Plastome sequencing of South American Podocarpus species reveals low rearrangement rates despite ancient Gondwanan disjunctions

María Paula Quiroga (paulaquiroga@comahue-conicet.gob.ar)
INIBIOMA: Instituto de Investigaciones en Biodiversidad y Medioambiente https://orcid.org/0000-0001-8301-3998

E. Zattara
INIBIOMA: Instituto de Investigaciones en Biodiversidad y Medioambiente

Gustavo Souza
Federal University of Pernambuco: Universidade Federal de Pernambuco

Andrea Pedrosa-Harand
Universidade Federal Rural de Pernambuco

Andrea C. Premoli
INIBIOMA: Instituto de Investigaciones en Biodiversidad y Medioambiente

Research Article

Keywords: chloroplast genome, Gondwanan conifers, South America

Posted Date: April 28th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1547449/v1

Abstract

I. Background: Historical reconstructions within Podocarpaceae have provided valuable information to disentangle biogeographic scenarios that begun 65 Mya. However, early molecular phylogenies of Podocarpaceae failed to agree on the intergeneric relationships within the family. The aims of this study were to test whether plastome organization is stable within the genus *Podocarpus*, to estimate the selective regimes affecting plastome protein-coding genes, and to strengthen our understanding of the phylogenetic relationships and biogeographic history.

II. Methods and Results: We sequenced the plastomes of four South American species from Patagonia, southern Yungas, and Brazilian subtropical Amazon. We compared their plastomes to those published from Brazil, Africa, New Zealand, and Southeast Asia, along with representatives from other genera within Podocarpaceae as outgroups. The four newly sequenced plastomes ranged in size between 133,791 bp and 133,991 bp. Gene content and order among chloroplasts from South American, African and Asian *Podocarpus* were conserved and different from the plastome of *P. totara*, from New Zealand. Most genes showed substitution patterns consistent with a conservative selective regime. Phylogenies inferred from either complete sequences or protein coding regions were mostly congruent with previous studies, but showed earlier branching of *P. salignus*, *P. totara* and *P. sellowii*.

III. Conclusions: Highly similar and conserved plastomes of African, South American and Asian species suggest that *P. totara* plastome should be revised and compared to other species from Oceanic distribution. Furthermore, given such structural conservation, we suggest plastome sequencing is not useful to test whether genomic order can be climatically or geologically structured.

Introduction

The genus *Podocarpus* (yellowwoods or mañíos) is the largest genus in the Podocarpaceae family, both in diversity and in distribution. It is divided into two monophyletic subgenera, *Podocarpus* and *Foliolatus*. The subgenus *Podocarpus* occurs in all continents of Gondwanan origin (*i.e.*, South America, Africa, Oceania and Australia), while *Foliolatus* is found mainly in Asia [1]. Biogeographically, this distribution is the product of vicariance and dispersal events that took place during the last 65 Mya [2, 3]. Historical reconstructions within Podocarpaceae combining fossil information with phylogenetic and phylogeographic tools have provided valuable information to test and disentangle biogeographic scenarios. However, early molecular phylogenies of Podocarpaceae were built using a handful chloroplast genes and a small number of nuclear DNA regions [4–7] and failed to agree on the intergeneric relationships within the family [8]. More recent work using entire chloroplast genome sequences of species from 14 out of the 19 described genera of Podocarpaceae identified frequent rearrangements in gene order, mostly due to inversions and translocations; these changes are proposed to be driven by the presence of intermediate-size repeats located near the boundaries of colinear blocks [8, 9].

The chloroplast genome (or plastome) of terrestrial plants is approximately 120-220kb in size and consists of two copies of the inverted repeats (IRs) that separate the small and large single copy regions, SSC and LSC respectively [10, 11]. In gymnosperms, the plastome is highly variable in size and organization, and this variation is often used in phylogenetics to identify families, subgenera, and genera [12]. Yet, certain families remain poorly sampled such as only eight of the 187 species within Podocarpaceae have been sequenced so far [13]. Also, the loss of the large IR has been reported in several species, mainly conifers [14–16]. Furthermore, many rearrangements can be observed in these plastomes, some of which appear to have played an important role in their evolution [15–17]. A frequent change is the deletion of genes, as reported within the inverted repeat region in Cupressophyta (Araucariales and Cupressales) and Pinaceae [18]. In Cupressoidaea, Sudianto *et al.* [19] suggest that added complexity results from the existence of isomeric plastomes. As in other species of Podocarpaceae [16, 17], the plastome of *Retrophyllum piresii* lacks one of the IRs. Similarly, the genus *Lagostrobos*, having the largest plastome described so far for the family (151.5 kb), lacks the IR region but hosts numerous intergenic spacer repeats. Wu and Chaw [20] concluded that in cupressophytes, mutation rates have a critical role in driving the evolution of plastomic size, while plastomic inversions evolve in a neutral manner. As repetitions play an important role in plastomic rearrangements [8, 21, 22], plastome sequencing and *de novo* assembly can be considered an efficient tool to understand phylogenetic relationships at different taxonomic levels, as well as to investigate the structural and functional evolution of plants [23].

A comparison of plastomes of two *Podocarpus* species, *P. lambertii* and *P. totara*, found differences in internal gene order due to four large inversions of about 20,000 bp in length each [17]. Since *P. lambertii* occurs in Brazil, while *P. totara* occurs in New Zealand, and given the vicariant biogeographic history of the genus, it could be hypothesized that the large geographic distance between them correlates with genetic differences, reflected both as nucleotide substitutions and plastome rearrangements. Thus, geographic distance, evolutionary time and different adaptive traits could explain plastome structural differences found between these two congener species [24]. Alternatively, these differences could have resulted from changes in a single lineage leading to either species rather than reflecting a trend for the whole genus. By comparing plastomes from only two species, these two alternatives cannot be distinguished.

More than half of the species of *Podocarpus* occur in South America. They are distributed in four clades (Quiroga et al. 2016) that are geographically structured into southern, tropical, and subtropical distributions, the latter of which is sister to the African subtropical clade and thus a good system to test the above hypotheses. Therefore, we sequenced the plastomes of four additional species of South American *Podocarpus*: *P. nubigenus* and *P. salignus*, found along the Patagonian temperate forests, *P. parlatorei* found at the southern Yungas rainforest, and *P. sellowii*, mainly distributed in the Atlantic rainforest, but also present in other Brazilian biomes in subtropical forests. We compared their plastomes with those available for *P. lambertii* from Mata Atlantica (southern Brazil) [17], *P. latifolius* and *P. milanjianus* from Africa [25], *P. totara* from New Zealand [26], and *P. neriifolius* from Southeast Asia [27], along with representatives from other genera within Podocarpaceae as outgroups. The aim of this work

was to analyze stability (or lack thereof) of plastome organization within the genus *Podocarpus*, estimate the selective regimes affecting plastome protein-coding genes, and strengthen our current understanding of the phylogenetic relationships and biogeographic history of *Podocarpus*.

Methods

Plant leaf tissue was collected from individuals of four *Podocarpus* species within their natural distributions. These South American species studied *de novo* here were previously analyzed by means of chloroplast and nuclear sequences that yielded three clades (following [2]): an austral clade (*P. nubigenus* and *P. salignus*), a subtropical clade (*P. parlatorei*), and a tropical clade (*P. sellowii*). Total DNA from samples of *P. nubigenus, P. salignus* and *P. parlatorei* was extracted from 2 g of fresh leaves, using a modified ATMAB protocol (adapted by Doyle and Doyle [28]). For *P. sellowii*, genomic DNA was extracted using a modified CTAB (cetyltrimethylammonium bromide) protocol [28], adapted by Ferreira and Grattapaglia [29].

Sequencing

The high-quality DNA was sheared, ligated to specific adapters and sequenced on the Illumina HiSeq 2500/4000 platform, to generate between 2 and 4 Gb of sequence for each species in the form of 150 bp paired-end reads. Sequencing libraries, sequencing and adapter trimming was performed by BGI Genomics (https://www.bgi.com). Ten micrograms of *P. sellowii* genomic DNA was sequenced on Illumina HiSeq 2500 platform for generating 1 Gb of 250-bp paired-end reads at the Max Planck-Genome-Centre Cologne, Germany.

Assembly and annotation

Raw reads were quality trimmed using Trimmomatic 0.39 [30]. Trimmed reads were then assembled into chloroplast genomes using the Assembly by Reduced Complexity (ARC) approach [31] (https://ibest.github.io/ARC/): reads were mapped to the *Podocarpus latifolius* plastome sequence [32] using bowtie2 [33], and mapping reads were then assembled *de novo* using SPAdes [34]. Contigs from the ARC assembly were imported to Geneious 8.1.9 [35] and further assembled using this software's *de novo* assembler. Although the bowtie2 mapping reference was used only to retrieve cpDNA reads from the total gDNA library and not as a scaffold for assembly, we repeated the assembly pipeline using the plastome sequence of *P. totara* [26] to ensure that genome structure was not biased by the choice of the initial reference seed. Due to the presence of inverted repeats, joining some of the contigs was ambiguous: contigs were spliced into alternative assemblies, and the alternatives were tested first based on the frequency of reads spanning the different possible junctions, and later using PCR validation with *ad-hoc* designed primers (see below). The genomes were annotated using Geneious: each assembly was aligned to existing annotated genomes from other *Podocarpus* species, annotations for regions having over 75% similarity in sequence were automatically transferred to the new genomes, and annotations were manually curated through comparisons with annotated published plastomes from other species.

After running ARC using short reads as input, we used the Geneious assembler to further assemble the ARC output. In all cases, this resulted in a fraction of ARC contigs assembling into 2–5 larger contigs; non-assembling ARC contigs were discarded. This smaller set of contigs was manually assembled, and PCR experiments were used to test for the correct orientation of assembled contigs (see next paragraph), resulting in four final whole plastome assemblies. We used either *P. latifolius* or *P. totara* (which differ in their plastome arrangement) as initial references, and found no differences in the final result, showing that the initial mapping reference used by ARC did not bias the resulting assemblies.

PCR verification of assembly ambiguities

To test the alternative assemblies, we designed five primers targeted to conserved 20 bp regions near the inverted repeats flanking the junction points: 90F (TCGGGTGCTTATTTGGTGCA); 659F (CACTCCCCGTTGTTCCTCAA); 1515F (GCATTACGCCCAAAACGGTT); 955R (AATCCCTCCTGGCTCACAAC); 2767R (CGAACGAAATGCCAGTGACC). Using highly stringent PCR conditions and using genomic DNA from each species as a template, we tested different combinations of these primers, predicting amplification of fragments of specific sizes or lack of amplification for each alternative assembly.

Genome-wide alignment and structure analysis

We used progressive Mauve [36], as implemented in Geneious, to perform a genome-wide alignment and identify locally collinear blocks (LCBs) across our four new *Podocarpus* assemblies plus available plastomes from 13 species of Podocarpaceae (NCBI accession): *Saxegothaea conspicua* (AP018906), *Pherosphaera fitzgeraldii* (AP018903), *Dacrycarpus imbricatus* (NC_034942), *Dacrydium cupressinum* (AP018900), *Dacrydium elatum* (NC_045880), *Retrophyllum piresii* (NC_024827), *Afrocarpus gracilior* (AP018899), *Nageia nagi* (NC_023120), *Podocarpus latifolius* (NC_042224), *Podocarpus lambertii* (NC_023805), *Podocarpus milanjianus* (MT019686), *Podocarpus neriifolius* (MN882702), and *Podocarpus totara* (NC_020361). Most species analyzed here, except outgroups, belong to the *Podocarpus* subgenus; only *Podocarpus neriifolius* corresponds to subgenus *Foliolatus* which is found in Asia and Oceania.

Sequence alignment and phylogenetic tree reconstruction

Whole genome alignment yielded a series of locally collinear blocks (LCBs), each consisting of a multiple sequence alignment (MSA) of a portion of the plastome with conserved synteny across all species. We concatenated MSAs for all LCBs, and used this dataset to infer phylogenetic relationships among all 17 species. We used a maximum likelihood approach and a Bayesian approach. Maximum likelihood inference was performed using RAxML 7.2.8 [37], using a general time reversible model with gamma-distributed rates plus a proportion of invariant sites and a rapid hill-climbing algorithm. Bayesian inference was performed using MrBayes 3.2.6 [38], using the model above, and running four heated chains for 1,100,000 iterations. Since alignment of non-coding sequences is often ambiguous and could potentially bias inference, we also extracted all protein-coding genes from MSAs, concatenated them and ran the same analyses described above on this dataset.

Substitution rate

We estimated a single ratio of nonsynonymous to synonymous substitution rates ω = dN/ dS for each protein-coding gene. This ratio measures the strength and mode of natural selection acting on protein genes, with ω > 1 indicating positive (adaptive or diversifying) selection, ω = 1 indicating neutral evolution, and ω < 1 indicating negative (purifying) selection [39]. We used CODEML from the PAML package [40], specifying NSites = 0 and model = 0. In order to detect a change in selective regime within the genus *Podocarpus* for any genes, we ran the analysis for all 17 species of Podocarpaceae, and then repeated it only for the 9 species of *Podocarpus.* The ratio ω summarizes gene evolution rates, and can be an informative feature, because it can identify which genes are the most (or least) conserved and also identify genes that may have undergone periods of adaptive evolution.

Results

Genome assembly and content

The circular plastomes of the four species ranged in size between 133,791 bp and 133,991 bp (Table 1, Fig. 1), well within the range of the previously sequenced *Podocarpus* plastomes. The %GC ranged between 37.0 and 37.2% (Table 1). The gene content was mostly similar (Table 1): one of the two *trnD-GUC* genes in *P. sellowii* could have suffered the deletion of 59 out of 74 bases towards the 3' end due to 112 bases missing (although low quality base calls flanking them hints that this loss might be an assembly artifact).

Table 1
Chloroplast genome features of *Podocarpus nubigenus, P. salignus P. sellowii*, and *P. parlatorei*, and other genomes of *Podocarpus* species previously available.

Species	size (bp)	%GC	genes	tRNA	rRNA	CDS	Reference
P. nubigenus	133,862	37.2	119	33	4	82	Present work
P. salignus	133,902	37.1	119	33	4	82	Present work
P. sellowii	133,791	37.0	118	32	4	82	Present work
P. parlatorei	133,991	37.2	119	33	4	82	Present work
P. latifolius / P. milanjianus	134020	37.2	117	31	4	82	[32] / [25]
P. lambertii	133734	37.1	118	31	4		[17]
P. neriifolius	133878	37.1	119	33	4	82	[27]
P. totara	133259						[26]

We found that the four species were mostly syntenic, with no evidence of major rearrangements (Fig. 2). Average plastome-wide pairwise sequence identity was 96.8% (min: 96.3%; max 97.5%). They were also syntenic to the available plastomes of *P. neriifolius, P. milanjianus, P. latifolius* and *P. lambertii* within the *Podocarpus* genus, and with those of *Nageia nagi, Afrocarpus gracilior* and *Dacrydium elatum*. Across the *Podocarpus* genus, average plastome-wide pairwise sequence identity was 96.3% (min: 93.2% between *P. neriifolius* and *P. sellowii/P. nubigenus*; max 97.5% between *P. parlatorei* and *P. lambertii*). Sequence identity between the four *Podocarpus* species and *Retrophyllum, Afrocarpus* and *Nageia* ranged between 91.4% and 92.5%, while identity between them and *Dacrycarpus* and *Dacrydium* ranged between 86.1% and 88.4%. They were on average 81.4% identical to *Phaerosphera*, and 78.6% identical to *Saxegothea*.

Phylogenetic analysis

Runs of RAxML and MrBayes yielded the same best tree topology, whether from the whole plastome alignment or from concatenated coding regions (Fig. 2). This topology recovered the same relationships among Podocarpaceae genera found by a previous report [8]. Within the genus *Podocarpus*, the Southeast Asian species *P. neriifolius* of *Foliolatus* subgenus branches off first, the South American *P. salignus* branches off next, followed by New Zealand's *P. totara* and the South American *P. sellowii* and *P. nubigenus*. Two clades separate last, both comprising subtropical species: the African *P. milanjianus* and *P. latifolius*, and the South American *P. lambertii* and *P. parlatorei*.

Reconstruction of plastome rearrangements

Comparing the plastome structure of all species to their phylogenetic relationships shows that the arrangement found in all *Podocarpus* species, except for *P. totara*, is conserved in 11 out of the 17 species analyzed (Fig. 2), differing from the arrangement found in the outgroup and monotypic South American species *Saxegothaea conspicua* by the inversion of a single collinear block in the latter (Fig. 2, LCB 6). A reconstruction of the minimum events of translocation and inversion suggests that the plastome arrangement seen within most *Podocarpus* species is likely ancestral for Podocarpaceae. From this hypothetical ancestral arrangement, we reconstruct an inversion event affecting LCB6 in the lineage leading to *Saxegothaea conspicua*; an inversion and event affecting LCBs 9 and 10, followed by translocation of LCB 9 in the lineage leading to *Pherosphera fitzgeraldii*; a single inversion of a large block comprising LCBs 3–7 in the ancestral lineage including *Dacrycarpus* and *Dacrydium*, followed by an inversion of the same block in the *Dacrydium elatum* lineage that effectively reverted to the ancestral arrangement; a single inversion event affecting LCBs 9 to 2 in the *Retrophyllum piresii* lineage; and an inversion of a large block comprising LCBs 3–7, followed by a translocation of LCB 4 to become located between LCB 6 and LCB 7 in *P. totara*.

Synonymous and non-synonymous substitution rates

We estimated a single rate, single site ω (model M0 in PAML, see Methods) for each protein coding gene across the 17 species of Podocarpaceae. The distribution of ω values shows that most genes have a ω

ratio of less than 0.5 (Fig. 3A, upper half), and that in all cases dN/dS ratios are below 1 (Fig. 3B). This implies that non-synonymous mutations are being selected against, as expected from the overall conservative chloroplast genomes. Notice, however, that some genes have much larger tree lengths, driven by faster dS rates (*psbT*), faster dN rates (*matK, ycf1* and *ycf2*), or both (*clpP*).

We estimated the single rate, single site ω value again, but now using as input a subset comprising only the nine *Podocarpus* species, to detect any change in selective regime compared to the family background rates. Overall, rates are similar, although there might be a slight skew towards higher ω values within the genus (Fig. 3A, lower half). However, we detected some genes showing noticeably more relaxed (i.e., ω closer to 1 and thus under less intense negative selection) dN/dS ratios within the genus, like *psbK* and *rps7*, while a few others showed stronger conservation, like *psbM* and *psaM* (Fig. 3C). A single gene, *rps2* (*30S ribosomal protein S2, chloroplastic*), seems to have switched from being on average under negative selection across the family to being under positive selection within the genus.

Discussion

Previous work has revealed considerable diversity in chloroplast genome organization in gymnosperms in general [12], and within Podocarpaceae in particular [8]. *Podocarpus* is the most diverse genus within the family Podocarpaceae, and a comparison of the plastomes of the first two sequenced species, *P. lambertii* and *P. totara*, suggested that plastome instability might also be a general property at the intrageneric level [17]. After sequencing and analyzing the plastomes of four South American species, plus three recently sequenced species from Africa and Southeast Asia, we found that all seven species shared the same organization with *P. lambertii*. The fact that no differences were found in the internal genome arrangement of the South American and African species reflects a shared biogeographic past within the *Podocarpus* lineage which in turn shows a highly conserved evolutionary history. The differential arrangement of *P. totara* within the lineage of the genus *Podocarpus* is striking, since even *P. neriifolius*, that belongs to the subgenus *Foliolatus*, has the same arrangement as other species within subgenus *Podocarpus*. One possibility would be a misassembly; however, the *P. totara* plastome assembly was performed by skilled researchers, using gap closing and verification through PCR (Peter Lockhart, pers. comm.). Thus, it is possible that, even if plastome rearrangements are not common within the genus, they are likely enough to have happened in at least the one lineage leading to *P. totara*. Therefore, other species with an Oceanic distribution should be considered in future analyses. If plastome rearrangements are confirmed, this could open *Podocarpus* as a potential system to study the underlying mechanisms.

We observed full synteny among plastomes of *P. nubigenus*, *P. salignus*, *P. parlatorei*, *P. sellowii*, *P. lambertii*, and *P. latifolius/P. milanjianus*, all of these being significantly different from that of *P. totara*, which occurs in New Zealand. A biogeographic connection has been suggested between subtropical Africa and South America, probably related to geographic and paleoclimatic belts at such latitudes [2]. Interestingly, the orderings are similar also with *P. neriifolius*, a species that belongs to the subgenus *Foliolatus* and has an Asian distribution. Thus, plastome conservation has maintained synteny

throughout the tectonic history of the group. Given this degree of conservation, plastome sequencing is not useful to test whether genomic order can be climatically or geologically structured.

Our analysis of family- and genus-wide substitution rates confirms the overall conservative evolutionary rates of plastomic protein-coding genes and agrees with previous estimates [8]. Interestingly, a single gene, *rps2* seems to have changed selective regime in the *Podocarpus* lineage. This gene, encoding the S2 subunit of the plastid ribosome, has also been reported to switch regimes in other unrelated lineages, like *Vicia* (Fabaceae), *Arabidopsis* (Brassicaceae) or parasitic Scrophulariaceae [41–43]. Considering that, at least for *Arabidopsis*, allelic variants of this gene are linked to variance in disease resistance and susceptibility, it is tempting to speculate on the selective pressures that might have affected *Podocarpus* lineages throughout its biogeographic history. A more targeted study of *rps2* sequence variability, both within- and between species, would be however needed before ascribing a differential role for this gene in *Podocarpus* evolution.

The phylogenetic trees including all plastid genes inferred using RAxML and MrBayes yield an arrangement quite consistent with the topology that was previously published for *Podocarpus* [2, 44], despite the lower number of taxa included here. The Asian species *P. neriifolius* showed a sister group relationship to the rest of the *Podocarpus* clade, as expected given it is part of the subgenus *Foliolatus* [2]. Similarly, the warm-temperate South American *P. salignus* showed a sister group relationship to a clade containing the austral *P. totara* from New Zealand, the South American tropical *P. sellowii* and austral *P. nubigenus*, and a subtropical clade containing South American *P. lambertii* and *P. parlatorei*, and its sister group the African *P. milanjianus* and *P. latifolius*. This subtropical clade containing taxa currently geographically disjunct was interpreted as an indication of a past Atlantic subtropical biogeographical corridor [2].

Declarations

Author contributions

M. Paula Quiroga: Conceptualization, Investigation, Validation, Resources, Writing – Original draft, Writing – Review & Editing; **Eduardo E. Zattara**: Data curation, Formal analysis, Funding Acquisition, Resources, Visualization, Writing – Original draft, Writing – Review & Editing; **G. Souza:** Investigation, Plant collection, Formal analysis, Writing – Review & Editing; **A. Pedrosa-Harand**: Investigation, Resources, Data curation, Formal analysis, Funding Acquisition, Writing – Review & Editing; **Andrea C. Premoli**: Conceptualization, Funding acquisition PICT 2015-1563. MPQ, EEZ and ACP are CONICET researchers.

Acknowledgments

We thank Tiago Esposito, from the Federal University of Pernambuco, for collecting the *P. sellowii* individual and its DNA extraction (Coordenação de Aperfeiçoamento de Pessoal de Nível – CAPES, Financial code 001), and Bruno Huettel, from the Max Planck-Genome-Centre Cologne, Germany, for sequencing *P. sellowii* genomic DNA. We thank the Fundação de Amparo à Ciência e Tecnologia do

Estado – FACEPE (grant APQ-0349-2.02/15) for partial financial support. Financial support of Agencia Nacional de Promoción Científica y Técnica PICT 2015-1563.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Mill RR (2014) A monographic revision of the genus *Podocarpus* (Podocarpaceae): I. Historical review. Edinb J Bot 71:309–360. https://doi.org/10.1017/S0960428614000146
2. Quiroga MP, Mathiasen P, Iglesias A et al (2016) Molecular and fossil evidence disentangle the biogeographical history of Podocarpus, a key genus in plant geography. J Biogeogr 43:372–383. https://doi.org/10.1111/jbi.12630
3. Wu J-Y, Chen H, Ruan S-C et al (2021) Fossil leaves of Podocarpus subgenus Foliolatus (Podocarpaceae) from the Pliocene of southwestern China and biogeographic history of Podocarpus. Rev Palaeobot Palynol 287:104380. https://doi.org/10.1016/j.revpalbo.2021.104380
4. Conran JG, Wood GM, Martin PG et al (2000) Generic relationships within and between the gymnosperm families Podocarpaceae and Phyllocladaceae based on an analysis of the chloroplast gene rbcL. Aust J Bot 48:715–724. https://doi.org/10.1071/bt99062
5. Sinclair WT, Mill RR, Gardner MF et al (2002) Evolutionary relationships of the New Caledonian heterotrophic conifer, Parasitaxus usta (Podocarpaceae), inferred from chloroplast trnL-F intron/spacer and nuclear rDNA ITS2 sequences. Plant Syst Evol 233:79–104. https://doi.org/10.1007/s00606-002-0199-8
6. Biffin E, Conran JG, Lowe AJ (2011) Podocarp Evolution: A Molecular Phylogenetic Perspective. Ecology of the Podocarpaceae in tropical forests. Smithsonian Institution Scholarly Press, Washington DC, pp 1–20
7. Knopf P, Schulz C, Little DP et al (2012) Relationships within Podocarpaceae based on DNA sequence, anatomical, morphological, and biogeographical data. Cladistics 28:271–299. https://doi.org/10.1111/j.1096-0031.2011.00381.x
8. Sudianto E, Wu C-S, Leonhard L et al (2019) Enlarged and highly repetitive plastome of Lagarostrobos and plastid phylogenomics of Podocarpaceae. Mol Phylogenet Evol 133:24–32. https://doi.org/10.1016/j.ympev.2018.12.012
9. Qu X-J, Fan S-J, Wicke S, Yi T-S (2019) Plastome Reduction in the Only Parasitic Gymnosperm Parasitaxus Is Due to Losses of Photosynthesis but Not Housekeeping Genes and Apparently Involves the Secondary Gain of a Large Inverted Repeat. Genome Biol Evol 11:2789–2796. https://doi.org/10.1093/gbe/evz187

10. Palmer JD (1983) Chloroplast DNA exists in two orientations. Nature 301:92–93. https://doi.org/10.1038/301092a0
11. Knox EB (2014) The dynamic history of plastid genomes in the Campanulaceae sensu lato is unique among angiosperms. Proceedings of the National Academy of Sciences 111:11097–11102. https://doi.org/10.1073/pnas.1403363111
12. Chaw S-M, Wu C-S, Sudianto E (2018) Chapter Seven - Evolution of Gymnosperm Plastid Genomes. In: Chaw S-M, Jansen RK (eds) Advances in Botanical Research. Academic Press, pp 195–222
13. Lubna AS, Khan AL et al (2021) The dynamic history of gymnosperm plastomes: Insights from structural characterization, comparative analysis, phylogenomics, and time divergence. The Plant Genome 14:e20130. https://doi.org/10.1002/tpg2.20130
14. Hirao T, Watanabe A, Kurita M et al (2008) Complete nucleotide sequence of the Cryptomeria japonicaD. Don. chloroplast genome and comparative chloroplast genomics: diversified genomic structure of coniferous species. BMC Plant Biol 8:70. https://doi.org/10.1186/1471-2229-8-70
15. Yi X, Gao L, Wang B et al (2013) The Complete Chloroplast Genome Sequence of Cephalotaxus oliveri (Cephalotaxaceae): Evolutionary Comparison of Cephalotaxus Chloroplast DNAs and Insights into the Loss of Inverted Repeat Copies in Gymnosperms. Genome Biol Evol 5:688–698. https://doi.org/10.1093/gbe/evt042
16. Wu C-S, Chaw S-M (2014) Highly rearranged and size-variable chloroplast genomes in conifers II clade (cupressophytes): evolution towards shorter intergenic spacers. Plant Biotechnol J 12:344–353. https://doi.org/10.1111/pbi.12141
17. Vieira L, Faoro N, Rogalski H et al (2014) The Complete Chloroplast Genome Sequence of Podocarpus lambertii: Genome Structure, Evolutionary Aspects, Gene Content and SSR Detection. PLoS ONE 9:e90618. https://doi.org/10.1371/journal.pone.0090618
18. Mower JP, Vickrey TL (2018) Structural Diversity Among Plastid Genomes of Land Plants. Adv Bot Res 85:263–292. https://doi.org/10.1016/bs.abr.2017.11.013
19. Qu X-J, Wu C-S, Chaw S-M, Yi T-S (2017) Insights into the Existence of Isomeric Plastomes in Cupressoideae (Cupressaceae). Genome Biol Evol 9:1110–1119. https://doi.org/10.1093/gbe/evx071
20. Wu C-S, Chaw S-M (2016) Large-Scale Comparative Analysis Reveals the Mechanisms Driving Plastomic Compaction, Reduction, and Inversions in Conifers II (Cupressophytes). Genome Biology and Evolution 8:3740–3750. https://doi.org/10.1093/gbe/evw278
21. Weng M-L, Blazier JC, Govindu M, Jansen RK (2014) Reconstruction of the Ancestral Plastid Genome in Geraniaceae Reveals a Correlation between Genome Rearrangements, Repeats, and Nucleotide Substitution Rates. Mol Biol Evol 31:645–659. https://doi.org/10.1093/molbev/mst257
22. Sveinsson S, Cronk Q (2014) Evolutionary origin of highly repetitive plastid genomes within the clover genus (Trifolium). BMC Evol Biol 14:228. https://doi.org/10.1186/s12862-014-0228-6
23. Vieira L, Rogalski N, Faoro M H, et al (2016) The plastome sequence of the endemic Amazonian conifer, Retrophyllum piresii (Silba) C.N.Page, reveals different recombination events and plastome

isoforms. Tree Genetics & Genomes 12:10. https://doi.org/10.1007/s11295-016-0968-0

24. Xu X, Wang D (2021) Comparative Chloroplast Genomics of Corydalis Species (Papaveraceae): Evolutionary Perspectives on Their Unusual Large Scale Rearrangements.Frontiers in Plant Science11
25. Migliore J, Lézine A-M, Hardy OJ (2020) The recent colonization history of the most widespread Podocarpus tree species in Afromontane forests. Ann Bot 126:73–83. https://doi.org/10.1093/aob/mcaa049
26. Marshall CW (2013) Genetic diversity and relationships of New Zealand totara (Podocarpus totara): a thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Genetics at Massey University, Manawatu, New Zealand. Thesis, Massey University
27. Xie C, Liu D, Nan C (2020) The complete chloroplast genome sequence of Podocarpus neriifolius (Podocarpaceae). Mitochondrial DNA Part B 5:1962–1963. https://doi.org/10.1080/23802359.2020.1756478
28. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bull 19:11–15
29. Ferreira M, Grattapaglia D (1996) Introdução ao uso de marcadores moleculares em análise genética, 2nd edn. Centro Nacional de Pesquisa de Recursos Geneticos e Biotecnologia, Brasilia, DF
30. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. https://doi.org/10.1093/bioinformatics/btu170. Bioinformatics btu170
31. Hunter SS, Lyon RT, Sarver BAJ et al (2015) Assembly by Reduced Complexity (ARC): a hybrid approach. for targeted assembly of homologous sequences
32. Saina JK, Li Z-Z, Gichira AW et al (2019) The complete plastome of real yellow wood (Podocarpus latifolius): gene organization and comparison with related species. Holzforschung 73:525–536. https://doi.org/10.1515/hf-2018-0155
33. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Meth 9:357–359. https://doi.org/10.1038/nmeth.1923
34. Bankevich A, Nurk S, Antipov D et al (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021
35. Kearse M, Moir R, Wilson A et al (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199
36. Darling AE, Mau B, Perna NT (2010) progressiveMauve: Multiple Genome Alignment with Gene Gain, Loss and Rearrangement. PLoS ONE 5:e11147. https://doi.org/10.1371/journal.pone.0011147
37. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033

38. Ronquist F, Teslenko M, van der Mark P et al (2012) MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Syst Biol 61:539–542. https://doi.org/10.1093/sysbio/sys029
39. Jeffares DC, Tomiczek B, Sojo V, dos Reis M (2015) A Beginners Guide to Estimating the Non-synonymous to Synonymous Rate Ratio of all Protein-Coding Genes in a Genome. In: Peacock C (ed) Parasite Genomics Protocols. Springer, New York, NY, pp 65–90
40. Yang Z (2007) PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol Biol Evol 24:1586–1591. https://doi.org/10.1093/molbev/msm088
41. Li C, Zhao Y, Xu Z et al (2020) Initial Characterization of the Chloroplast Genome of Vicia sepium, an Important Wild Resource Plant, and Related Inferences About Its Evolution.Frontiers in Genetics11
42. Mauricio R, Stahl EA, Korves T et al (2003) Natural Selection for Polymorphism in the Disease Resistance Gene Rps2 of Arabidopsis thaliana. Genetics 163:735–746. https://doi.org/10.1093/genetics/163.2.735
43. 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. PNAS 94:7367–7372. https://doi.org/10.1073/pnas.94.14.7367
44. Leslie AB, Beaulieu JM, Rai HS et al (2012) Hemisphere-scale differences in conifer evolutionary dynamics. Proceedings of the National Academy of Sciences 109:16217–16221. https://doi.org/10.1073/pnas.1213621109

Figures

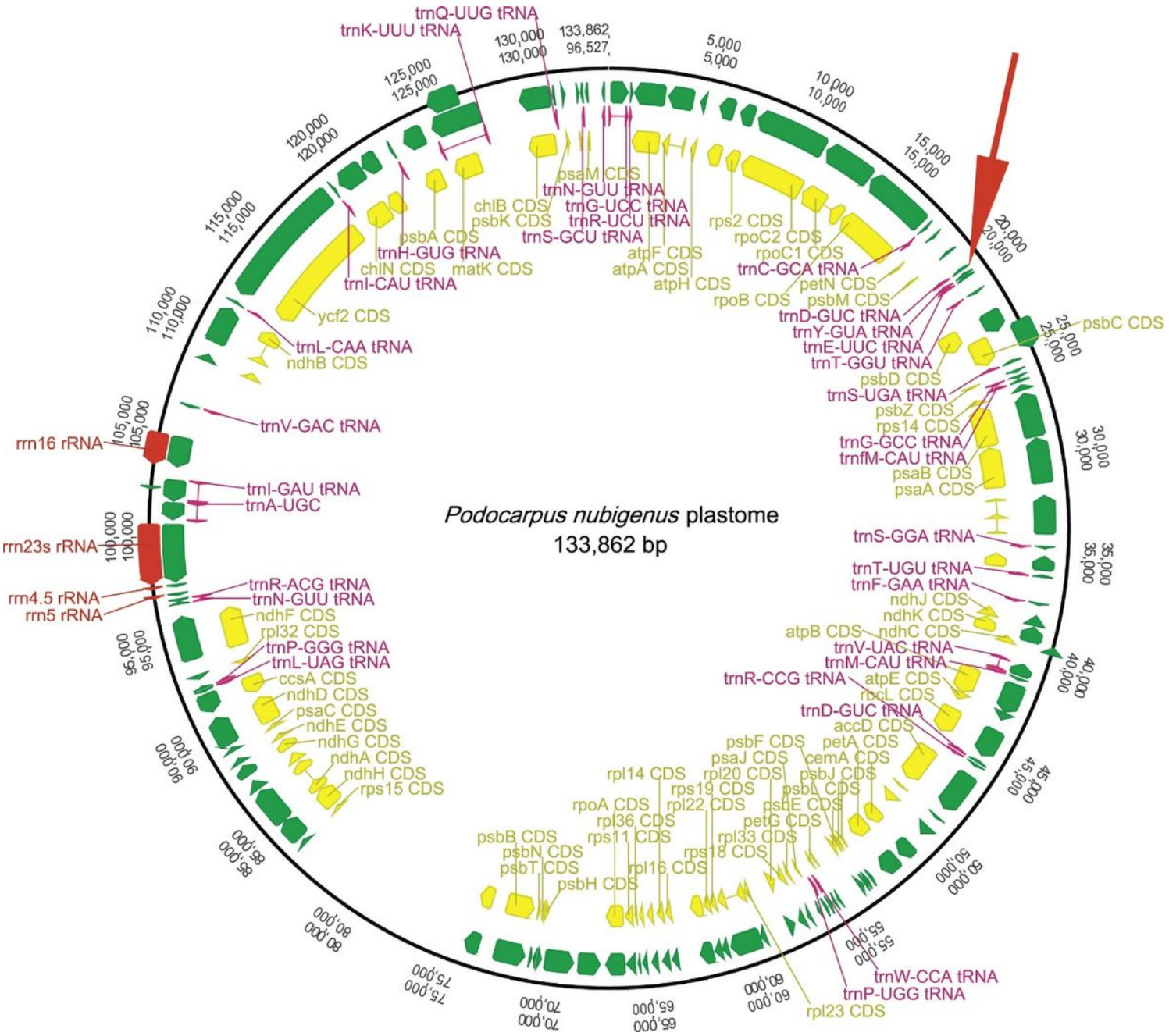

Figure 1

Genomic structure of the circular plastome of *Podocarpus nubigenus*. *Podocarpus salignus, P. sellowii,* and *P. parlatorei* share this same general structure, except for the partial deletion of one of the trnD-GUC genes in *P. sellowii* (red arrow).

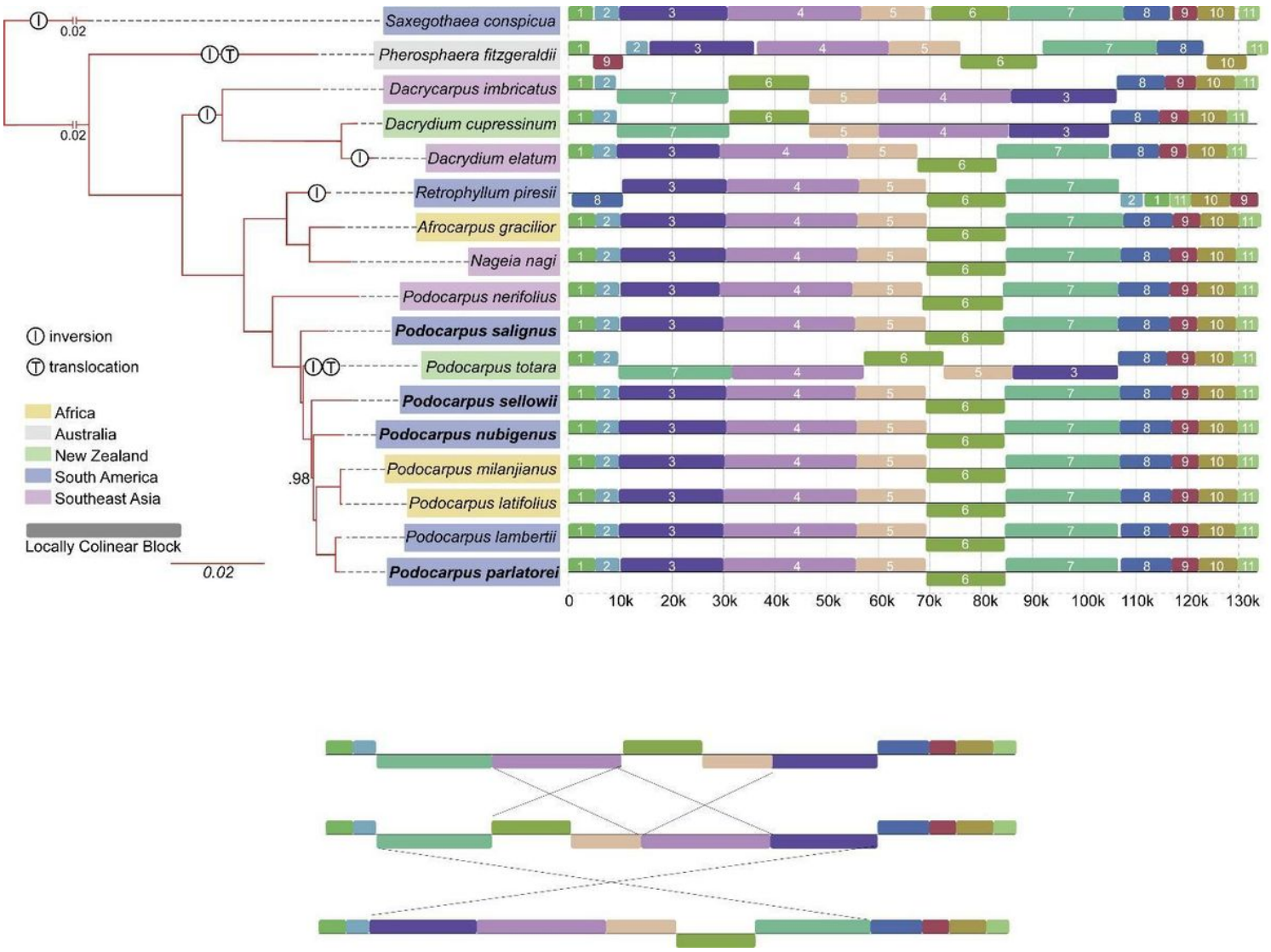

Figure 2

Phylogeny and plastome rearrangements among 17 species of Podocarpaceae. Phylogenetic relationships on the left represent MrBayes output as inferred from whole plastome alignment; posterior probabilities for all branches were >0.99 except where indicated. The pair of vertical bars breaking the first two branches represent an abridged distance of 0.02 substitutions/site. Inferred inversions (I) and translocations (T) indicated within circles in branches. The continental distribution of each species is shown as colored cartridges around species names. Species in bold were sequenced in the present work. The block diagram on the right represents the arrangement of locally colinear blocks (LCBs), as reconstructed by the progressive Mauve package. The horizontal axis represents position along the plastome, with position 0 located arbitrarily at the end of the *psbA* gene (whose CDS is in the reverse strand).

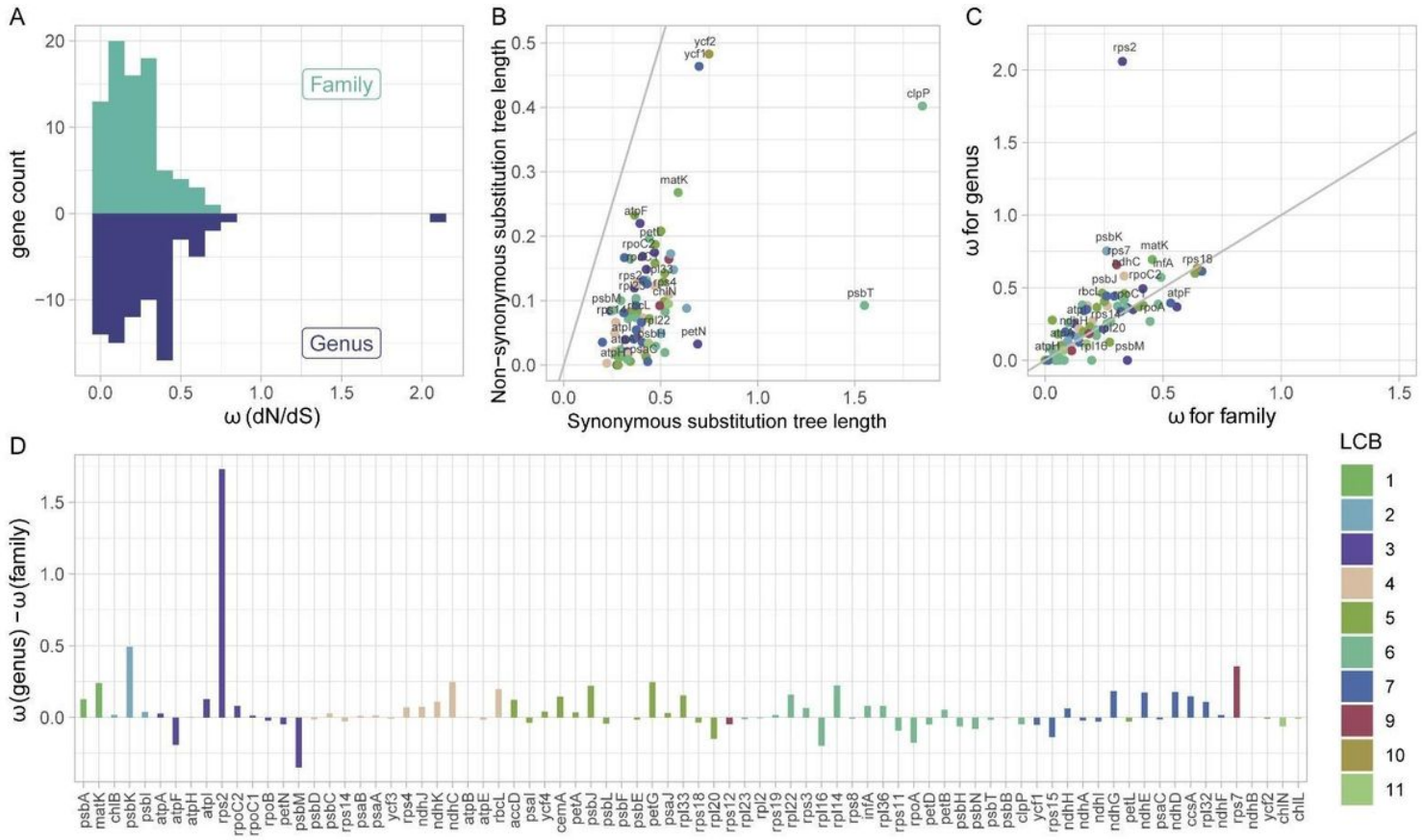

Figure 3

Values of dN/dS ratios (ω) estimated for each protein-coding chloroplast gene. All values of ω were estimated as a single average value across all branches from the phylogeny in Figure 1, using either all 17 species of Podocarpaceae (family estimates) or only the 9 species of *Podocarpus* (genus estimates). Values of ω < 1 indicate negative selection against non-synonymous substitutions, ω ≈ 1 indicate neutral evolution, and ω > 1 suggest positive selection. Colors in B-D indicate the locally colinear block (LCB) in which a gene is located. A) Histograms of ω for all 82 protein-coding genes estimated for the family (upper bars) or for the genus (lower bar). B) Plot of total tree lengths for each gene, measured as non-synonymous versus synonymous substitutions. C) Plot comparing ω values estimated at the genus *versus* family levels. D) Difference in ω value between genus and family for each gene, plotted following the order of the genes along the plastome.