). Here, selective pressure on high turnover in the plastid is reduced, because the parasite obtains at least parts of the required organic compounds from its host rather than synthesizing them. Because of the reduced need for a rapid assembly of the thylakoid photosynthesis machinery, purifying selection is relaxed not only in photosynthesis genes, but also in genes encoding the transcription and translation machinery, allowing for indels to accumulate and dN and dS to increase. This finding is in line with the relaxation of selection in the genes for the plastid-encoded polymerase (*rpo*) seen already in the photosynthetic but obligate parasites *S. americana*, *S. hermonthica* (Table S1), mistletoes (19), and the repeatedly observed overall higher evolutionary rates in parasites (Fig. 1). These changes may also involve adaptations in the plastid housekeeping apparatus (Fig. S4). The hypothesized turnover rate-dependent rate shifts will be attenuated in genes that are required over longer periods, such as the ATP synthase (*atp*) genes or RuBisCO (*rbcL*), possibly because they take over or continue to carry out alternative functions (23, 24). Therefore, nonphotosynthetic parasites such as *M. californica*, *Orobanche* (11), or some *Cuscuta* species (7, 25), which all retain intact genes for the ATP synthase despite the loss of other photosynthesis genes, also have lower base-level evolutionary rates. The eventual deletion of all dispensable regions may reconstitute the compactness of the plastid chromosome with its typically low amounts of nongenic and low-complexity DNA regions (1).

A Model of Plastome Evolution in Parasites. Based on our data and previous research (4–7, 9, 11, 19, 25–35), we here propose a model of plastid genome evolution under relaxed selective constraints (Fig. 3). This model is applicable to many other secondarily heterotrophic lineages within primarily phototrophic clades other than Orobanchaceae such as algae and mycoheterotrophic plants.

Parasitism relaxes constraints on the NADH complex that is essential for electron cycling around photosystem I under stress, rendering *ndh* genes the first, to our knowledge, to be functionally lost from the plastome (5, 9). More dramatic changes concur with the transition to obligate parasitism, which relieves photosynthesis and, concomitantly, on plastid housekeeping functions of functional

Table 2. Phylogenetic MCMC-GLMMs

Model	SWs*
μ ~ indels + gene loss + pt size + lifestyle	0.334
μ ~ GR + gene loss + pt size + lifestyle	0.270
μ ~ indels + gene loss + pt size	0.157
dS ~ GR + gene loss + pt size + lifestyle	0.255
dS ~ indels + gene loss + pt size + lifestyle	0.225
dS ~ indels + gene loss + pt size	0.142
dS ~ GR + pt size + lifestyle	0.092
dN ~ indels + gene loss + pt size + lifestyle	0.388
dN ~ GR + gene loss + pt size + lifestyle	0.276
dN ~ indels + gene loss + pt size	0.147
dN, dS ~ indels + gene loss + pt size	0.346
dN, dS ~ GR + pt size + lifestyle	0.213
dN, dS ~ indels + pt size + lifestyle	0.145
ω ~ indels	0.617
ω ~ indels + lifestyle	0.247

GR, genome rearrangements based on locally collinear blocks; μ, total rate, pt, plastome.
*Calculated separately over the set of models per response variable(s), and only models with a cumulative weight of 0.7 are shown.

constraints (losses 1 and 2 in Fig. 3). This first phase of selection relaxation is characterized by a steady increase of microstructural changes and the acceleration of dN and dS. Following this episode of selectional shift, evolutionary rates in the plastome evolve at a new equilibrium, perhaps matching the modified transcript and protein requirements (11). This molecular evolutionary regime shift is repeated once selective constraints on plastid proteins (e.g., ATP synthase) that continue to function for a longer period are lost or functionally replaced (functional loss 3 in Fig. 3). The relaxation of selective pressure on photosynthesis, alternative photosynthesis-unassociated functions, and on the plastid housekeeping machinery may be linked to the increasing specialization on external carbon (e.g., via distinct host systems with improved efficiency of nutrient acquisition), but the precise coevolutionary mechanisms remain unclear. The lifestyle-specific shifts of evolutionary rate regimes are accompanied by a reduced plastome GC content that may in part be due to relaxed constraints on codon use or nutrient economy (5, 36), although the changing GC content apparently has no direct influence on dN and dS (Table 2). However, low GC content correlates with increases in the amount of structural rearrangements (37) including the deletion of dispensable DNA (5), all factors that directly influence the substitution rates of plastid genes (Table 2) (11).

Reasons for the retention of minimal plastomes in most nonphotosynthetic plants investigated so far are varied. The proportion and nature of lineage-specifically retained genes and nonessential DNA potentially relate to inefficient or impossible protein import, highly reduced translational apparatus, regulatory coupling of genes for biological processes, coordinate assembly and cotranslation of partnered proteins, divergences in the genetic code (including modified start codons), or posttranscriptional editing that all represent barriers for functional gene transfer (38). Even if all protein genes are lost, the plastid-encoded L-glutamyl-tRNA (*trnE*) may still be required for tetrapyrrole biosynthesis in plants, which requires its functional transfer to the nuclear genome because the nuclear equivalent cannot replace the initiator function (38; but see ref. 35). However, a minimal plastome with only one tRNA that is nearly indifferent from the cytosolic tRNA species would be difficult to detect, even with high-throughput methods and in situ visualization techniques. Functional recompartmentalization of molecular biological processes and functional replacement of plastid-encoded genes through, for instance, functional intracellular gene transfer may, however, eventually relieve the organelle of the pressure to retain own genetic material (losses 4 and 5 in Fig. 3), potentially leading to plastids without plastomes [as in *Rafflesia* (32); *Polytomella* (33)].

Materials and Methods

Taxon Sampling and Plastome Sequencing. Using whole-genome shotgun sequencing (454 FLX and Illumina, paired-end), we reconstructed the plastomes of four holoparasites and two hemiparasites (Table S2) in addition to the Orobanchaceae already sequenced (5, 6, 11), using the same experimental and bioinformatic procedures. We queried combined builds of transcriptome data from *Phelipanche aegyptiaca* and *S. hermonthica* (39) to assist the annotation of plastid genes. Details are provided in *SI Materials and Methods*.

Analysis of Gene Content and Evolutionary Rates. Based on a matrix of all unique plastid genes and the established phylogenetic relationships (10, 11), we reconstructed gene losses at ancestral nodes by using BayesTraits 2 (40) under the multistate option and 500 maximum likelihood (ML) attempts. We calculated the probability of a functional complex to have retained its function at each node by averaging over the ML estimates of probabilities of its contributing genes to be intact (Fig. S1). Following automated, codon-wise alignments by using prank 0.14 (41), analyses of relative dN and dS, and of the total rate in all plastid protein genes were carried out in HyPhy 2.1–2.3

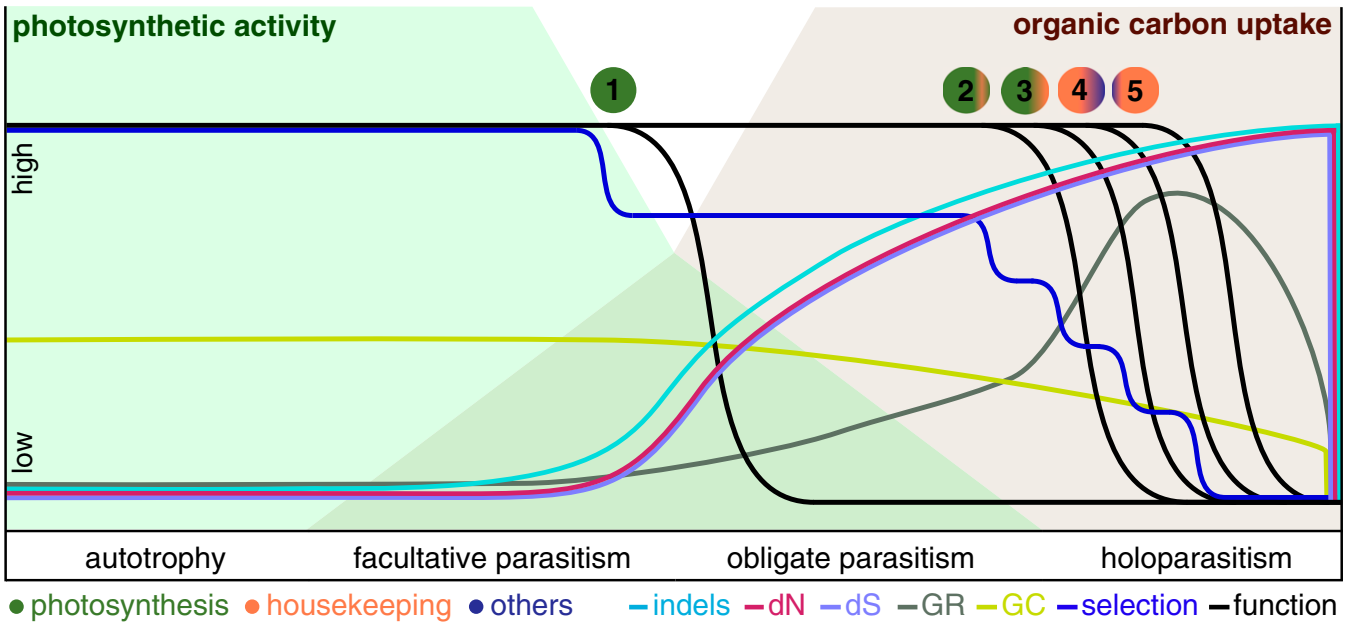

Fig. 3. General model of plastome evolution under relaxed functional constraints. The model illustrates the evolution of genomic traits (GC, GC content; GR, genomic changes including gene deletions), functionality of different plastid-encoded complexes (blue, other metabolic genes; green, photosynthesis genes; orange, housekeeping genes), evolutionary rates (dN, dS) as well as selection intensity and their associations during the transition from autotrophy with high photosynthetic activity (green background) and no uptake of external organic carbon to holoparasitism with maximal uptake (brown background). As the specialization on the heterotrophic lifestyle increases, more plastid-encoded complexes are lost, altering evolutionary rates and the genome structure. Numbers indicate the predicted functional reductions starting with the loss of nonessential or stress-relevant genes (*ndh* genes) (1) followed by primary photosynthesis genes (*pet* and *psa*/*psb* genes) and the plastid-encoded polymerase (PEP) (2), genes that have a prolonged function (e.g., *atp* genes, *rbcL*) and nonessential housekeeping genes (3), other metabolic genes (e.g., *accD, clpP, ycf1*/*2*) (4), and, finally, the remaining housekeeping genes, including *trnE* (5). Unlike most factors that are likely to show a rather monotonic trend as parasitism evolves, selectional strength may experience several periods of relaxation and intensification. The associations of evolutionary rates, genomic traits, and functional losses between 1 and 3 have been observed, whereas those after 3 are hypothetical. For simplicity, we here depict the functional reduction of a plastome with a coding capacity of flowering plants; more or fewer lineage-specific reconfigurations may occur in other lineages.

(42) using custom batch scripts and the MG94×GTR_3×4 codon model; *ycf1* was excluded because of uncertain homology assessment. Changes of ω and of the strength of selection measured by *k* were tested with a series of branch-site random effects likelihood methods (13, 14). For identifying the best lifestyle model per gene, different test branch sets were defined and evaluated by using Akaike weights. Details of all procedures are provided in *SI Materials and Methods*.

Analysis of Evolutionary Rates, Selection Pressures, and Genetic Factors. We used traitRate 1.1 (15) with 100 stochastic mappings and LRTs to test for lifestyle-dependent rate changes. To this end, we compared the log likelihoods of a null model that assumes no trait dependency versus a model that includes a trait parameter for each tested rate (total rate, dN, or dS) on the set of universally retained genes (Fig. 1 and Fig. S1) and performed parametric bootstrap analyses (as in ref. 15) of the best trait-rate models using 200 replicates. We measured rearrangements by locally collinear blocks (Fig. S2) (as in refs. 5 and 11). We distinguished lifestyle as nonparasite, facultative hemiparasite, obligate hemiparasite, generalist holoparasite, or specialized holoparasite, and life history as annual, biennial, and perennial (Table S2). Indels were coded with SeqState 1.4 (43) using SIC (44), and their occurrence was reconstructed over the Orobanchaceae tree by using the R packages ape (45) and phangorn (46). Gene losses per branch were calculated as the percentage of nonessential unique genes lost, where nonessential refers to a gene that has been lost in one or more of the study taxa. Phylogenetic MCMC-GLMMs (16) were computed by using unfixed and least-informative priors for both the variance-covariance matrices of the random effects and the residuals, and variance-corresponding hyperpriors. Factors were hierarchically reduced by significance until no better model fit was obtained according to BIC. We considered only models in the final set until the cumulative Schwarz weight per response exceeded 0.7. We used phylogenetic principal component (47) regression to model associations between selection pressure and the probability of the plastid-encoded fraction of a functional complex to have retained its function (from BayesTraits ML probabilities per complex). All components that together explained at least 90% of the variance were used as fixed effects in phylogenetic MCMC-GLMM analysis (as above). To evaluate the time series of genetic changes, we traced dN, dS, and indels on dated phylogenies, which we obtained by using penalized likelihood (48) implemented in ape (45), setting the root age boundary to 51–71 million years ago (11). Details of all coevolutionary analyses are provided in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank S. Renner (Munich) and T. Rattei (Vienna) for access to genome data of some holoparasites; and J. Naumann (Pennsylvania State University) and two anonymous reviewers for valuable comments on an earlier version of this manuscript. This work was supported by Austrian Science Fund FWF Grant 19404 (to G.M.S.); National Science Foundation Grants DBI-0701748 and DBI-1238057 (to C.W.d.); and the German Academic Exchange Service (S.W.).

1. Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D (2011) The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol Biol* 76(3-5):273–297.
2. Jansen RK, Ruhlman TA (2012) Plastid genomes of seed plants. *Genomics of Chloroplasts and Mitochondria*, eds Bock R, Knoop V (Springer, Dordrecht, The Netherlands), pp 103–126.
3. Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci USA* 84(24):9054–9058.
4. Young ND, dePamphilis CW (2005) Rate variation in parasitic plants: Correlated and uncorrelated patterns among plastid genes of different function. *BMC Evol Biol* 5(1):16.
5. Wicke S, et al. (2013) Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *Plant Cell* 25(10):3711–3725.
6. Wolfe KH, Morden CW, Palmer JD (1992) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. *Proc Natl Acad Sci USA* 89(22):10648–10652.
7. 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(1):45.
8. Lohan AJ, Wolfe KH (1998) A subset of conserved tRNA genes in plastid DNA of nongreen plants. *Genetics* 150(1):425–433.
9. Barrett CF, et al. (2014) Investigating the path of plastid genome degradation in an early-transitional clade of heterotrophic orchids, and implications for heterotrophic angiosperms. *Mol Biol Evol* 31(12):3095–3112.
10. McNeal JR, Bennett JR, Wolfe AD, Mathews S (2013) Phylogeny and origins of holoparasitism in Orobanchaceae. *Am J Bot* 100(5):971–983.
11. Cusimano N, Wicke S (2016) Massive intracellular gene transfer during plastid genome reduction in nongreen Orobanchaceae. *New Phytol* 210(2):680–693.
12. Lapointe F-J, Garland T, Jr (2001) A generalized permutation model for the analysis of cross-species data. *J Classif* 18(1):109–127.
13. Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K (2015) RELAX: Detecting relaxed selection in a phylogenetic framework. *Mol Biol Evol* 32(3):820–832.
14. Kosakovsky Pond SL, et al. (2011) A random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol* 28(11):3033–3043.
15. Mayrose I, Otto SP (2011) A likelihood method for detecting trait-dependent shifts in the rate of molecular evolution. *Mol Biol Evol* 28(1):759–770.
16. Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *J Stat Softw* 33(2):1–22.
17. Uribe-Convers S, Duke JR, Moore MJ, Tank DC (2014) A long PCR-based approach for DNA enrichment prior to next-generation sequencing for systematic studies. *Appl Plant Sci* 2(1):1300063.
18. Wickett NJ, et al. (2011) Transcriptomes of the parasitic plant family Orobanchaceae reveal surprising conservation of chlorophyll synthesis. *Curr Biol* 21(24):2098–2104.
19. Petersen G, Cuenca A, Seberg O (2015) Plastome evolution in hemiparasitic mistletoes. *Genome Biol Evol* 7(9):2520–2532.
20. Gaut B, Yang L, Takuno S, Eguiarte LE (2011) The patterns and causes of variation in plant nucleotide substitution rates. *Annu Rev Ecol Evol Syst* 42(1):245–266.
21. Moriyama T, Sato N (2014) Enzymes involved in organellar DNA replication in photosynthetic eukaryotes. *Front Plant Sci* 5(5):480.
22. Wicke S, Schneeweiss GM (2015) Next generation organellar genomics: potentials and pitfalls of high-throughput technologies for molecular evolutionary studies and plant systematics. *Next Generation Sequencing in Plant Systematics, Regnum Vegetabile*, eds Hörandl E, Appelhans M (Koeltz Scientific, Koenigstein, Germany), pp 9–50.
23. Kamikawa R, et al. (2015) Proposal of a twin arginine translocator system-mediated constraint against loss of ATP synthase genes from nonphotosynthetic plastid genomes. *Mol Biol Evol* 32(10):2598–2604.
24. Leebens-Mack J, dePamphilis C (2002) Power analysis of tests for loss of selective constraint in cave crayfish and nonphotosynthetic plant lineages. *Mol Biol Evol* 19(8):1292–1302.
25. McNeal JR, Kuehl JV, Boore JL, dePamphilis CW (2007) Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. *BMC Plant Biol* 7(1):57.
26. Delannoy E, Fujii S, Colas des Francs-Small C, Brundrett M, Small I (2011) Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. *Mol Biol Evol* 28(7):2077–2086.
27. Logacheva MD, Schelkunov MI, Nuraliev MS, Samigullin TH, Penin AA (2014) The plastid genome of mycoheterotrophic monocot *Petrosavia stellaris* exhibits both gene losses and multiple rearrangements. *Genome Biol Evol* 6(1):238–246.
28. Logacheva MD, Schelkunov MI, Penin AA (2011) Sequencing and analysis of plastid genome in mycoheterotrophic orchid *Neottia nidus-avis*. *Genome Biol Evol* 3:1296–1303.
29. Schelkunov MI, et al. (2015) Exploring the limits for reduction of plastid genomes: A case study of the mycoheterotrophic orchids *Epipogium aphyllum* and *Epipogium roseum*. *Genome Biol Evol* 7(4):1179–1191.
30. Lam VKY, Soto Gomez M, Graham SW (2015) The highly reduced plastome of mycoheterotrophic *Sciaphila* (Triuridaceae) is colinear with its green relatives and is under strong purifying selection. *Genome Biol Evol* 7(8):2220–2236.
31. Barrett CF, Davis JI (2012) The plastid genome of the mycoheterotrophic *Corallorhiza striata* (Orchidaceae) is in the relatively early stages of degradation. *Am J Bot* 99(9):1513–1523.
32. Molina J, et al. (2014) Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Mol Biol Evol* 31(4):793–803.
33. Smith DR, Lee RW (2014) A plastid without a genome: Evidence from the nonphotosynthetic green algal genus *Polytomella*. *Plant Physiol* 164(4):1812–1819.
34. Naumann J, et al. (2016) Detecting and characterizing the highly divergent plastid genome of the nonphotosynthetic parasitic plant *Hydnora visseri* (Hydnoraceae). *Genome Biol Evol* 8(2):345–363.
35. 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(1):189–201.
36. 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 tRNA and ribosomal protein genes. *J Mol Evol* 35(4):304–317.
37. Müller AE, et al. (1999) Palindromic sequences and A+T-rich DNA elements promote illegitimate recombination in *Nicotiana tabacum*. *J Mol Biol* 291(1):29–46.
38. Barbrook AC, Howe CJ, Purton S (2006) Why are plastid genomes retained in nonphotosynthetic organisms? *Trends Plant Sci* 11(2):101–108.
39. Yang Z, et al. (2015) Comparative transcriptome analyses reveal core parasitism genes and suggest gene duplication and repurposing as sources of structural novelty. *Mol Biol Evol* 32(3):767–790.
40. Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Syst Biol* 53(5):673–684.
41. Löytynoja A, Goldman N (2005) An algorithm for progressive multiple alignment of sequences with insertions. *Proc Natl Acad Sci USA* 102(30):10557–10562.
42. Pond SL, Frost SDW, Muse SV (2005) HyPhy: Hypothesis testing using phylogenies. *Bioinformatics* 21(5):676–679.
43. Müller K (2005) SeqState: Primer design and sequence statistics for phylogenetic DNA datasets. *Appl Bioinformatics* 4(1):65–69.
44. Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Syst Biol* 49(2):369–381.
45. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20(2):289–290.
46. Schliep KP (2011) phangorn: Phylogenetic analysis in R. *Bioinformatics* 27(4):592–593.
47. Revell LJ (2009) Size-correction and principal components for interspecific comparative studies. *Evolution* 63(12):3258–3268.
48. Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Mol Biol Evol* 19(1):101–109.