Molecular Phylogenetics and Evolution 163 (2021) 107214

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Introgression is widespread in the radiation of carnivorous *Nepenthes* pitcher plants

Mathias Scharmann [a,b,*], Andreas Wistuba [c], Alex Widmer [a]

[a] Institute of Integrative Biology, ETH Zurich, Universitätstrasse 16, 8092 Zürich, Switzerland
[b] Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland
[c] Friedhofweg 4, 88437 Maselheim, Germany

ARTICLE INFO

Keywords:
Evolutionary radiation
Carnivorous plants
Nepenthes
Introgression
Phylogenetic networks
Admixture

ABSTRACT

Introgression and hybridization are important processes in plant evolution, but they are difficult to study from a phylogenetic perspective, because they conflict with the bifurcating evolutionary history typically depicted in phylogenetic models. The role of hybridization in plant evolution is best documented in the form of allopolyploidizations. In contrast, homoploid hybridization and introgression are less explored, although they may be crucial in adaptive radiations. Here we employ genome-wide data (ddRAD-seq, transcriptomes) to investigate the evolutionary history of *Nepenthes*, a radiation of c. 160 species of iconic carnivorous plants mainly from tropical Asia. Our data indicates that the main radiation is only c. 5 million years old, and confirms previous bifurcating phylogenies. However, due to a greatly expanded number of loci, we were able test for the first time the long-standing hypotheses of introgression and historical hybridization. The genus presents one very clear case of organellar capture between two distantly related but sympatric groups. Furthermore, all *Nepenthes* species show introgression signals in their nuclear genomes, as uncovered by a general survey of ABBA-BABA-like statistics. The ancestor of the rapid main radiation shows ancestry from two deeply diverged lineages, as indicated by phylogenetic network analyses. All major clades of the main radiation show further introgression both within and between each other, as suggested by admixture graphs. Our study supports the hypothesis that rapid adaptive radiations are hotspots of introgression in the tree of life, and highlights the need to consider non-treelike processes in evolutionary studies of *Nepenthes* in particular.

1. Introduction

Introgression is emerging as a fundamental process in the diversification of life in general and plants in particular, for example in the form of hybrid speciation or the exchange of genetic variants between species (e.g. Abbott et al., 2013; Folk et al., 2018). Introgression is the process through which species or populations exchange parts of their genomes through hybridization and subsequent repeated back-crossing. Introgressed genetic variants may facilitate environmental adapation (Racimo et al., 2015) and promote rapid adaptive radiations (Meier et al., 2017; Seehausen et al., 2014). In phylogenetics, however, introgression is usually considered an exception rather than the norm. This is at least partially due to the methodological challenges that are encountered when seeking robust empirical evidence of introgression in a phylogenetic framework.

Phylogenies aim to establish evolutionary relationships in the form of a tree-like graph which connects observed entities via hypothetical common ancestors and lines of descent (Felsenstein, 2004; Maddison, 1997). Species trees, in particular, have the scope to represent the descent of species, and thus should reflect all genetic loci or genes which make up the species, i.e. the genome. As such, a species tree is necessarily a multi-locus phylogeny and can be understood as a summary over many gene trees (geneaologies), and as a cloud or container of gene trees (Edwards, 2009; Maddison, 1997). However, the genes of any given species do not necessarily share the same history, and hence gene trees within any nuclear and between nuclear and organellar genomes may be conflicting. Discordance or incongruence among gene trees frequently arises from biological processes rather than from methodological artefacts, and it must not be discounted when summarising genomic data into species trees (Degnan and Rosenberg, 2009; Hahn and Nakhleh, 2016; Maddison, 1997). If introgression occurred, it can become apparent as incongruent gene trees with lower evolutionary divergence

* Corresponding author at: Institute of Integrative Biology, ETH Zurich, Universitätstrasse 16, 8092 Zürich, Switzerland.
E-mail address: msph52@gmail.com (M. Scharmann).

https://doi.org/10.1016/j.ympev.2021.107214
Received 24 January 2021; Received in revised form 14 May 2021; Accepted 25 May 2021
Available online 28 May 2021

between species against a baseline of non-introgressed gene trees. The problem is that introgression is just one of several confounding causes of gene tree incongruence.

The most important and ubiquitous source of gene tree incongruence is incomplete lineage sorting (ILS), or 'deep coalescence', which is observed when mutations have arisen as segregating polymorphisms in ancestral populations and then became 'sorted' into descendants (Degnan and Rosenberg, 2009; Edwards, 2009). In sexual populations, mutations do not cause instant 'substitutions' but instead transient polymorphims that remain in a population for a time proportional to the population size. Time to fixation of a mutation can take longer than the time in between speciation events, resulting in gene geneaologies incongruent with the species tree. ILS is expected to be particularly prevalent in rapid radiations, at both young and ancient phylogenetic nodes with short branch lengths (Degnan and Rosenberg, 2009).

How do species tree inference methods perform in the face of gene tree incongruence? A diverse array of concepts and computational strategies exists for summarising phylogenomic datasets into species trees (reviewed by e.g. Liu et al., 2015). The simplest is to concatenate all loci and fit a bifurcating phylogeny to the entire "supermatrix" as one (Rokas et al., 2003). Although concatenation ignores ILS and other conflicting signals, its accuracy can be similar to or better than that of "ILS-aware" strategies even if ILS is present (Mirarab et al., 2016; Tonini et al., 2015). "ILS-aware" strategies in general adopt a coalescent model to handle ILS, and, for example, construct bifurcating species trees as "supertrees" in the form of democratic majority-rule summaries of quartet topologies contained in gene trees (e.g. ASTRAL, Mirarab et al., 2014; Zhang et al., 2018b). While many species tree methods thus handle ILS with more or less appropriate results, the detection and handling of introgression is generally beyond their scope.

Many targeted introgression inference methods employ population allele frequencies and are thus not suitable to phylogenetic sampling schemes and timescales (Twyford and Ennos, 2012). Phylogenetic networks, i.e. graphs which contain closed circles, rather than only bifurcations, can be estimated for scenarios in which both ILS and introgression operate, using certain assumptions to ease the problem of identifiability (Solís-Lemus et al., 2017; Solís-Lemus and Ané, 2016; Wen et al., 2018; Zhang et al., 2018a). Unfortunately, current implementations are computationally limited to networks with relatively few species (Folk et al., 2018; Wen et al., 2018). To overcome this problem, a two-step procedure to analyse introgression in species phylogenies can be used. First, a bifurcating species tree is fitted to the data using concatenation or ILS-aware methods. Second, the same data is scanned for signatures of introgression against the bifurcating hypotheses. Hence, a bifurcating phylogeny serves as a null-model to describe the data, and signals which are not expected under bifurcation remain to be explained, potentially by introgression. One prominent test to identify non-treelike signals and to distinguish introgression from ILS is Patterson's D-test ("ABBA-BABA test", Durand et al., 2011; Green et al., 2010). A suite of similar statistics has recently been developed (Blischak et al., 2018; Hahn and Hibbins, 2019; Hibbins and Hahn, 2019; Pease and Hahn, 2015). Other methods specifically evaluate the probability that particular gene trees, such as mitochondrial or plastid phylogenies, are not resulting from ILS in a species tree, which presents a test for organellar introgression (e.g. Folk et al., 2017; Joly, 2012).

Similar to computational tools, phylogenomic sequencing strategies are manifold. As whole genome sequencing remains challenging and is not necessary for many purposes, transcriptomic (RNA-seq) and reduced-representation genome sequencing (RRS) methods are popular choices (e.g. McKain et al., 2018). Transcriptomes yield thousands of genic sequences which are informative from young radiations to the most deeply diverged plants (One Thousand Plant Transcriptomes Initiative, 2019; Wickett et al., 2014), but are relatively costly and require live material. Among RRS methods, restriction-site associated DNA sequencing (RAD-seq, e.g. Baird et al., 2008; Peterson et al., 2012) occupies the opposite end of the spectrum: it scales to many samples because of a lower price and yield tens of thousands of short anonymous loci, but its usefulness is limited to population genomics and phylogenomics of shallow and intermediate levels of sequence divergence (Cariou et al., 2013; Eaton and Ree, 2013; Nadeau et al., 2013; Wagner et al., 2013). Another common approach uses target-capture probes to sub-sample genomes through hybridization capture, which is robust over shallow and deep levels of sequence divergence but limited to a predefined set of just a few hundred loci (e.g. Johnson et al., 2019).

Here we investigate the phylogeny of a radiation of iconic carnivorous plants through reconstruction of species trees and introgression. *Nepenthes* (Nepenthaceae) are long-lived vines or shrubs and produce strongly modified leaves in the form of a pitcher (Cheek and Jebb, 2001; Clarke, 2001, 1997; Ellison and Adamec, 2018; Juniper et al., 1989; McPherson, 2011, 2009). While understanding of the biomechanics and physiology of *Nepenthes* traps has made significant progress in recent years, much is still unclear about trap functional diversity and its evolution within the genus that contains at least 160 species, distributed from Madagascar to New Caledonia, and from northeast India and southern China to northern Australia (Cheek and Jebb, 2001, 2013; Clarke, 1997, 2001; Clarke et al., 2018b; McPherson, 2009, 2011). This evolutionary radiation has produced tremendous morphological variation of the pitcher trap, such that all its anatomical parts were modified, extended or eliminated in different combinations. Additional diversity exists at the level of prey spectra (e.g. Chin et al., 2014) and presumably in the digestive mechanisms, as the pitcher fluids range from slightly to extremely acidic, from watery to viscid (Bonhomme et al., 2011), from a few millilitres to more than one litre in volume, and contain different solute proteins (Biteau et al., 2013) as well as different communities of interacting animals and microbes (Beaver, 1983; Bittleston, 2018; Kitching, 2000; Sickel et al., 2016; Takeuchi et al., 2015). In most *Nepenthes* habitats, multiple species co-occur and often hybridize naturally (Peng and Clarke, 2015). The dynamics of the *Nepenthes* radiation, as well as factors promoting speciation, sympatric coexistence and functional divergence of the pitcher traps, remain elusive. They may include adaptations in carnivory among other more conventional drivers of plant speciation (Givnish, 2010; Rieseberg and Willis, 2007; Widmer et al., 2009). An essential resource to test such hypotheses is knowledge of the evolutionary history.

Nepenthaceae belong to the "non-core Caryophyllales" (Chase et al., 2016; Cuénoud et al., 2002), and form a clade with Ancistrocladaceae, Dioncophyllaceae, Droseraceae, and Drosophyllaceae, of which the latter three comprise carnivorous genera of predominantly flypaper-type traps. Indeed, such traps may be ancestral to this clade of five families, but may have been lost in Ancistrocladaceae and some Dioncophyllaceae (Heubl et al., 2006; Renner and Specht, 2011). While the pitcher traps of *Nepenthes* could thus be derived from flypaper-type traps, they present a spectacular case of functional convergence at molecular and morphological scales with the pitcher traps of *Cephalotus* (Oxalidales) and Sarraceniaceae (Ericales) (Bauer et al., 2013; Fleischmann et al., 2018; Fukushima et al., 2017; Juniper et al., 1989). Uncovering the phylogeny of the crown *Nepenthes* radiation was the aim of several previous efforts (Mullins, 2000; Meimberg et al., 2001; Meimberg and Heubl, 2006; Alamsyah and Ito, 2013; Nauheimer et al., 2019; Murphy et al., 2019). While each study speculated that observed gene tree incongruence is due to introgression, or that hybridization played a major role in speciation, these hypotheses were never tested. The present study aims to close this gap using an appropriately large sample of genome-wide markers.

We focussed in particular on incongruent phylogenomic patterns and aimed at distinguishing ILS from introgression as their causes. We furthermore asked what the age of *Nepenthes* is. Because of the ease with which contemporary species hybridize in nature and cultivation, and because of a possibly young age of the radiation, we expected strong gene tree incongruence due to both ILS and introgression.

2. Methods

2.1. Plant material and sequence data generation

We included a total of 243 individual *Nepenthes* plants representing at least 144 morphospecies and one natural hybrid, along with 26 outgroup taxa (Table S1). Our sampling is nearly exhaustive for accepted *Nepenthes* taxa and relevant outgroups. We generally sampled one individual per species but included multiple individuals where possible.

ddRAD-seq libraries (Peterson et al., 2012) were prepared on DNA extracted with silica column kits (Nucleospin Plant II, Macherey Nagel, Düren, Germany) from fresh or liquid-preserved (Camacho-Sanchez et al., 2013) leaf material, using enzymes EcoRI and TaqI. Libraries were sequenced in 96-plex pools on Illumina HiSeq 2500 instruments, with varying single-end read lengths (101, 126, 151). Data was deposited at the European Nucleotide Archive (ENA) under project PRJEB37794. For transcriptomes, total RNA was extracted from *Nepenthes* leaves or pitchers using the Total RNA Mini Kit (Plant) (Geneaid Biotech Ltd, New Taipei City, Taiwan) with the "PRB" lysis buffer, using two lysis steps and a DNAse digest. RNA-seq libraries were prepared from 0.5 μg total RNA using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA), and sequenced on an Illumia HiSeq 2500 for 126 bp paired-end reads. Data were deposited under ENA project PRJEB37797. Additional sequence data was downloaded from public sources in the form of transcriptomes (Bemm et al., 2016; Fukushima et al., 2017; Walker et al., 2017) or large alignment files (Yang et al., 2015).

2.2. ddRAD-seq read processing & nuclear data

We genotyped ddRAD samples following a modified dDocent pipeline (Puritz et al., 2014). The principle steps are read demultiplexing and quality filtering, denovo assembly of reference contigs ("RAD-tags"), read mapping, variant calling and variant quality filtering. Raw reads were stringently filtered for an accurate barcode sequence by enforcing Phred score >=20 in each of the first five bases (the barcode). We then demultiplexed using process_radtags (Catchen et al., 2013) without the barcode-rescue option and quality-filtered reads to at least Phred score 20 in all sliding windows of 15% of the read lengths. Reads passing filter (mean 1.5 million reads per sample, minimum 133,000, maximum 4.3 million) were used to construct reference contigs based on their frequency among all reads. Because we combined data from Illumina sequencing runs with varying read lengths, we included unique 96 bp reads and 121 bp reads if they occurred at least 25 times, but unique 146 bp reads were allowed at the lower threshold of six occurrences as only few samples had these long reads. Unique reads passing these thresholds were clustered at identity 0.9 in VSEARCH (Rognes et al., 2016). Clusters were globally aligned by MUSCLE (Edgar, 2004) and majority consensus contigs formed. These contigs were clustered once again using CD-HIT-EST (Li and Godzik, 2006) at identity 0.9 and the longest contigs retained as the final RAD-tags. Reads of each sample were mapped against this reference using bwa-mem (Li, 2013), and variants were called by freebayes (Garrison and Marth, 2012). Only bi-allelic single nucleotide polymorphisms (SNPs) were retained. The resulting .vcf files were quality filtered by removing any contigs with reads mapped from two control libraries (water templates), or with any SNPs showing heterozygote excess (Hardy-Weinberg tests, p <= 0.05), and by retaining only SNPs with allele balance between 0.25 and 0.75, and quality-depth ratio > 0.25. Individual genotypes were further filtered for minimum and maximum read depth cut-offs corresponding to at least three reads and at most the mean plus two standard deviations of the read depth per individual. Putative organellar RAD-contigs were identified by mapping against a *Nepenthes* plastome (Yao et al., 2019) and a mitochondrial genome (Gruzdev et al., 2018), and removed from the .vcf files. Each *Nepenthes* sample had ~ 100,000–2,000,000 post-filtering SNPs, but the outgroups only ~ 800–6,000 SNPs. Further filters for missingness were specific to downstream analyses (see below).

2.3. Organellar ddRAD data

Demultiplexed ddRAD reads were mapped to a plastid genome assembly of *N. mirabilis* (Genbank NC_041271.1, Yao et al., 2019), and a mitochondrial genome assembly of *N. ventricosa* × *alata* (Genbank MH798871.1, Gruzdev et al., 2018), using bwa-mem. Haploid SNP and indel genotypes were called by freebayes. To generate alignment files, variants were applied to the reference genomes, and regions with less than three reads aligned were masked by gap characters, using bcftools consensus (Li et al., 2009).

2.4. Transcriptomic data processing

Orthology inference from transcriptomes followed the pipeline of (Yang and Smith, 2014) with slight modifications. In brief, transcriptomes were de novo assembled by Trinity and coding sequences predicted by transdecoder (Grabherr et al., 2011; Haas et al., 2013). After clustering with mcl (van Dongen, 2001) from an all-versus-all BLASTP comparison table, the clusters were aligned on the peptide sequence, and maximum-likelihood trees inferred on the reconstituted DNA alignments. To form finer clusters, excessively long branches were trimmed and the trees cut on long internal branches. The alignment, tree inference and pruning/cutting procedure was iterated once again and the final clusters accepted as homologous gene sets. We then generated the final ortholog gene sets by pruning paralogs according to the maximum-inclusion strategy. Fourty-four ortholog gene sets matched the *Nepenthes* organelles (aligned with bwa to Genbank MH798871.1, Genbank NC_041271.1) and were excluded, so that the transcriptome data used in this study contains only nuclear sequences.

2.5. Preparation of datasets, tree estimation and molecular dating

We prepared a total of eight phylogenomic datasets for the purposes of species tree inference, rooting of the species tree, molecular dating, and testing of introgression. The datasets differed in their composition of ddRAD-seq and transcriptome data, the sub-genomes which they represent, the data missingness, and were analysed with several tools (Table 1).

As SNP-only alignments are known to distort the branch length estimates of maximum-likelihood algorithms (Leaché et al., 2015), we included invariant sites in the ddRAD data by inserting SNP genotypes into the reference RAD-tag sequence, with heterozygotes as ambiguity code.

To determine the topology within the *Nepenthes* species tree relative to outgroups (rooting), we created the dataset 'rooting low missingness' (RLM). We selected a subset of 24 samples representing main lineages of *Nepenthes* and seven outgroup taxa, and for these we constructed an alignment matrix that contained ddRAD data and transcriptome data together. We strictly filtered genes for presence in all 20 transcriptome samples, and RAD-tags for presence in 20 out of 24 ddRAD samples. We used RAxML v. 8.2.4 (Stamatakis, 2014) to generate a concatenated supermatrix species tree of RLM under the GTR + CAT substitution model, and calculated branch support values with the SH-like likelihood ratio test. As an alternative to the supermatrix, we inferred the separate gene trees from all transcriptome loci and ddRAD-loc using RAxML (GTR + CAT), and fed them into ASTRAL v. 5.1.2 (Mirarab et al., 2014; Zhang et al., 2018b) to construct a supertree. 'Gene trees' from individual ddRAD-loci always contained very large polytomies, as expected from the low number of SNPs per ddRAD-locus (usually less than ten). However, ASTRAL naturally handles gene tree polytomies by considering all of the implied quartets.

To estimate species trees for the genus *Nepenthes* itself, we used exclusively ddRAD data, from 235 *Nepenthes* samples (incl. six duplicates) and no outgroups. To evaluate the effect of the number of

Table 1
Overview of the phylogenomic data matrices and their purposes in this study. Alignments, tree files and further outputs are deposited at Mendeley Data (https://data.mendeley.com/datasets/fp77hgb466/2, DOI: https://doi.org/10.17632/fp77hgb466.2), and raw sequencing data at the European Nucleotide Archive (ENA), projects PRJEB37794 and PRJEB37797.

dataset	supporting hypothesis	inference methods	Figure number	subgenomes	site presence filter	N Nepenthes	N outgroup	N ddRAD-seq	N transcriptomes	N total	loci/gene trees ddRAD-seq	loci/gene trees transcriptome	alignment length	gaps	alignment patterns	total gene trees
RLM	root of Nepenthes species tree	supermatrix: RAxML	S1	nuclear	low missingness	20	7	24	20	27	5652	1334	2,724,817	32.01%	326,469	na
		supertree: ASTRAL	1	nuclear	low missingness	20	7	24	20	27	5652	1334	na	na	na	6986
DATING	Nepenthes divergence time	supermatrix: RAxML + RelTime	2	any	low missingness	14	24	0	38	38	0	2216	1,023,400	32.76%	573,765	2216
M212	Nepenthes species tree	supermatrix: RAxML	4, S3	nuclear	in >= 90% of samples	235	0	235	0	235	4477	0	670,157	12.59%	68,453	na
		supertree: ASTRAL	4, S2	nuclear	in >= 90% of samples	235	0	235	0	235	4477	0	na	na	na	4477
M118		supermatrix: RAxML	4, S4	nuclear	in >= 50% of samples	235	0	235	0	235	57,556	0	8,226,999	44.41%	1,223,250	na
M4		supermatrix: ExaML	3, 4	nuclear	in >= 4 samples	235	0	235	0	235	365,053	0	46,195,628	82.33%	5,646,448	na
PLAS	Nepenthes organellar tree	RAxML	5 left, S5, S6	plastid	in >= 70% of samples	235	0	235	0	235	1 (plastid genome)	0	11,550	8.10%	3200	na
MITO		RAxML	5 right, S7, S8	mitochondrial	in >= 70% of samples	235	0	235	0	235	1 (mitochondrial genome)	0	26,010	7.39%	4228	na
M4	test introgression	D3 statistic (255668 tests)	6, 7	nuclear	in >= 4 samples	235	0	235	0	235	365,053	0	46,195,628	82.33%	5,646,448	na
STR		SNAQ (45 sets of 6 taxa, 30 runs each)	8B	nuclear	overall in >= 10 out of 17; per taxa set in >= 4 out of 6	10 (4 per taxa set)	7 (2 per taxa set)	0	17	17 (6 per taxa set)	0	8508	na	na	na	8508 total, 5984–6625 per taxa set
M212		SNAQ (211 sets of 6 taxa, 30 runs each)	8A	nuclear	in >= 90% of samples	6 per taxa set	0	6 per taxa set	0	234 (6 per taxa set)	4477	0	na	na	na	4477
		TreeMix (9 taxa sets, 48 runs each)	9, S9	nuclear	in 100% of samples for each of 9 different sample sets	various (71–234 per taxa set)	0	234	0	234	variable	0	various (21119–51228 per taxa set)	0.00%	na	na

included loci and missing data, we built three different datasets: the most stringently filtered matrix required SNPs to be present in at least 212 (90%) of samples (M212), the intermediate stringency dataset required SNP presence in at least 118 (50%) of samples (M118), and for the most relaxed dataset we filtered SNPs for presence in at least four samples (M4). Supermatrix trees were inferred using RAxML as described above for M212 and M118, but for the much larger alignment M4 we used ExaML v. 3.0.20 (Kozlov et al., 2015; GTR + CAT). A supertree was estimated using ASTRAL and RAxML 'gene trees' (as above) for M212, but ASTRAL failed to complete for the larger datasets M118 and M4.

To generate organellar phylogenies, we filtered the columns of alignments resulting from variant application to the mitochondrial and plastid reference genomes for data presence in at least 70% of individuals (datasets PLAS and MITO, Table 1). Maximum-likelihood phylogenies were inferred using RAxML with the GTRGAMMA model and uncertainty assessed through 1,000 rapid bootstraps.

The dataset 'SNAQ-transcriptomes' (STR, Table 1) was specifically compiled to test for introgression with gene trees. To this end, genes from the transcriptome orthology inference (see above) were filtered to contain at least 10 taxa, resulting in 8,058 genes. These were then further filtered down according to more specific criteria (see section *phylogenetic networks* below).

Divergence times for *Nepenthes* were estimated by combining phylogenomic data with fossil calibrated dates of Angiosperm diversification (Magallón et al., 2015). The DATING supermatrix (Table 1) contained transcriptomic data (one-to-one orthologs) from 14 *Nepenthes* and 24 outgroups (Table S1). These outgroups were selected because their divergence times had previously been estimated (Magallón et al., 2015), and curated phylogenomic sequence data were available for them (Yang et al., 2015). Thus, we took 18 samples from the peptide alignment of 1,122 orthologs of Yang et al. (2015), and emended this alignment by sequences from 20 additional transcriptomes (these were also used for the rooting analysis in the RLM dataset). The ortholog sequences in the additional transcriptomes were identified by BLAST searches against Yang's *Nepenthes* cf. *alata* sequence ("WQUF", reciprocal best hits, c. 1,100 found in each transcriptome), and re-aligned by MUSCLE (Edgar, 2004). Furthermore, this alignment was extended by 1,094 genes from the transcriptome part of the RLM alignment (i.e. orthologs identified in this study, present in all of the 20 additional transcriptomes). The added genes were non-redundant with respect to Yang's set of genes, as tested by BLASTP search of the majority consensi. The final DATING alignment was filtered for minimum column occupancy of 30% and minimum 100 amino acid sites per gene. A supermatrix species tree was inferred using RAxML (PROTCATWAG). The nodes of this tree were dated with RelTime (Tamura et al., 2018, 2012) as implemented in MEGA-CC version 7.0.26–1 (Kumar et al., 2016). We specified the WAG substitution model with five gamma-distributed rate categories and invariant sites, and similar clock rates were merged on one standard error. Brassicaceae (*Arabidopsis thaliana*) was specified as the outgroup. For 13 nodes that were also present in the Angiosperm time-tree of Magallón et al. (2015), we supplied absolute time calibrations in the form of upper and lower limits on age, i.e. the upper and lower bounds of the 95% credibility intervals of Magallón et al. (2015) (Table S2).

Phylogenetic trees were rooted or re-rooted using Dendroscope v. 3.5.7 (Huson and Scornavacca, 2012) to avoid swapping of support values (Czech et al., 2017), and then plotted using the R package APE v. 5.3 (Paradis and Schliep, 2019).

Alignments, tree files and further outputs are deposited at Mendeley

Data (https://data.mendeley.com/datasets/fp77hgb466/2, DOI: https://doi.org/10.17632/fp77hgb466.2).

2.6. Testing cyto-nuclear discordance

We evaluated the degree to which discordance of organellar (plastid and mitochondrial) phylogenies with a bifurcating species tree was expected under a neutral coalescent model similar to the method of (Folk et al., 2017). In brief, we used DendroPy v. 4.4.0 (Sukumaran and Holder, 2010) to simulate 10,000 gene trees under the containing ASTRAL supertree M212. The branch lengths of this species tree are in coalescent units (ratio Ne / generations), and were originally estimated from diploid, biparentally inherited nuclear markers. However, organellar markers are generally expected to be haploid and maternally inherited in plants, and therefore should experience only 1/4 of the Ne of nuclear markers, or in other words, four-fold increased rates of coalescence and genetic drift relative to nuclear markers. Hence, for the simulation of organellar markers, branch lengths of the species tree were multiplied by the factor four. The frequency of bi-partitions in simulated organellar trees was then annotated on the observed organellar trees using the function prop.clades in APE (Paradis and Schliep, 2019).

2.7. non-parametric introgression tests (D3)

We used the D3 statistic (Hahn and Hibbins, 2019) to test for signals consistent with introgression in dataset M4 (Table 1). This metric is theoretically similar to Patterson's D-statistic (Durand et al., 2011; Green et al., 2010; Patterson et al., 2012), also called ABBA-BABA. Patterson's D-test exploits the expectations that on the tree of four taxa ((A,B),(C,D)); two site patterns at the tips of the tree should be equally common if evolution proceeded under a bifurcating tree and ILS alone, whereas significant deviations can be interpreted as introgression. D3 applies the same rationale, but on the basis of genetic distances. It furthermore has the important practical advantages that it (1) uses trios of the configuration ((A,B)C); rather than quartets of samples, and hence (2) the identity of the introgression pair within the test is not ambiguous, unlike in Patterson's D-test where *a priori* assumptions about non-introgressing outgroups are necessary. D3 was found to have almost the same power and false-positive rate as Patterson's D (Hahn and Hibbins, 2019).

To explore signals consistent with introgression in the *Nepenthes* radiation as a whole, we applied D3 tests to a large number of automatically generated test trios. To generate these trios, we developed an algorithm (using the ETE toolkit, Huerta-Cepas et al., 2016) that finds all possible test trios that follow the topology of a rooted, bifurcating species tree (supermatrix tree, M4). First, all possible pairs of taxa within a clade were formed, and then to each pair all possible taxa from the sister clade, i.e. taxa basal to the pair, were added to form trios. This process suggested computationally unfeasible numbers of D3 test trios - therefore, redundant trios were discarded entirely, and partially redundant trios (trios that share one pair of taxa) were randomly downsampled so that each pair of taxa would appear in at most ten test trios. Similar procedures to sample taxa sets from a given species tree were implemented by Malinsky et al. (2020) and Lambert et al. (2019) who retained partially redundant taxa sets.

The D3 test was implemented in a python script, with pairwise genetic distance per site calculated as the absolute difference in allele frequency (thus including heterozygous sites). The significance of a test was assessed by 100 bootstrapped datasets with blocks of 20 sites (Hahn and Hibbins, 2019), and tests were considered significant at $p <= 5\%$.

Large numbers of significant D3 test results were summarised as *pairwise* signals of introgression by the following logic: each test trio ((A,B),C); was decomposed into the two non-sister pairs A-C and B-C. The sign of the D3 statistic indicates which of these two pairs has the lower genetic distance and therefore in the framework of this test has a signature of introgression, while the *absolute* value of D3 is the magnitude of the signature in that pair. The test provides no evidence for or against introgression in the second pair of the trio. Inspecting all significant D3 test trios, we thus collected the absolute D3 statistics for all pairs of samples and extracted their maximum D3 for visualization. The pairwise matrix of D3 statistics was ordered according to the topology of the species tree from which the test trios were generated, and plotted as a heatmap in R.

2.8. Phylogenetic networks (SNAQ)

We inferred phylogenetic newtorks using the SNAQ method from the PhyloNetworks package v. 0.10.0 (Solís-Lemus et al., 2017). This algorithm uses the topology (but not the branch lengths) of unrooted gene trees to fit an unrooted phylogenetic network with up to m directed migration edges, based on maximum pseudo-likelihood. Importantly, SNAQ models both ILS and migration (introgression), but it places a constraint on the network architecture, namely that no edge may be part of more than one introgression loop ("level one networks"). The computational feasibility of SNAQ and the interpretability of its results are limited by the number of taxa and the resolution of the gene trees on which the analysis is based. Therefore, we used SNAQ to test for introgression among only few taxa and sets of gene trees that were taxa-complete and relatively well resolved. In effect, these constraints restricted the use of SNAQ to the main lineages of *Nepenthes* and outgroups.

In the first analysis, we used gene trees from ddRAd loci (dataset M212), which allowed to include *N. danseri* but no outgroups. We generated 211 sets of taxa, each of which represented all five lineages in the EDG (Early Diverging Grade: *N. pervillei*, *N. madagascariensis*, *N. distillatoria*, *N. khasiana*, *N. danseri*) and one lineage of the EC (Eastern Clade). Thus, we exhaustively explored all species and samples from the EC, using each once, but we used each sample from lineages in the EDG several times (randomly chosen). This strategy effectively treats the multiple species belonging to a lineage as replicates, which produce technically but not phylogenetically independent results. A SNAQ analyis was conducted for each set of taxa, using all gene trees from dataset M212 that could be pruned to at least four taxa from the specific set of taxa. We sequentially estimated networks from $m = 0$ through $m = 2$ with ten independent runs each. We retained the best plausible network of $m = 2$ for further analysis, based on ranking by maximum likelihood and network plausibility. Candidate networks were deemed plausible if the direction of migration edges did not conflict with a biologically meaningful rooting, here on *N. pervillei*. Finally, results from the 211 SNAQ analyses were summarised by extracting the most common network, and the frequency of its migration edges was counted among all 211 networks.

In the second analysis, we used gene trees from transcriptomes which allowed to include outgroups (dataset STR). Following the same strategy as above, we generated 45 taxa sets as technical replicate observations of the same phylogenetic events. Each taxa set consisted of six taxa, including two outgroup taxa, the EDG species *N. khasiana* and *N. pervillei*, and two *Nepenthes* of the EC (Eastern Clade), thereby sampling all 45 possible pairs of EC *Nepenthes* and using the other samples multiple times. The remaining steps were done as described above (gene tree pruning, SNAQ runs, identification of one best plausible network for $m = 2$, summarising of results).

2.9. Admixture graphs (TreeMix)

We fitted admixture graphs with the tool TreeMix v. 1.13 (Pickrell and Pritchard, 2012). TreeMix uses allele frequency information from multiple observed populations (or species) to fit an admixture graph using maximum-likelihood under a model of pure genetic drift. To avoid the computationally prohibitive search of phylogenetic network space, the TreeMix heuristic first optimises a bifurcating tree and examines residual population co-variation that is poorly explained by it, and then

places a number of migration edges on the tree to account for the residuals.

We ran TreeMix analyses for each of nine different sets of taxa, which were selected to test specific hypotheses on admixture between the main lineages of *Nepenthes* or within each of these lineages. Allele frequency data were generated by filtering from ddRAD dataset M212 (Table 1) specifically for each taxa set, allowing no missing data. This resulted in 21,000–51,000 SNPs per each taxa set. Each TreeMix analysis explored the number of migration edges from zero to up to 12, and was replicated 48 times with different random number seeds. Replicates with the highest likelihood were used for visualization using a custom plotting function in R. To evaluate how many migration edges were meaningful, we inspected the relative increase in log-likelihood of the fitted admixture graph upon adding further migration edges.

3. Results

3.1. Position of the phylogenetic root of Nepenthes

Analysis of transcriptomes and ddRAD data replicate previous reports that Nepenthaceae is closely affiliated with Droseraceae and Drosophyllaceae - Ancistrocladaceae (Fig. 1). Sister to Nepenthaceae were Drosophyllaceae - Ancistrocladaceae rather than the Droseraceae, but gene tree discordance at this deep phylogenetic node was extreme with less than 50% of gene trees agreeing with this arrangement (QS = quartet support, ASTRAL), as previously reported (Walker et al., 2017). Virtually all gene trees agreed that Nepenthaceae are monophyletic (>99% QS). The crown of the genus *Nepenthes* can be divided into a species-poor early diverging grade (EDG) with mostly western distribution, and an eastern clade (EC) that can be further divided into two species-rich clades. The EDG holds *N. pervillei* from the Seychelles as sister to all other *Nepenthes*, while all further species with a western distribution branch off sequentially, starting with *N. madagascariensis* followed by a sister clade of the two species from the Indian subcontinent, *N. distillatoria* and *N. khasiana*, then *N. danseri* from Waigeo near New Guinea. We noted strong gene tree discordance within the genus *Nepenthes*, implied by the relatively low QS throughout the basal backbone (Fig. 1 B), ranging from 56% to 67% (Fig. 1 A). Relatively high congruence (86% QS) was observed at the node separating the EC from *N. danseri*, but much lower QS were again observed within the subclades of the EC.

Notwithstanding gene tree incongruence, the exactly identical topology was obtained by both supertree and supermatrix (concatenation) inference methods, the latter showing full SH-like support throughout the tree (Figure S1). The combination of transcriptomes with ddRAD data in both supermatrix and supertree methods was robust: for two species we included two representatives with different data, one with both transcriptome and ddRAD data and a second one with only ddRAD and the transcriptome entirely as missing characters (*N. khasiana*, *N. pervillei*). In both species, the two representatives clustered together with very high QS. We therefore interpret the placement of species with only one type of data, especially the critically important early diverged *Nepenthes* with only ddRAD data (*N. danseri*, *N. distillatoria*, *N. madagascariensis*), as reliable.

3.2. DNA sequence divergence and molecular clock dating of the Nepenthes radiation

The *Nepenthes* radiation comprises at most c. 3.7% DNA sequence divergence over all sites, and 8.9% at the more neutrally evolving fourfold degenerate sites (FFDS), as counted in the codons of 1,334 genes in dataset RLM (Table S3). The high values are found between the most basal lineage *N. pervillei* and any other species, while *N. khasiana*, the second-most basal lineage in this dataset, was c. 2.2% (5.3% at FFDS) diverged from any species of the main Southeast Asian radiation. Within the latter clade, sequence divergence at all sites ranged from 0.35% to 1.47%, and from 0.8% to 3.4% at FFDS.

We seeked to estimate how evolutionary divergence in sequence translates to time in years using a molecular clock. The topology of the DATING supermatrix tree (Fig. 2) was fully congruent with the RLM *Nepenthes* species tree (Fig. 1) in respect to *Nepenthes* and its allies, and all branches recieved full SH-like support. Nepenthaceae split from its bona-fide sister lineage, the extant Drosophyllaceae - Ancistrocladaceae, in the late Cretaceaous to early Tertiary at c. 76 (CI 58.3–93.8) Mya. However, the separation from Droseraceae was also dated to approximately the same time frame, consistent with the strong gene tree conflict at this node (see above). The extant *Nepenthes* crown is marked by *N. pervillei,* whose ancestors separated from the remaining *Nepenthes* lineage at c. 12.3 (CI 6.4–18.2) Mya. *N. khasiana* from India last shared a common ancestor with the Eastern Clade (EC) at c. 7.6 (CI 3.3–11.9) Mya. The two sub-clades in the EC diverged onwards from c. 5.5 (CI 2.1–9.0) Mya. This is the upper limit for the ages of all approx. 150 species in the EC.

3.3. Nuclear ddRAD species trees

We next investigated the *Nepenthes* radiation with only nuclear ddRAD data for 235 individual *Nepenthes* samples. Supermatrix analysis of dataset M4 resulted in a well resolved species tree that showed full SH-like support on 93.5% of its internal branches (Fig. 3). The basal

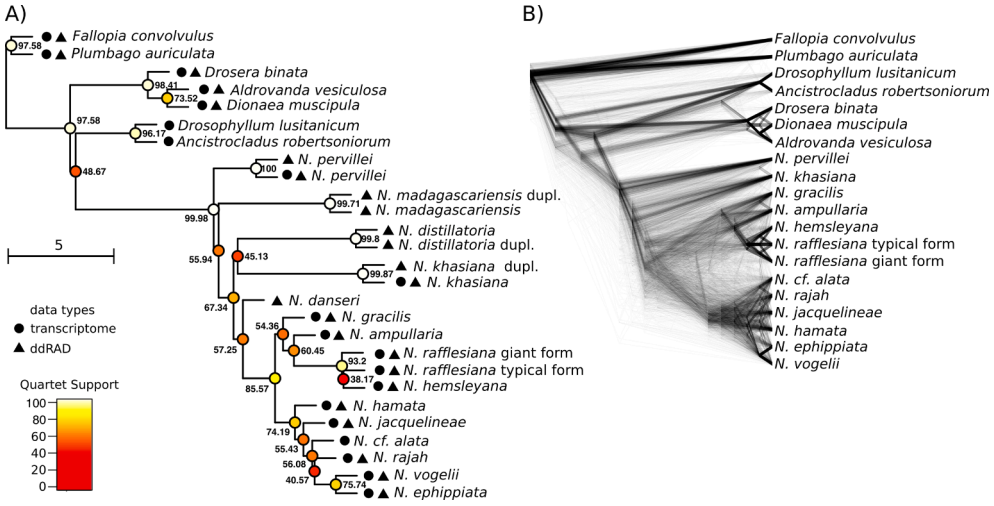

Fig. 1. Rooted species tree for *Nepenthes* with relevant outgroups and representatives of selected lineages within the genus. A) Supertree (ASTRAL) from combined transcriptome and ddRAD data with complete representation of putative early diverged lineages within the genus, inferred from dataset RLM. Node labels indicate gene tree incongruence as the Quartet Support, which is the percentage of gene tree-induced quartets supporting the species tree's topology. Scale bar shows branch lengths in coalescent units (the ratio of effective population size to time in generations). B) Gene tree incongruence illustrated by the 1334 transcriptome gene trees (Densitree plot, true branch lengths not shown) that contributed to the supertree on the left (in addition to gene trees from ddRAD loci).

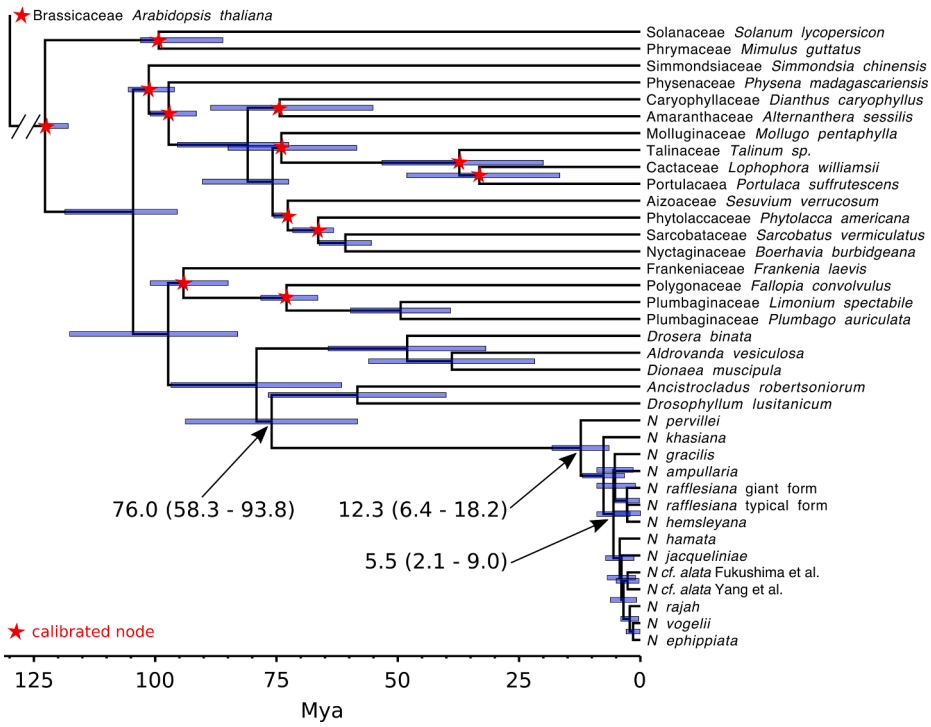

Fig. 2. Timetree of *Nepenthes* and selected representatives of diverse angiosperm orders as outgroups, generated from the amino acid sequences of 2,216 genes (dataset DATING) by RelTime (Tamura et al., 2012). Blue node bars indicate Confidence Intervals, and calibrated nodes (Magallón et al., 2015) are marked by red asterisks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

topology recovered exactly the RLM and DATING trees, i.e. the same EDG (see section 3.1). The power of the M4 dataset is demonstrated in that it resolved plausible phylogeographic structure within taxa, such as grouping of *N. pervillei* samples by the islands of Mahe and Silhouette, and grouping of populations within *N. gracilis*, *N. bicalcarata*, *N. ampullaria*, and *N. rafflesiana* according to their spatial distances (Fig. 3). As another indicator of the robustness of the inference, we conducted three additional bifurcating species tree analyses of the same nuclear ddRAD data (two additional supermatrices differing in column occupancy: M118, M212; supertree on dataset M212, Table 1). These revealed inconsistencies throughout the backbone of the radiation (Fig. 4), while the topology within sub-clades remained largely stable (Figures S2, S3, S4). Together, the four alternative analyses delineated the main structure of the *Nepenthes* crown radiation. Ten main groups with at least four species each are defined by consistent appearance in all four alternative species trees (Figures 3, 4, S2, S3, S4):

1. early diverging grade (EDG): The species from the far Western range of the genus (Madagascar, Seychelles, Sri Lanka and North-East India) and *N. danseri* from Waigeo Island near New Guinea can be distinguished as a grade against all other *Nepenthes* (the "Eastern clade").
2. *mirabilis*-group. This group is a clade and sister to all other species in the Eastern clade, comprising widespread lowland species of both Sundaland and Sahul. It corresponds exactly to "clade 1″ sensu Murphy et al. (2019). A nested sub-clade of lowland to alpine endemics from Sulawesi, New Guinea, and New Caledonia corresponds to the *N. tomoriana* group of Murphy et al. (2019).

The following groups form the second major sub-clade of the Eastern Clade (EC), and correspond to "clade 2" sensu Murphy et al. (2019):

3. Tentaculatae. This clade is consistently placed as the earliest branching within "clade 2", and contains highland species from Borneo and Sulawesi. It is fully congruent with section Tentaculatae (Cheek and Jebb, 2016a).
4. Insignes. This clade branches off subsequent to the Tentaculatae and occurs in the Philippines and New Guinea. It is congruent with sect. Insignes sensu Clarke et al. (2018b) except that we do not find *N. campanulata* to be a member.
5. Pyrophytae. This clade unites all endemic species from the Malay Peninsula and Indochina and is broadly congruent with sect. Pyrophytae (Cheek and Jebb, 2016b) sensu Clarke et al. (2018b).
6. Montanae. A large radiation of all species endemic to Sumatra and Java, which exactly corresponds to sect. Montanae sensu Clarke et al. (2018b). This clade is consistently sister to the Pyrophytae (Fig. 4).
7. *copelandii*-group. This clade is a radiation endemic to Mindanao (southern Philippines), and may be sister to either the *philippinensis*-group or the *maxima*-group (see below).
8. *philippinensis*-group. A clade of mostly north-western Philippine endemic species, with small sub-clades restricted to Palawan, Luzon or Visayan islands. At least one species reached the Mindanao region (*N. viridis*). The group may be sister to the *copelandii*-group, or to a *rajah-copelandii-maxima*-group clade (Fig. 4).
9. *rajah*-group. A small clade of four species endemic to ultramafic highlands of northern Borneo. Despite striking morphological and geographical connections, it is not sister to Palawan's highland species (*philippinensis*-group), but either to the maxima-group, a *maxima-copelandii*-group clade, or to *N. campanulata*.
10. *maxima*-group. A large radiation that dominates mid-elevation to highland environments with endemic species in Borneo. Nested within the Bornean radiation is a sub-clade with *N. maxima* and its allies from Sulawesi, the Moluccas and New Guinea.

The *mirabilis*-group, *copelandii*-group, *philippinensis*-group, *rajah*-group and maxima-group were here named after the oldest described taxon in that clade, or after particularly well-known species.

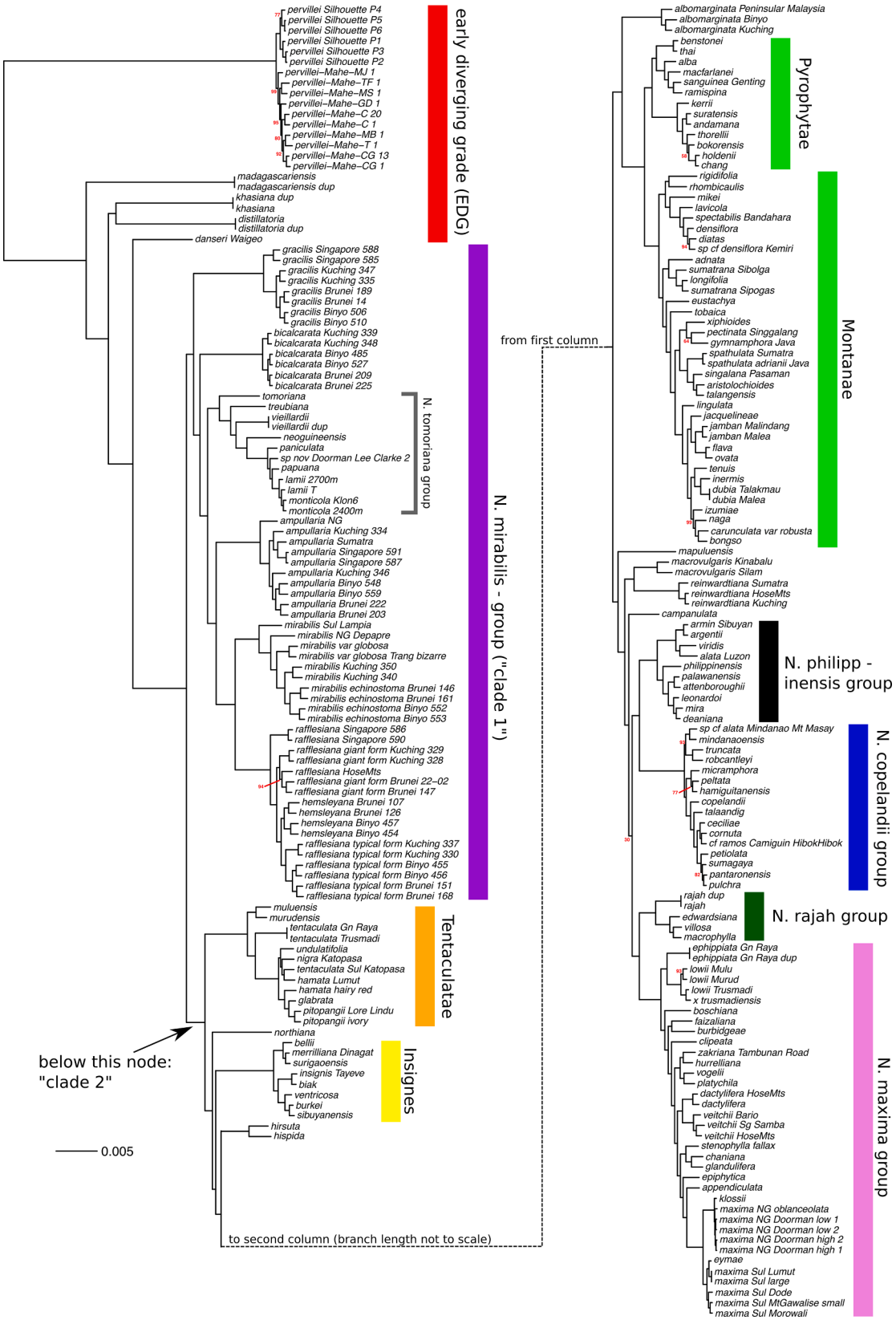

Fig. 3. Bifurcating species tree for the genus *Nepenthes* estimated by maximum-likelihood from a concatenated supermatrix of ddRAD data (dataset M4). Major lineages that were consistently grouped in this and alternative species trees (M118, M212 supermatrix trees, M212 supertree) are named and indicated by vertical bars. The colouring scheme is consistent with that in the organellar gene trees (Fig. 5). SH-like support values are shown only if less than 100. Scale bar indicates number of expected substitutions per site.

In addition to the EDG and the nine major clades, within "clade 2" there are six species-poor, isolated lineages which all occur on Borneo:

1. *N. northiana* branches off the "clade 2" backbone just after the basal Tentaculatae (supermatrices M4, M118, supertree M212), but may also be sister to the Insignes (supermatrix M212). It is a limestone specialist with a very small range in southwest Sarawak.
2. *N. hirsuta* and *N. hispida* form a clade and inhabit the shrub layer of Bornean heath and lower montane forests; they branch off after the Insignes within "clade 2", or later, after the divergence of Montanae-Pyrophytae.
3. *N. mapuluensis* (east Borneo) shares a similar morphology and habitat type with *N. northiana*, but it was never affiliated with that species in any dataset. It appears to derive directly from the "clade 2" backbone.
4. *N. albomarginata*, a lowland to mid-elevation species widespread in Sundaland, may be sister to Montanae-Pyrophytae or diverge directly from the backbone of "clade 2".
5. *N. reinwardtiana* and *N. macrovulgaris* (the former widespread in Borneo and Sumatra, the latter endemic to northern Borneo) formed a clade that was consistently sister to a clade that unites *N. campanulata* with the *rajah*-, *maxima*-, *copelandii*-, and *philippinensis*-groups.
6. *N. campanulata*, a specialist of overhanging limestone cliffs in Borneo, was sister to the rajah group, or sister to a clade uniting the *philippinensis*-, *copelandii*-, *rajah*-, and *maxima*-groups.

3.4. Organellar phylogenies and cyto-nuclear discordance

We analysed the plastid and mitochondrial information contained in ddRAD data separately from the nuclear data. Read mapping statistics showed that *Nepenthes* ddRAD samples contained 0.6–22.7% of plastid-derived reads, which covered 6–13.7% of the sites of the plastome at least three reads deep (alignment PLAS, Table 1). The resulting phylogeny re-capitulated certain lineages already detected in the nuclear species trees, but also showed several important incongruences (Fig. 5, left). Species from the Indian subcontinent (*N. distillatoria*, *N. khasiana*) did not form a clade but instead *N. distillatoria* was placed in between the basal-most species *N. pervillei* and *N. madagascariensis*, while *N. khasiana* was placed closer to the Malesian taxa. *N. danseri* joins Malesian taxa in a large unresolved multifurcation. From this multifurcation emerge several well supported clades, among some isolated species. Yet of the groups identified in the nuclear species trees, only the Tentaculatae, the *copelandii*-group, the *rajah*-group as well as some of the species-poor isolated lineages are recovered as monophyletic in the plastid tree. The large *mirabilis*-group was split into a sub-group consisting of the New Guinean endemics around *N. treubiana*, whereas the predominantly west-malesian species were grouped in a mixed clade also uniting the Montanae and Pyrophytae. The *philippinensis*-group was unresolved with the *maxima*-group. Some species for which we included multiple accessions (*N. albomarginata*, *N. ampullaria*, *N. gracilis*) contained rather divergent plastids from distant clades. Remarkably, all New Guinean members of the *maxima*-group, i.e. six samples including accessions of *N. maxima*, *N. klossii* and *N. oblanceolata*, which were firmly placed with Sulawesian and Bornean relatives in the nuclear species tree, were grouped with full bootstrap support within Insignes. More precisely, plastids of these New Guinean *N. maxima* - allies are sister to the plastids of the New Guinean species *N. biak* and *N. insignis*, members of a clade (nuclear trees) which is otherwise restricted to the Philippines.

We tested whether this case of cyto-nuclear incongruence, peculiarly corresponding to geographic overlap of *Nepenthes* from two distant nuclear groups, could be the result of ILS alone. We simulated neutral gene trees under the bifurcating species tree (M212 supertree), to generate an expected distribution of plastid phylogenies if only ILS had occurred. None of 10,000 simulated plastid phylogenies replicated the observed plastid phylogeny. In particular, it was not possible to generate any gene tree in which the New Guinean branch of the *maxima*-group formed a clade with any Insignes (Figure S5). This result was robust to the assumed rate of coalescence within species, i.e. it did not matter whether the branch lengths of the species tree were re-scaled to reflect the lower effective population sizes of organellar loci, or left at their original (nuclear) lengths (Figure S6). In contrast, bi-partitions contained in the nuclear species tree topology, including the monophyly of *N. maxima* sensu lato, were common in simulations. This analysis strongly rejects ILS as an explanation for the observed plastid-nuclear incongruence.

The mitochondrial portion of the ddRAD data comprised 0.9–16.7% of reads per sample, thereby covering 4–22% of the sites of the mitochondrial reference genome at a depth of three or more reads (dataset MITO, Table 1). Notably the proportion of SNPs in the MITO alignment was only about half that found in the PLAS alignment, and the mitochondrial phylogeny was overall much less resolved. Nevertheless, the mitochondria of *Nepenthes* (Fig. 5, right) display exactly the same basal topology as the plastids, i.e. the sequential branching of *N. pervillei*, then *N. distillatoria*, *N. madagascariensis*, then *N. khasiana*, but *N. danseri* in a very large polytomy with Malesian taxa. This polytomy furthermore holds an unresolved mix of species and small sub-clades from distant nuclear lineages, namely from the *mirabilis*-group, Montanae, Pyrophytae, and Insignes. Some of the main nuclear groups are clades in the mitochondrial tree, and directly join this polytomy, namely the *rajah*-group, Tentaculatae, the *copelandii*-group, and a clade in which the *philippinensis*-group is sister to most members of the *maxima*-group. Mirroring the plastid tree, cyto-nuclear incongruence was also evident in the mitochondrial tree, and all New Guinean species of the *maxima*-group were firmly placed sister to *N. insignis* from New Guinea. As for the plastids, this topology was never expected under species tree M212 and ILS alone, implying that ILS can not explain this case of mito-nuclear incongruence (Figures S7 and S8).

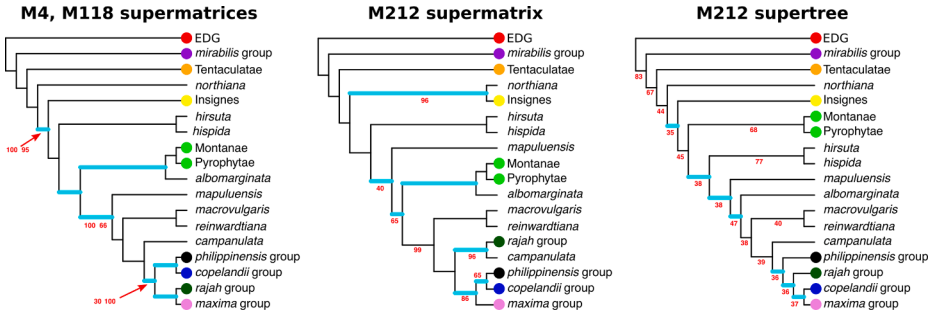

Fig. 4. Three different backbone topologies among the major clades of *Nepenthes*' Eastern Clade subclade 2 are suggested by nuclear ddRAD data, depending on filtering of the sequence data and tree inference methods. Shown are the results for three supermatrices of different occupancy (RAxML, SH-like support values in red, showing only support values less than 100) and one supertree method (ASTRAL, Quartet Support in red). The topology in M4 and M118 supertrees was identical, but SH-like support values differed at certain nodes (annotated in red). Unstable bi-partitions, i.e. those not consistently present in all four species trees, are highlighted in turquois; each tree contains seven of such unstable bi-partitions. EDG = early diverging grade. Full trees are presented in Figures 3, S2, S3 and S4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

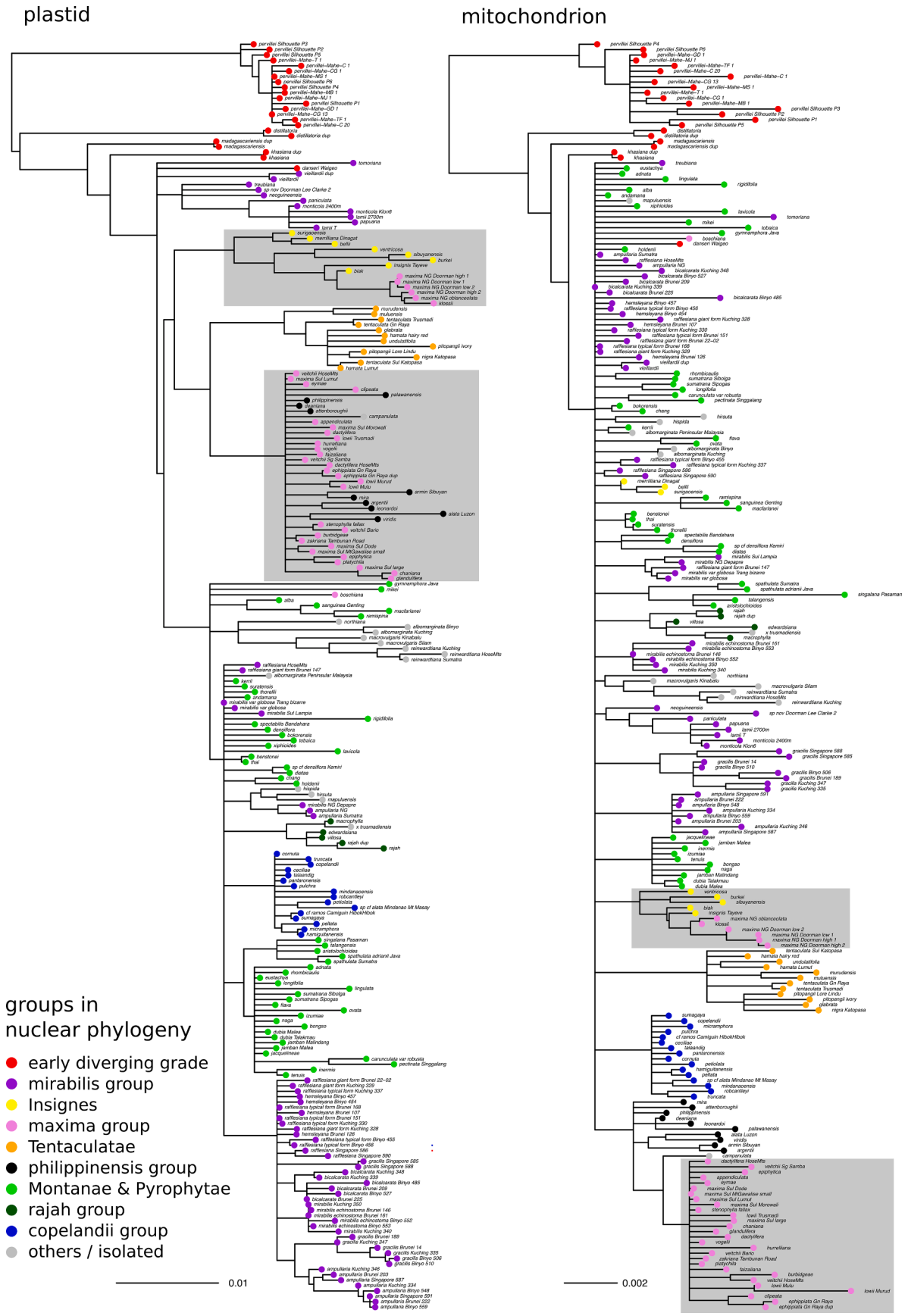

Fig. 5. *Nepenthes* organellar phylogenies from ddRAD markers aligning to organellar reference genomes (datasets PLAS and MITO). Bi-partitions receiving less than 70% bootstrap support are collapsed to polytomies. Each sample is labelled (in colour) with the group it is assigned to based on the nuclear species trees. Grey boxes highlight the Insignes and *maxima*-groups, showing organellar capture. Scale bars indicate number of expected subsitutions per site.

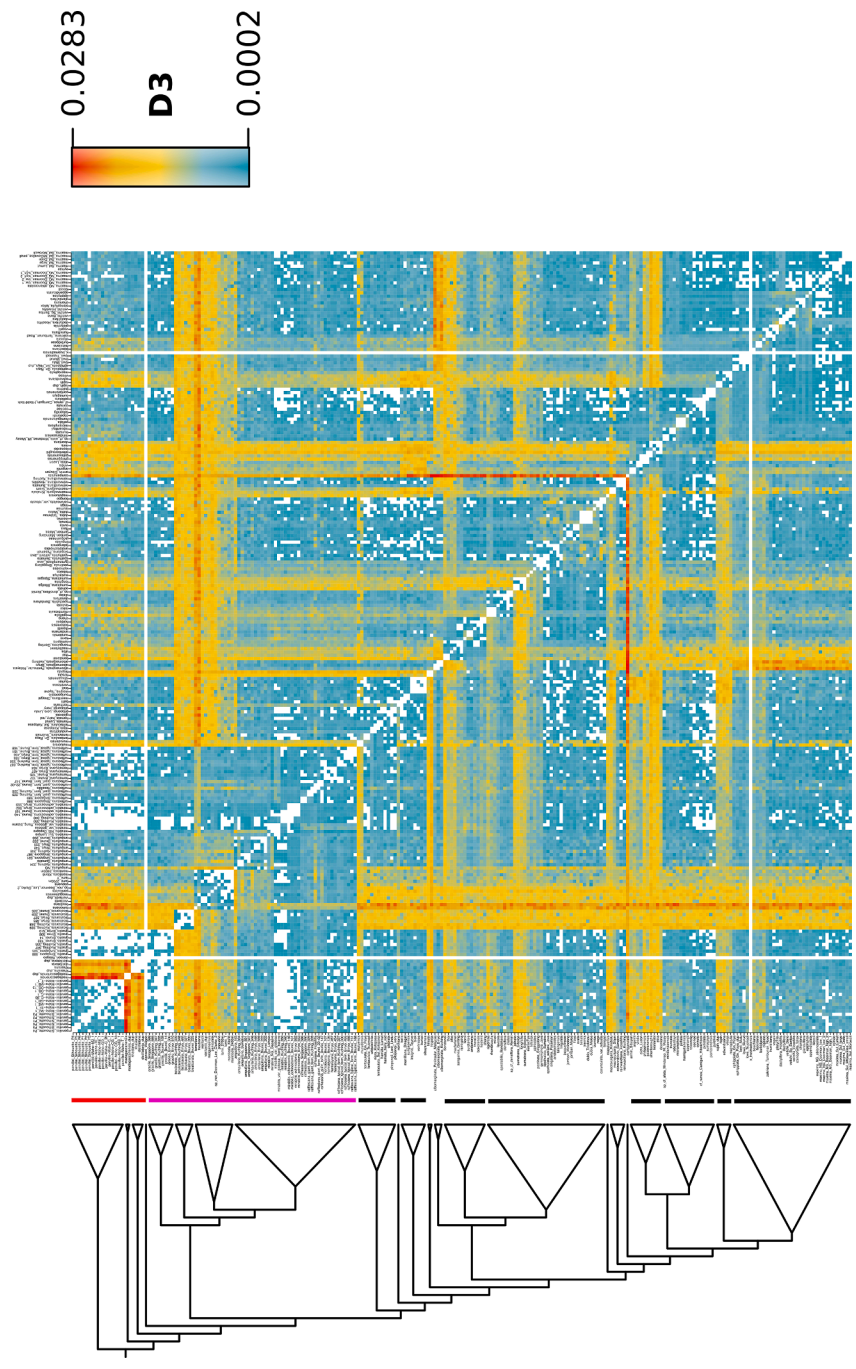

Fig. 6. Patterns of introgression in the genus *Nepenthes* as measured by the D3 statistic (pairwise matrix, right) against species tree M4 (left), both derived from the same data (alignment M4). Results for *N. danseri* and the natural F1 hybrid *N.* × *trusmadiensis* are omitted here and and shown in Fig. 7 because of their higher scale.

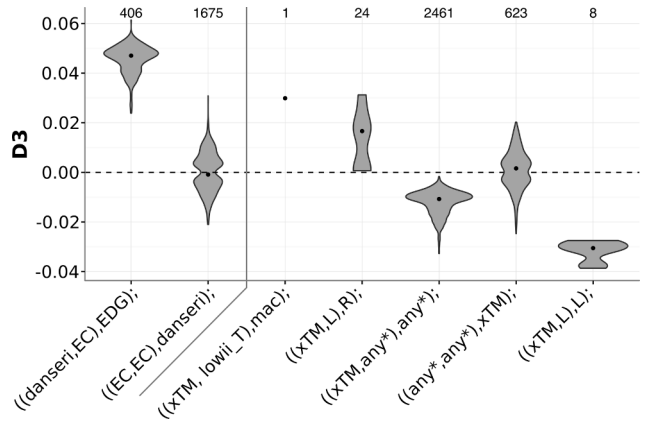

Fig. 7. Introgression statistics D3 against species tree M4 for *N. danseri* (left) and the natural F1 hybrid *N.* × *trusmadiensis* (*N. lowii* × *N. macrophylla* from location Mount Trus Madi, Sabah; xTM, right). The X-axis shows the taxa and topology used in the D3 tests [((A,B),C);]. EC = eastern clade, EDG = early diverging grade, lowii_T = *N. lowii* Trusmadi, mac = *N. macrophylla*, L = *N. lowii* and *N. ephippiata*, R = rajah-group, any* = any other taxa excluding L and R. D3 is shown as violin plots with the median indicated by a dot and the number of D3 values printed atop (number of significant tests at the p = 5% threshold). This figure complements Fig. 6 by showing the D3 values that were omitted there because of higher scale.

3.5. Tests of introgression in Nepenthes nuclear data

3.5.1. General survey of D3 statistics

We scanned for introgression signals against our largest supermatrix species tree (M4, Fig. 3) using the D3 statistic and the alignment underlying this tree. The tree induced 2,135,445 initial D3 test trios, which contained 27,413 unique species pairs. After downsampling to at most ten trios per sample pair, we retained 255,669 unique trios (11.97%) for D3 tests. Significant D3 was found in almost all testable pairs of samples, and varied over three orders of magnitude (Fig. 6, Fig. 7). The D3 statistics appeared to be phylogenetically structured, i.e. clades of the species tree consistently showed higher or lower values. However, interpretation of these D3 signals is challenging as illustrated by the results for *N.* × *trusmadiensis*. This taxon is undoubtedly a natural F1 hybrid between *N. macrophylla* and *N. lowii* from Mount Trus Madi, Borneo (Marabini, 1983). As expected, the test trio of *N.* × *trusmadiensis* with its true parental populations showed a very high and significant D3 statistic (Fig. 7, right). But when the true parental populations were replaced by *N. lowii* accessions from other locations, or by closely related species (*N. ephippiata*, any members of the r*ajah*-group), we also obtained high D3 statistics. Furthermore, trios that did not include the true parents nor close relatives also yielded high D3 statistics, and frequently so for the pair excluding *N.* × *trusmadiensis*. Hence, the "control" of a known F1 hybrid demonstrated that the hybrid status of one sample in a D3 test trio may result in significant D3 values not only for that sample, but also in significant D3 statistics between distant species, creating misleading signals of introgression, or "spillover-effects". Hence, the explorative D3 survey across the species tree showed that introgression may affect most *Nepenthes* taxa, but it could not pinpoint which particular lineages actually participated in introgression. Therefore, we used this survey to generate hypotheses which we subsequently tested by methods which do consider the non-independence of introgression signals on a tree.

3.5.2. Phylogenetic networks with SNAQ

The D3 survey showed the highest signals of introgression between *N. danseri* and other EDG species (Fig. 7, left), even higher than D3 for the hybrid 'control' sample *N.* × *trusmadiensis*. Thus, we hypothesized that *N. danseri* is introgressed with ancestry from both the EC and the EDG. To test this hypothesis, we constructed phylogenetic networks around *N. danseri*, the EDG, and the EC using SNAQ. Surprisingly, SNAQ never inferred *N. danseri* as a donor or recipient of introgression in our search over a large number of alternative taxa sets. Instead, SNAQ consistently inferred networks in which the Eastern Clade, the sister

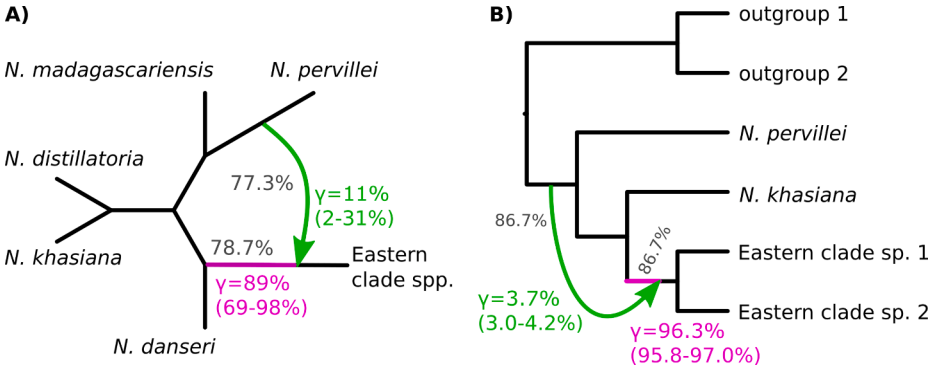

Fig. 8. Phylogenetic networks inferred with SNAQ among the basally branching lineages of *Nepenthes*. Coloured edges are hybridization edges and contribute the annotated ancestry proportions γ (averages, minimum and maximum). Light grey values indicate the proportion of taxon sets in which the corresponding hybridization edge was found. A) Most common network based on ddRAD data (M212, lacking an outgroup) for 211 taxa sets and independent SNAQ runs. Each taxon set contained samples from the five named species of the early diverging grade (EDG), and one sample from the Eastern clade, thereby exhaustively exploring all species and samples from the Eastern clade. B) Most commonly found network based on transcriptome data, among 45 taxon sets and independent SNAQ runs. Each taxon set contained *N. khasiana*, *N. pervillei*, two randomly sampled outgroups, and two samples from the Eastern clade, thus exhaustively exploring all pairwise combinations of samples from the Eastern clade. Since this network is rooted with outgroups, the horizontal dimension indicates evolutionary time. The green arrow does, however, not suggest hybridization between lineages living in distant times. Instead, this arrow indicates that a lineage branched off the species tree before *N. pervillei* and survived at least until the hybridization event, in which it acted as the minor donor. Apart from the admixed EC, "pure" descendants of this hybridization donor lineage are not known, hence it does not appear as a tip (terminal branch) on the right of the tree. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

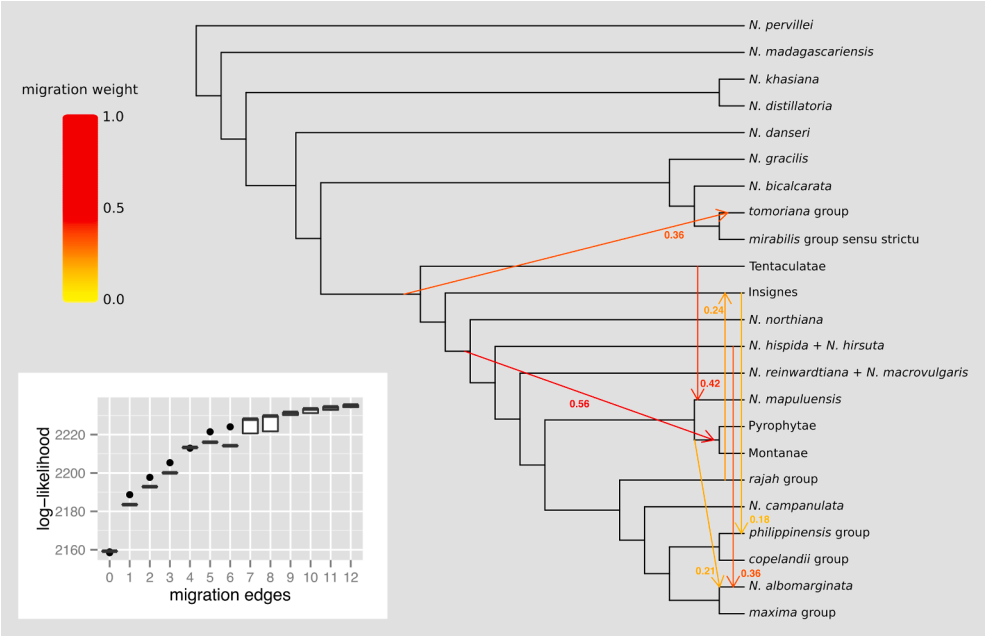

Fig. 9. Maximum-likelihood admixture graph with seven migration edges (TreeMix) for the *Nepenthes* radiation. Allele frequency data for 23 sub-groups of the *Nepenthes* radiation were used (dataset M212). This plot shows topology only, branch lengths are meaningless. Inset on the lower left shows the log-likelihoods of 48 replicate TreeMix runs, each with migration edge parameters from zero to 12. The *mirabilis*-group sensu strictu here comprises *N. ampullaria*, *N. mirabilis* sensu lato, and *N. rafflesiana* sensu lato.

lineage of *N. danseri*, received introgression (gamma = 2–31%) from the *N. pervillei* branch (Fig. 8 A). This single introgression event already maximised SNAQ's pseudo-likelihood, as further events were not fitted even when we allowed them in the model. It thus appeared that *N. danseri* did not introgress, but instead its very strong D3 signals could be "spillover-effects" of introgression between other lineages, as previously exemplified by the *N.* × *trusmadiensis* case. Indeed, a very high D3 for *N. danseri* - EC is consistent with introgression from a lineage basal to all extant *Nepenthes* into the EC, since this can reduce the genetic distance *N. danseri*-EC (simulations confirm this conclusion; see Mendeley Data, doi: https://doi.org/10.17632/fp77hgb466.2).

When using ddRAD data, where no outgroup was available, SNAQ placed the introgression donor lineage with *N. pervillei*. However, these networks were unrooted, and therefore did not imply that the *N. pervillei* lineage was the introgression donor, which is anyway unlikely given the geographic distance between Seychelles and Southeast Asia. To include an outgroup, we repeated the SNAQ analysis with transcriptomes (dataset STR, Fig. 8 B). Despite the differences in taxon sampling and method of sequence data generation, 86.7% of SNAQ networks based on transcriptomes were congruent with those from ddRAD data, inferring one hybridization event into the stem of the EC. Astoundingly, the introgression donor (gamma c. 3.7%) was not grouped with *N. pervillei* but instead it descended directly from the stem of the whole genus. This would imply a previously unknown, hypothetical *Nepenthes* lineage that split off the *Nepenthes* stem lineage before *N. pervillei*, and then survived through time at least until it hybridized with a lineage sister to *N. danseri*, i.e. the stem of the EC. Such hypothetical "archaic ghost" *Nepenthes* could now be extinct, or else not yet discovered.

Following up on this unexpected result, we investigated the gene trees more closely. Genes introduced to the EC from a hypothetical ancient lineage that diverged before *N. pervillei* should show a topology in which the EDG forms a sister clade to the EC, or is nested within the EC. Furthermore, the phylogenetic distance between EC and EDG should be greater in these discordant gene trees than in the species-tree

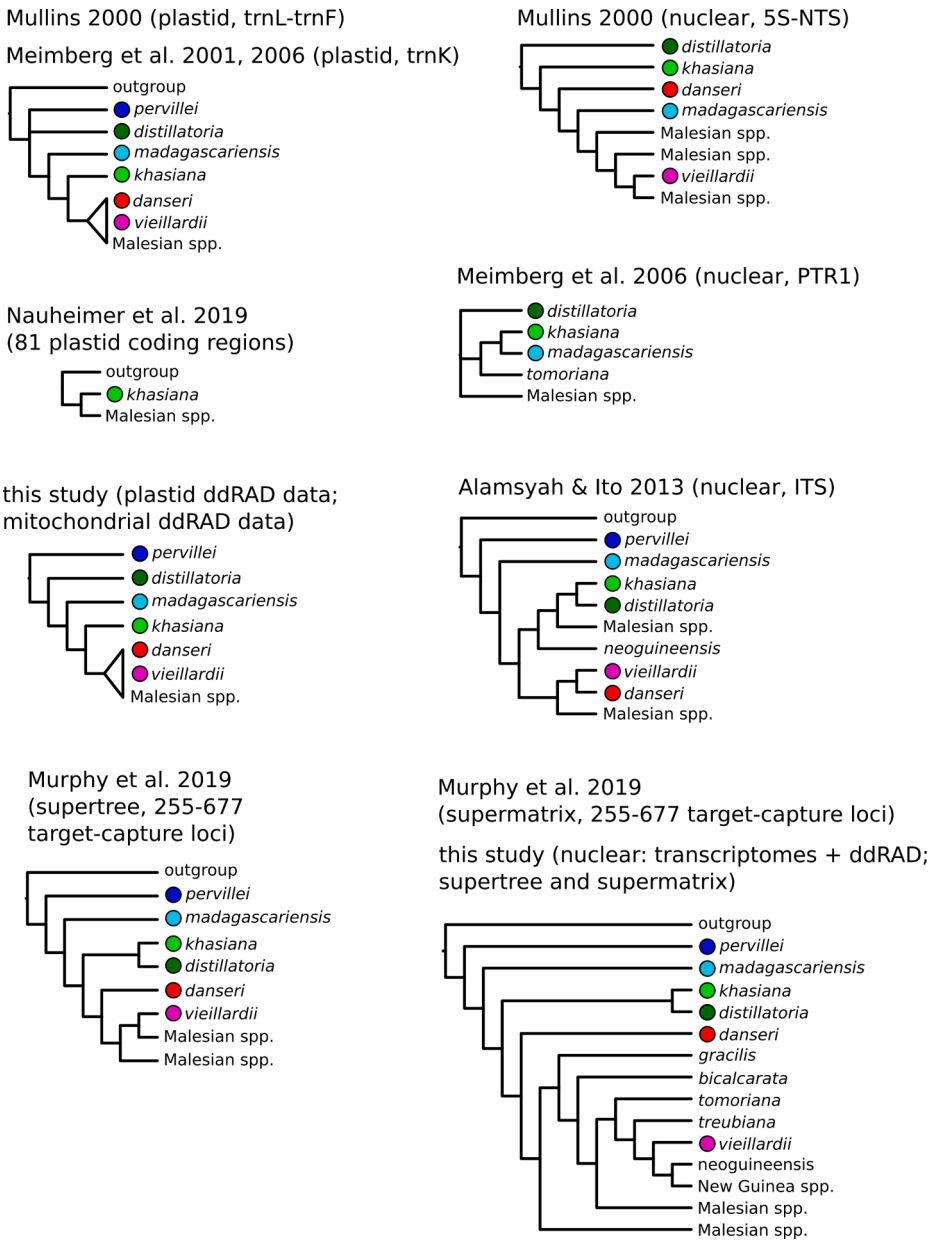

Fig. 10. Overview of bifurcating topologies at the base of the genus *Nepenthes*, showing all alternatives reported by this and previous studies that produced original data, and the results from this study. Triangles here mean that *N. danseri* resp. *N. vieillardii* are unresolved in a multifurcation with other Malesian taxa. *N. masoalensis* not shown.

concordant gene trees. Indeed, among the 1,334 gene trees in dataset RLM, 43 (3.2%) showed the EDG as a sister clade to the EC, and in 79 trees the EDG was nested in the EC. The minimum phylogenetic distance (ML branch length) between EC and EDG *Nepenthes* was significantly greater in the 43 gene trees with reciprocal monophyly than in those concordant with the base of the species tree (species tree-like = 0.019, reciprocal monophyly = 0.023; p = 0.01104; 100,000 permutations).

Together, the SNAQ analyses of ddRAD and transcriptomes (Fig. 8) suggest that an introgression event produced the EC, with a large contribution from a lineage sister to *N. danseri* and a smaller contribution from an unknown lineage that diverged before any other extant *Nepenthes*.

3.5.3. Admixture graphs with TreeMix

To provide an overview of plausible introgression events during the main *Nepenthes* radiation that may explain the ubiquitous D3 introgression signals, we fitted admixture graphs using TreeMix. We arranged the samples (data M212 without *N.* × *trusmadiensis*) in nine different subsets, each to test introgression among a specific set of clades as identified in bifurcating species trees.

In a global admixture graph, all ddRAD *Nepenthes* samples except the known hybrid *N.* × *trusmadiensis* were aggregated as 23 populations according to the clades (Fig. 9). The log-likelihood increased upon addition of admixture events, and did not reach a plateau with 12 admixture events, the largest number we explored, suggesting that substantial co-variance still remained unexplained. Admixture graphs for the remaining eight subsets of taxa are shown in Figure S9 A-H. These results suggested four major admixture events among the main clades of the *Nepenthes* crown radiation:

1. The *tomoriana*-group showed a c. 60% genomic contribution from a lineage sister to *N. ampullaria* or *N. mirabilis* and *N. rafflesiana* s.l., and a c. 40% contribution from a lineage diverged at the base of "clade 2".
2. The common ancestor of the Pyrophytae and Montanae had c. 50% contributed by two deeply diverged lineages within "clade 2" each; the closest living relative of one of the donors could be *N. mapuluensis* and the other a basal lineage in "clade 2".
3. The *philippinensis*-group showed a c. 80% genomic contribution from near the *copelandii*-group, and c. 20% were contributed from Insignes.
4. Insignes were modelled as c. 75% derived from a basal lineage in "clade 2" and c. 25% from the *rajah*-group lineage

Further introgression events have likely ocurred within each of the ten species-rich clades, and often such that the likelihood of the admixture graphs did not reach a plateau upon adding up to 12 admixture edges (Figure S9 A-H). Substantial admixture was also detected for three of the five species-poor lineages of "clade 2": *N. mapuluensis* was modelled as sister to Montanae-Pyrophytae with strong introgression from the more basal diverged Tentaculatae, whereas *N. albomarginata* may have > 20% ancestry from each of three lineages, namely from the base of the *maxima*-group, the common ancestor of Pyrophytae-Montanae, and the *N. hirsuta* lineage. No admixture events were proposed to involve the remaining phylogenetically isolated lineages, *N. reinwardtiana-N. macrovulgaris*, and *N. campanulata*. We note that TreeMix did not propose significant admixture involving the EDG species.

4. Discussion

4.1. Bifurcating Nepenthes phylogenies are consistent across studies and methods

Molecular phylogenies of *Nepenthes* have been studied for over 20 years. The pioneering work of Mullins (2000), Meimberg et al. (2001, 2006), and Alamsyah and Ito (2013) provided fundamental insights and motivated novel hypotheses that later projects aimed to test. The last years saw a renewed interest in phylogenetics of *Nepenthes* with four independent studies being published (Biswal et al., 2018; Bunawan et al., 2017; Murphy et al., 2019; Nauheimer et al., 2019). While all of these provided new insights, three studies (Biswal et al., 2018; Bunawan et al., 2017; Nauheimer et al., 2019) had a regional focus or limited taxon sampling and used few independent genetic markers. Murphy et al. (2019) produced the most extensive phylogeny of *Nepenthes* yet, using target-capture probes resulting in 255–677 loci, on 197 samples from 151 described or putative *Nepenthes* species. The present study is similar in taxon sampling (243 samples from at least 144 *Nepenthes* species and one natural hybrid), but expands the scale of genomic data by a factor of ten to thousand-fold (4,500 to > 350,000 loci from ddRAD-seq, and 1,300 to 8,500 ortholog genes from transcriptomes), and provides separate phylogenies for the nuclear, plastid and mitochondrial sub-genomes. This expanded genomic survey gave the statistical power for the first analysis of introgression in the *Nepenthes* radiation.

Despite the stark methodological differences, an independently generated ddRAD dataset combined with transcriptomes (this study) and target capture data (Murphy et al., 2019) yielded some remarkably consistent results. Bifurcating nuclear species trees estimated by the two studies fully agree on the early diverging grade (EDG), that *N. danseri* is sister to a clade of all other Malesian *Nepenthes* (the Eastern Clade, EC), and two large clades within the EC (Fig. 4). Furthermore, both studies show that *N. vieillardii* is not within the EDG, a hypothesis originally inspired by plastid gene trees and the isolated geographic distribution in New Caledonia (Meimberg et al., 2001; Meimberg and Heubl, 2006; Mullins, 2000). While our organellar phylogenies do not reject this hypothesis, they provide little clarification since *N. vieillardii* resides in very large polytomies. Nuclear data on the other hand place *N. vieillardii* within the EC, as sister to New Guinean species of the *mirabilis*-group resp. *tomoriana*-group (Fig. 4). These relationships are found in supermatrix and supertree approaches in the present study, while Murphy et al. (2019) find a different position for *N. vieillardii* in their supertree (Fig. 10), which they, however, explain as a consequence of high missing data.

Regarding the topology at the base of the genus *Nepenthes*, it is noteworthy that, despite differences in taxon sampling and methodology, all plastid phylogenies (Meimberg et al., 2001; Meimberg and Heubl, 2006; Mullins, 2000; Nauheimer et al., 2019) are fully congruent with one another and also largely congruent with the plastid and mitochondrial gene trees of the present study (Fig. 10). We add confidence that the basal multifurcation in the plastid tree is not an artefact, and we clearly place *N. pervillei* organelles as the earliest branching followed by those of *N. distillatoria*. Nuclear markers of early studies could not reveal which *Nepenthes* lineage diverged first, due to lack of outgroups or nuclear data from *N. pervillei* (Meimberg and Heubl, 2006; Mullins, 2000). Nuclear data from the present study (RNA-seq), target capture (Murphy et al., 2019), and the ITS (Alamsyah and Ito, 2013) all agree that *N. pervillei* is the most diverged of all living *Nepenthes* lineages, in agreement also with its morphology. The nuclear data from the three studies also support a sister relationship for *N. khasiana* and *N. distillatoria*. However, we show here that *N. distillatoria*'s organelles diverged from those of other *Nepenthes* more basally than its nuclear genome. Although this could be a case of organellar capture from an older lineage into *N. distillatoria*, simulations suggest that some cyto-nuclear incongruence is expected at this node of the species tree under ILS alone, albeit with a probability of only 0.6% (Figure S5).

Both the present study and Murphy et al. (2019) identify the same "unstable" or rather phylogenetically isolated taxa within the Malesian "clade 2", namely *N. albomarginata*, *N. campanulata*, *N. hispida* - *N. hirsuta*, *N. northiana*, *N. mapuluensis*, and *N. reinwardtiana* - *N. macrovulgaris*. Furthermore, both studies agree in remarkable detail on many of the topologies within the main Malesian groups, for example that the New Guinean Insignes are sister to the northern branch of the

Philippine Insignes (e.g. *N. ventricosa*), not to the southern branch (e.g. *N. merrilliana*). A minor difference of the two studies concerns the placement of *N. thai*: whereas Murphy et al. (2019) place *N. thai* within the Indochinese sub-clade of the Pyrophytae as sister to *N. kerrii*, the present study resolves it as sister to *N. benstonei* in the Peninsular Malaysian sub-clade of the Pyrophytae. On the other hand, Murphy et al. (2019) placed *N. kongkandana* as sister to *N. benstonei*. Morphology and geography both suggest that *N. thai* is more similar to *N. benstonei* (Cheek and Jebb, 2009) than to *N. kerrii* or *N. kongkandana*, supporting the placement of *N. thai* found in this study. Furthermore, our study provides phylogenetic perspectives for a number of samples and species not included in previous studies, and we comment on their taxonomic relevance in Text S1 of the Supplemental Information. To conclude, the sequence data collected in the present study yield overall highly consistent results with those of previous studies, if approached only with bifurcating phylogenetic models.

4.2. Admixture is widespread in the Nepenthes radiation

We tested the hypothesis that bifurcating species trees are incomplete models of the *Nepenthes* radiation. This is important, because high levels of introgression can mask the history of lineage splitting (Fontaine et al., 2015), and confound phylogenetic analyses. We found strong gene tree conflict in the backbone of the genus, in both transcriptome and ddRAD data. Similar results were obtained by Murphy et al. (2019), but discussed only in terms of the degree of support it lends for bifurcating species trees. We inferred that gene tree conflict in *Nepenthes* is due to introgression rather than ILS and/or gene tree estimation error alone, highlighting the evolutionary consequences of introgression.

A clear case of organellar capture is presented by the Insignes-like organelles of the New Guinean *N. maxima* allies. Our plastid tree from ddRAd data is nearly perfectly congruent with the plastid tree from 81 complete plastid gene sequences of Nauheimer et al. (2019), as far as the taxon overlap goes. We recover Nauheimer et al's plastid clades A and B, albeit with lower sub-clade resolution. *Nepenthes* organellar gene trees are in many parts incongruent with the nuclear species trees, but we showed that it is not necessary to invoke introgression in most cases, because ILS suffices except for the Insignes-like organelles of the New Guinean *N. maxima* allies. The same case of plastid-nuclear incongruence was already observed by Mullins (2000), who suspected it to be the result of hybridization. Hypothetically, organellar capture from Insignes into the *maxima*-group may have occurred when the latter dispersed towards New Guinea from the West (Borneo, Sulawesi, Moluccas) and encountered an already established Insignes, producing fertile hybrids. While significant nuclear introgression signals (D3, Fig. 6) are found between New Guinean *N. maxima* allies and Insignes, they do not generally exceed in magnitude those between the Bornean or Sulawesian *maxima*-group and Insignes. It appears that nuclear Insignes material, which must have been present in an early hybrid population, was subsequently excluded and replaced by *N. maxima* material. Artificial hybrids between Insignes and *N. maxima* are viable and could be used to experimentally test this hypothesis and the mechanism behind it (Tsitrone et al., 2003) through observation of their fertility, meiotic recombination rates and fitness.

Introgression has shaped the radiation of *Nepenthes* in a much more profound way than occasional organellar capture, as suggested by ubiquitous D3 signals throughout the nuclear species tree. We stress that ubiquitous D3 patterns do not imply introgression between all recent species pairs, but likely reflect introgression among some of the ancestral lineages. Another challenge of interpretation is that significant D3 signals in a trio of species may reflect not only introgression within that trio or its ancestors, but also introgression with related species or lineages that were not themselves present in the trio (ghosts), akin to a "spillover-effect" (see also simulation study in Mendeley Data, DOI: https://doi.org/10.17632/fp77hgb466.2). The phylogenetic non-independence of ABBA-BABA like statistics on a tree is well known (Pease and Hahn, 2015). Another point to consider are alternative explanations for D3 signals, such as heterogeneity in substitution rates, and strong ancestral population structure (Hahn and Hibbins, 2019). However, strong variation between *Nepenthes* lineages in the rate of sequence evolution appear not very plausible because life history, a major predictor of substitution rates (Smith and Donoghue, 2008), is similar across the genus. Ancestral population structure on the other hand would have to persist through multiple speciation events to account for non-treelike patterns between distantly related lineages, a scenario that is not plausible, as pointed out by Malinsky et al. (2018). Thus, we are reasonably confident that a large part of the D3 signals in *Nepenthes* is due to introgression, even though D3 cannot simply be translated into a phylogenetic network model.

While most studies using ABBA-BABA-like statistics limit their introgression tests to a narrow set of *a priori* hypotheses, we have explored such patterns globally against the complete bifurcating species tree. Two recent studies have taken similar approaches: Malinsky et al. (2018) developed a method to summarise and visualise f_4 statistics (Green et al., 2010) from all possible combinations of taxa in the radiation of Lake Malawi Cichlids, and found that the majority of nodes of their species tree was affected. Lambert et al. (2019) discovered introgression in a radiation of Mexican spiny lizards by exhaustively searching against the bifurcating species tree for significant DFOIL (Pease and Hahn, 2015), a statistic that can indicate the direction of introgression. They summarised the results as proportions of significant tests between particular lineages. Two arguments stand in favour of such unguided, explorative analyses over isolated tests on selected lineages: (1) introgression may go undetected if it involves non-suspicious lineages, and (2) isolated tests could be misled because of the non-independence of introgression signals on a species tree.

An unexpected signal in our data is that of a deep admixture event at the base of the main EC *Nepenthes* radiation (Figure 8 and D3 statistics). The putative ancestral hybrid population had a major genomic contribution from a lineage sister to *N. danseri*, and a minor contribution from another lineage that diverged prior to all extant *Nepenthes*. Since no genetically pure descendants nor fossils of this lineage are known, and analogous to the field of human genetics (e.g. Durvasula and Sankararaman, 2020), we refer to this lineage as an "archaic ghost" population (AG). Patterns suggesting AG admixture of the EC are detected in both ddRAD and transcriptome sequences. We extrapolate that the divergence time of the AG and the recipient *N. danseri*-like lineage was at least around 7 My at the time of hybridization, given that AG must have diverged before *N. pervillei* and given the point estimates of our dated phylogeny. Admixture between *Nepenthes* lineages with 7 My divergence is not implausible from the viewpoint of reproductive barriers, as evident from the numerous man-made hybrids between deeply diverged extant lineages (e.g. *N. khasiana*-hybrids). At present, we can not exclude the possibility that the deep coalescence of hundreds (RNA-seq) or thousands of loci (ddRAD-seq), which was interpreted as AG admixture by SNAQ, has an alternative explanation, such as long-term balancing selection, or paralogy. The latter problem, i.e. mistaking paralogs for orthologs, can be exacerbated by polyploidy, and indeed *Nepenthes* shows several palaeo-polyploidizations (Walker et al., 2017). However, if paralogy, or more precisely the differential retention of paralogs (homeologs) between lineages, were an explanation for the AG admixture-like gene trees, it would be required that the same paralogs were lost multiple times independently in the sequential branches of the EDG (*N. pervillei*, *N. khasiana*, etc) while two or more paralogs survived through several speciation events up to the ancestor of the EC. It remains to be tested how realistic this scenario is for dozens or hundreds of loci, and given the presumably ancient timing of *Nepenthes*' polyploidizations and re-diploidizations. To corroborate the AG admixture hypothesis for *Nepenthes*, and to study the timing of admixture and selective regimes acting upon introgressed regions, future work should aim to place admixture in the context of a genome assembly, and replicate individuals of *N. danseri*, in comparison with those from putatively

introgressed EC taxa.

We used TreeMix as another parameterised approach to interpret the ubiquitous admixture signals in the whole radiation, because the limited resolution of gene trees from ddRAD-loci within the recently diverged EC, as well as the large number of species, made the application of phylogenetic network tools impossible. Within the EC, we found major admixture events at the base of five of the nine sub-radiations, and for three of the five species-poor lineages (Fig. 9). This suggests that admixture is part of the explanation why these sub-radiations and species-poor lineages were inconsistently arranged in the different datasets and bifurcating phylogenetic models of the present study (Fig. 4) and by Murphy et al. (2019). Dozens of further admixture events were detected within each sub-radiation, and these were often between species overlapping in their current geographic range. However, some caution is advised for TreeMix results, because the power of this method to infer the correct number and property of admixture events is not known. Furthermore, we have a sampling design with a single (or a few) samples from many species, while TreeMix models populations that diverge by neutral drift of ancestral variation only. Nevertheless, it is clear that due to the young age of the radiation and generally large effective population sizes (dioecious mating system), most genetic variation in *Nepenthes* must be ancestral (older than speciation), and homoplasy should be very rare.

4.3. What does admixture reveal about the biology of the radiation?

Evolutionary lineages vary dramatically in their rates of phenotypic and phylogenetic diversification, and *Nepenthes* is no exception. Although quantitative studies of diversification rates and trait evolution are still pending, it is clear that lineages of the EDG, which split at least about 5 Mya (*N. danseri*), presumably by geographic isolation (allopatry), have accrued almost zero net-diversification, that they are similar in their ecology, their floral and vegetative traits, and their carnivorous syndrome. During the same time, the EC has radiated to >150 species in many different habitat types, evolved spectacular morphological diversity with strong modifications of the carnivorous syndrome, and occurs in communities of multiple sympatric species. What can explain these stark differences? Most likely, repeated geographic isolation and novel ecological opportunities (soils, climates, prey spectra) played a role. However, *N. danseri* occurs in the same region as most of the EC (insular Southeast Asia / Australasia), yet it failed to diversify (with the possible exception of *N. weda* and *N. halmahera*; no molecular data is available). Therefore, we propose that introgression and hybrid speciation could be the key factor that allowed one lineage to radiate rapidly while those without remained in relative phenotypic stasis. A growing number of studies report that alleles underlying adaptation or reproductive isolation in recent rapid speciation events and radiations were not new, but much older than these events and brought together by introgression (reviewed in Marques et al., 2019). Our study suggests that introgression occurred throughout the *Nepenthes* radiation and particularly the EC, whereas no introgression events were proposed for the species-poor EDG lineages (nothwithstanding D3 signals, which could be due to "spillover-effects", see sect. 3.5.1). We hypothesize that introgression was a crucial factor supplying the genetic variants required for the rapid radiation of the EC. The "deep introgression" at the base of the EC between two lineages that diverged for several My could have been a key event that combined divergent adaptive alleles and thereby hot-started the EC radiation, similar to the "hybrid swarm" origin of Lake Victoria Cichlids (Meier et al., 2017), and species of cod (Árnason and Halldórsdóttir, 2019). The AG warrants further study, as methods are now available that can detect introgressed genomic regions even if the source population (species) is unknown (Plagnol and Wall, 2006). It may also be found that an AG contributed different alleles to the different lineages of the EC or that there were multiple AG admixture events, similar to the history of different modern human populations that harbor different archaic alleles (Vernot and Akey, 2014). Allele sorting and repeated hybrid speciation could have generated the sub-radiations of *Nepenthes* on different landmasses. Similar phenotypes and trapping syndromes have evolved seemingly repeatedly (Bauer et al., 2012; Moran and Clarke, 2010), but appreciation of introgression and ILS could mean that these represent hemiplasy rather than convergence in the strict sense. That ancestral, introgressed variation is important in plant radiations was also demonstrated by Pease et al. (2016) in wild tomatoes (Solanum sect. Lycopersicon). They report that introgression was one of three important sources of adaptive genetic variation in this radiation, in addition to de-novo mutations and sorting of ancestral variation through speciation processes. Future studies may test such hypotheses in *Nepenthes*, albeit the genetic basis of adaptation and reproductive isolation in the genus is presently unknown.

4.4. Appreciating ILS and introgression in Nepenthes - practical consequences for future studies

Bifurcating phylogenies based on few markers, or phylogenies from combined analyses of nuclear and plastid markers, were employed in several comparative evolutionary studies (Bauer et al., 2012; Gilbert et al., 2018; Merckx et al., 2015; Schwallier et al., 2017, 2016). In the light of expanded genomic data, these phylogenies were not fully representative of the actual history of *Nepenthes*. Instead there is a great amount of gene tree discordance in *Nepenthes*, due to ILS as well as admixture. We may thus expect that hemiplasy, the sharing of homologous alleles or traits due to the sorting of ancestral variation into descendants or introgression (Avise and Robinson, 2008; Hahn and Nakhleh, 2016), is common in the genus. Bifurcating species trees could be appropriate for some applications in taxonomy and systematics. However, ignoring the complex, network-like phylogenetic history of *Nepenthes* genomes could mislead macro-evolutionary studies of diversification rates, phylogeography, and phylogenetic comparative analyses. In *Quercus*, for example, introgression confounds naive inference of phylogeography and clade ages (McVay et al., 2017). Recent work substantiates the concern that phylogenetic comparative analyses for both Mendelian and quantitative traits can be flawed by ILS (Mendes et al., 2018) as well as introgression (Hibbins et al., 2020). Therefore, future studies that utilize *Nepenthes* phylogenies for evolutionary inference should aim to mitigate the risks of hemiplasy (e.g. Wu et al., 2018).

4.5. The age of Nepenthes and phylogeography

Using a bifurcating species tree, we dated the *Nepenthes* stem age to the late Cretaceous, consistent with several previous reports (Fleischmann et al., 2018; Nauheimer et al., 2019). Our point estimate and confidence interval for the *Nepenthes* stem age match almost exactly the estimates given in the source of the secondary calibrations (Magallón et al., 2015), although we did not constrain the ages of nodes close to *Nepenthes* and used a different dataset and dating algorithm. We acknowledge that molecular dating of this radiation with a pronounced history of introgression may be confounded (McVay et al., 2017), and therefore urge caution for susceptible nodes. The crown age and ages within the EDG, however, should be unproblematic in this regards. On the other hand, divergence time estimates should generally be understood with caution, because molecular dating methods can show poor performance under some circumstances, and in particular the accuracy of a posteriori calibration methods such as RelTime is the matter of ongoing debate (Beavan et al., 2020).

The crown age of *Nepenthes* was investigated by several previous studies and reviewed by Nauheimer et al. (2019), who pointed out methodological flaws or lack of detail in some of these reports (30 M years, Merckx et al., 2015; 15 M years, Biswal et al., 2018; 8.7 M years, Fleischmann et al., 2018). Nauheimer et al. found an age of c. 20 ("HPD" 9–35) Mya for the split of *N. khasiana*, using plastome data and the same secondary calibrations as the present study. This point estimate is considerably older than our estimate of 7.6 Mya, although confidence

intervals overlap. Here, we estimated the crown age using the basal-most species *N. pervillei* at 12 (6–18) Mya, and that the main radiation in Southeast Asia is 2–9 My old. This is younger than most previous estimates, but still consistent with the presence of *Nepenthes* pollen in Bornean coal deposits from the late Miocene (c. 7 Mya; Anderson and Muller, 1975). If correct, a younger age does not support Nauheimer et al.'s hypothesis that the docking of the Indian continental plate with the Asian mainland (~40 Mya) played a direct role in speciation of extant lineages, although it can not be excluded that ancestral *Nepenthes* reached Asia from West Gondwana via the Indian plate. Given that these plate tectonic events appear far too old to explain *Nepenthes*'s phylogeography, we hypothesize that the MRCA of *Nepenthes* occurred on a larger landmass from which it dispersed to several destinations, and subsequently went extinct. Before its extinction, a continental MRCA could have sequentially seeded descendants in the Seychelles, then Madagascar, and lastly the eastern Indian subcontinent, from where Southeast Asia was (re-)colonised. Candidate regions are Africa, Eurasia (including Southeast Asia) and India, all of which supported megathermal evergreen vegetation during the Miocene (Morley, 2011). The current distribution of *Nepenthes* together with a dated phylogeny implies that the long-distance dispersal events in this hypothesis are not unrealistic. The vast range of *N. mirabilis*, from China to Australia and from Sumatra to the Pacific, may be a modern analog of the hypothesised spread of ancestral Miocene *Nepenthes* around the Indian ocean.

5. Conclusion

The present study confirms bifurcating phylogenies of *Nepenthes* that were previously reported, and provides novel insight into the timing and frequency of introgression. Due to the genome-wide scale of locus sampling, and nearly exhaustive taxon sampling, which includes many recently described species that were never before available to phylogenetic analysis, we expect that our phylogenetic trees and networks will be an important resource for further evolutionary studies of *Nepenthes*, as well as systematics and pressing conservation issues (Clarke et al., 2018a). Appreciating that introgression and incomplete lineage sorting heavily shape the *Nepenthes* radiation, we advocate that these processes be taken into consideration and warrant further study of this rapid radiation of iconic carnivorous plants.

CRediT authorship contribution statement

Mathias Scharmann: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing. **Andreas Wistuba:** Methodology, Resources, Writing – review & editing. **Alex Widmer:** Funding acquisition, Resources, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Claudia Michel for excellent help in the wetlab. J. Danz, S. u. I. Hartmeyer, M. Zehnder, and U. Zimmermann kindly donated plant material from their private collections. We thank Abdul Hadzid Harith Tinggal, Ieney Daud (Telamba Homestay) and their family, as well as Ulmar Grafe, for hospitality and assistance in Brunei, and Lam Weng Ngai for assistance with field work in Singapore. Hirzi Luqman and Christoph Küffer are thanked for sampling in the Seychelles. For the granting of permits to sample and export plants we thank the Ministry of Industry and Primary Resources of Brunei Darussalam, Forest Department Sarawak and Sarawak Forestry Corporation, the Agri-Food and Veterinary Authority and the National Parks Board of Singapore, the Seychelles Agricultural Agency, and the Botanical Garden of the University of Zürich. Sequencing and computation were carried out with the Genetic Diversity Center Zürich and the Functional Genomics Center Zürich. This work was supported by the ETH Zürich and the Rübel foundation. Keygene N.V. owns patents and patent applications protecting its Sequence Based Genotyping technologies.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2021.107214. Further data is deposited at Mendeley Data, doi: https://doi.org/10.17632/fp77hgb466.2.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J.E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C.A., Buggs, R., Butlin, R.K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S.H., Hermansen, J.S., Hewitt, G., Hudson, A.G., Jiggins, C., Jones, J., Keller, B., Marczewski, T., Mallet, J., Martinez-Rodriguez, P., Möst, M., Mullen, S., Nichols, R., Nolte, A.W., Parisod, C., Pfennig, K., Rice, A.M., Ritchie, M.G., Seifert, B., Smadja, C.M., Stelkens, R., Szymura, J.M., Väinölä, R., Wolf, J.B.W., Zinner, D., 2013. Hybridization and speciation. J. Evol. Biol. 26, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x.

Alamsyah, F., Ito, M., 2013. Phylogenetic analysis of Nepenthaceae, based on Internal Transcribed Spacer nuclear ribosomal DNA sequences. APG Acta Phytotaxon. Geobot. 64, 113–126.

Anderson, J.A.R., Muller, J., 1975. Palynological study of a holocene peat and a miocene coal deposit from NW Borneo. Rev. Palaeobot. Palynol. 19 https://doi.org/10.1016/0034-6667(75)90049-4.

Árnason, E., Halldórsdóttir, K., 2019. Codweb: Whole-genome sequencing uncovers extensive reticulations fueling adaptation among Atlantic, Arctic, and Pacific gadids. Sci. Adv. 5, eaat8788. https://doi.org/10.1126/sciadv.aat8788.

Avise, J.C., Robinson, T.J., 2008. Hemiplasy: A New Term in the Lexicon of Phylogenetics. Syst. Biol. 57, 503–507. https://doi.org/10.1080/10635150802164587.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLOS ONE 3, e3376. https://doi.org/10.1371/journal.pone.0003376.

Bauer, U., Clemente, C.J., Renner, T., Federle, W., 2012. Form follows function: Morphological diversification and alternative trapping strategies in carnivorous *Nepenthes* pitcher plants. J. Evol. Biol. 25, 90–102. https://doi.org/10.1111/j.1420-9101.2011.02406.x.

Bauer, U., Scharmann, M., Skepper, J., Federle, W., 2013. 'Insect aquaplaning' on a superhydrophilic hairy surface: How *Heliamphora nutans* Benth. pitcher plants capture prey. Proc. R. Soc. B Biol. Sci. 280, 20122569. https://doi.org/10.1098/rspb.2012.2569.

Beavan, A.J.S., Donoghue, P.C.J., Beaumont, M.A., Pisani, D., 2020. Performance of A Priori and A Posteriori Calibration Strategies in Divergence Time Estimation. Genome Biol. Evol. 12, 1087–1098. https://doi.org/10.1093/gbe/evaa105.

Beaver, R.A., 1983. The communities living in *Nepenthes* pitcher plants: fauna and food webs. In: Frank, J.H., Lounibos, L.P. (Eds.), Phytotelmata: Terrestrial Plants as Hosts for Aquatic Insect Communities. Plexus Publishing Inc, Medford, New Jersey, pp. 129–159.

Bemm, F., Becker, D., Larisch, C., Kreuzer, I., Escalante-Perez, M., Schulze, W.X., Ankenbrand, M., de Weyer, A.-L.V., Krol, E., Al-Rasheid, K.A., Mithöfer, A., Weber, A.P., Schultz, J., Hedrich, R., 2016. Venus flytrap carnivorous lifestyle builds on herbivore defense strategies. Genome Res. 26, 812–825. https://doi.org/10.1101/gr.202200.115.

Biswal, D.K., Yanthan, S., Konhar, R., Debnath, M., Tandon, P., 2018. Phylogeny and Biogeography of Carnivorous Plant Family Nepenthaceae with reference to the Indian pitcher plant *Nepenthes khasiana* reveals an Indian subcontinent origin of *Nepenthes* colonization in South east Asia during the Miocene epoch. Front. Ecol. Evol. 6 https://doi.org/10.3389/fevo.2018.00108.

Biteau, F., Nisse, E., Miguel, S., Hannewald, P., Bazile, V., Gaume, L., Mignard, B., Hehn, A., Bourgaud, F., 2013. A simple SDS-PAGE protein pattern from pitcher secretions as a new tool to distinguish *Nepenthes* species (Nepenthaceae). Am. J. Bot. 100, 2478–2484. https://doi.org/10.3732/ajb.1300145.

Bittleston, L.S., 2018. Commensals of *Nepenthes* pitchers. In: Carnivorous Plants: Physiology, Ecology, and Evolution. Oxford University Press, pp. 314–332.

Blischak, P.D., Chifman, J., Wolfe, A.D., Kubatko, L.S., 2018. HyDe: A Python Package for Genome-Scale Hybridization Detection. Syst. Biol. 67, 821–829. https://doi.org/10.1093/sysbio/syy023.

Bonhomme, V., Pelloux-Prayer, H., Jousselin, E., Forterre, Y., Labat, J.-J., Gaume, L., 2011. Slippery or sticky? Functional diversity in the trapping strategy of *Nepenthes* carnivorous plants. New Phytol. 191, 545–554. https://doi.org/10.1111/j.1469-8137.2011.03696.x.

Bunawan, H., Yen, C.C., Yaakop, S., Noor, N.M., 2017. Phylogenetic inferences of *Nepenthes* species in Peninsular Malaysia revealed by chloroplast (trnL intron) and

nuclear (ITS) DNA sequences. BMC Res. Notes 10, 67. https://doi.org/10.1186/s13104-017-2379-1.
Camacho-Sanchez, M., Burraco, P., Gomez-Mestre, I., Leonard, J.A., 2013. Preservation of RNA and DNA from mammal samples under field conditions. Mol. Ecol. Resour. 13, 663–673. https://doi.org/10.1111/1755-0998.12108.
Cariou, M., Duret, L., Charlat, S., 2013. Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. Ecol. Evol. 3, 846–852. https://doi.org/10.1002/ece3.512.
Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: An analysis tool set for population genomics. Mol. Ecol. 22, 3124–3140. https://doi.org/10.1111/mec.12354.
Chase, M.W., Christenhusz, M.J.M., Fay, M.F., Byng, J.W., Judd, W.S., Soltis, D.E., Mabberley, D.J., Sennikov, A.N., Soltis, P.S., Stevens, P.F., 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. 181, 1–20. https://doi.org/10.1111/boj.12385.
Cheek, M., Jebb, M., 2016a. A new section in *Nepenthes* (Nepenthaceae) and a new species from Sulawesi. Blumea - Biodivers. Evol. Biogeogr. Plants 61, 59–62. https://doi.org/10.3767/000651916X691510.
Cheek, M., Jebb, M., 2016b. Nepenthes section Pyrophytae. Planta Carniv. 38, 44–45.
Cheek, M., Jebb, M., 2013. The *Nepenthes micramphora* (Nepenthaceae) group, with two new species from Mindanao, Philippines. Phytotaxa 151, 25–34. https://doi.org/10.11646/phytotaxa.151.1.2.
Cheek, M., Jebb, M., 2009. *Nepenthes* group Montanae (Nepenthaceae) in Indo-China, with *N. thai* and *N. bokor* described as new. Kew Bull. 64, 319. https://doi.org/10.1007/s12225-009-9117-3.
Cheek, M., Jebb, M., 2001. Nepenthaceae. Flora Malesiana 15, 1–157.
Chin, L., Chung, A., Clarke, C.M., 2014. Interspecific variation in prey capture behavior by co-occurring *Nepenthes* pitcher plants: Evidence for resource partitioning or sampling-scheme artifacts? Plant Signal. Behav. 9, e27930 https://doi.org/10.4161/psb.27930.
Clarke, C.M., 2001. Nepenthes of Sumatra and Peninsular Malaysia. Natural History Publications (Borneo), Kota Kinabalu.
Clarke, C.M., 1997. Nepenthes of Borneo. Natural History Publications (Borneo), Kota Kinabalu.
Clarke, C.M., Cross, A.T., Rice, B., 2018a. Conservation of carnivorous plants. In: Carnivorous Plants: Physiology, Ecology, and Evolution. Oxford University Press, pp. 375–387.
Clarke, C.M., Schlauer, J., Moran, J.A., Robinson, A.S., 2018b. Systematics and evolution of *Nepenthes*. In: Carnivorous Plants: Physiology, Ecology, and Evolution. Oxford University Press, pp. 58–69.
Cuénoud, P., Savolainen, V., Chatrou, L.W., Powell, M., Grayer, R.J., Chase, M.W., 2002. Molecular phylogenetics of Caryophyllales based on nuclear 18S rDNA and plastid rbcL, atpB, and matK DNA sequences. Am. J. Bot. 89, 132–144. https://doi.org/10.3732/ajb.89.1.132.
Czech, L., Huerta-Cepas, J., Stamatakis, A., 2017. A Critical Review on the Use of Support Values in Tree Viewers and Bioinformatics Toolkits. Mol. Biol. Evol. 34, 1535–1542. https://doi.org/10.1093/molbev/msx055.
Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24, 332–340. https://doi.org/10.1016/j.tree.2009.01.009.
Durand, E.Y., Patterson, N., Reich, D., Slatkin, M., 2011. Testing for ancient admixture between closely related populations. Mol. Biol. Evol. 28, 2239–2252. https://doi.org/10.1093/molbev/msr048.
Durvasula, A., Sankararaman, S., 2020. Recovering signals of ghost archaic introgression in African populations. Sci. Adv. 6, eaax5097. https://doi.org/10.1126/sciadv.aax5097.
Eaton, D.A.R., Ree, R.H., 2013. Inferring phylogeny and introgression using RADseq data: An example from flowering plants (*Pedicularis*: Orobanchaceae). Syst. Biol. 62, 689–706. https://doi.org/10.1093/sysbio/syt032.
Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797. https://doi.org/10.1093/nar/gkh340.
Edwards, S.V., 2009. Is a New and General Theory of Molecular Systematics Emerging? Evolution 63, 1–19. https://doi.org/10.1111/j.1558-5646.2008.00549.x.
Ellison, A.M., Adamec, L. (Eds.), 2018. Carnivorous Plants: Physiology, ecology, and evolution. Oxford University Press, Oxford, New York.
Felsenstein, J., 2004. Inferring Phylogenies, 2004th ed. Sinauer Associates Inc, Sunderland, Massachusetts, USA.
Fleischmann, A., Schlauer, J., Smith, S.A., Givnish, T.J., 2018. Evolution of carnivory in angiosperms. In: Carnivorous Plants: Physiology, Ecology, and Evolution. Oxford University Press, pp. 22–42.
Folk, R.A., Mandel, J.R., Freudenstein, J.V., 2017. Ancestral Gene Flow and Parallel Organellar Genome Capture Result in Extreme Phylogenomic Discord in a Lineage of Angiosperms. Syst. Biol. 66, 320–337. https://doi.org/10.1093/sysbio/syw083.
Folk, R.A., Soltis, P.S., Soltis, D.E., Guralnick, R., 2018. New prospects in the detection and comparative analysis of hybridization in the tree of life. Am. J. Bot. 364–375 https://doi.org/10.1002/ajb2.1018@10.1002/(ISSN)1537-2197.Tree-of-Life-virtual-issue.
Fontaine, M.C., Pease, J.B., Steele, A., Waterhouse, R.M., Neafsey, D.E., Sharakhov, I.V., Jiang, X., Hall, A.B., Catteruccia, F., Kakani, E., Mitchell, S.N., Wu, Y.-C., Smith, H.A., Love, R.R., Lawniczak, M.K., Slotman, M.A., Emrich, S.J., Hahn, M.W., Besansky, N.J., 2015. Extensive introgression in a malaria vector species complex revealed by phylogenomics. Science 347, 1258524. https://doi.org/10.1126/science.1258524.
Fukushima, K., Fang, X., Alvarez-Ponce, D., Cai, H., Carretero-Paulet, L., Chen, C., Chang, T.-H., Farr, K.M., Fujita, T., Hiwatashi, Y., Hoshi, Y., Imai, T., Kasahara, M., Librado, P., Mao, L., Mori, H., Nishiyama, T., Nozawa, M., Pálfalvi, G., Pollard, S.T., Rozas, J., Sánchez-Gracia, A., Sankoff, D., Shibata, T.F., Shigenobu, S., Sumikawa, N., Uzawa, T., Xie, M., Zheng, C., Pollock, D.D., Albert, V.A., Li, S., Hasebe, M., 2017. Genome of the pitcher plant *Cephalotus* reveals genetic changes associated with carnivory. Nat. Ecol. Evol. 1, 0059. https://doi.org/10.1038/s41559-016-0059.
Garrison, E., Marth, G., 2012. Haplotype-based variant detection from short-read sequencing. ArXiv12073907 Q-Bio.
Gilbert, K.J., Nitta, J.H., Talavera, G., Pierce, N.E., 2018. Keeping an eye on coloration: ecological correlates of the evolution of pitcher traits in the genus *Nepenthes* (Caryophyllales). Biol. J. Linn. Soc. 123, 321–337. https://doi.org/10.1093/biolinnean/blx142.
Givnish, T.J., 2010. Ecology of plant speciation. Taxon 59, 1326–1366.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. https://doi.org/10.1038/nbt.1883.
Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.-Y., Hansen, N.F., Durand, E.Y., Malaspinas, A.-S., Jensen, J.D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H.A., Good, J.M., Schultz, R., Aximu-Petri, A., Butthof, A., Höber, B., Höffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E.S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic, D., Kucan, Ž., Gušic, I., Doronichev, V.B., Golovanova, L.V., Lalueza-Fox, C., de la Rasilla, M., Fortea, J., Rosas, A., Schmitz, R.W., Johnson, P.L.F., Eichler, E.E., Falush, D., Birney, E., Mullikin, J.C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D., Pääbo, S., 2010. A draft sequence of the Neandertal genome. Science 328, 710–722. https://doi.org/10.1126/science.1188021.
Gruzdev, E.V., Mardanov, A.V., Beletsky, A.V., Ravin, N.V., Skryabin, K.G., 2018. The complete mitochondrial genome of the carnivorous flowering plant *Nepenthes* X *ventrata*. Mitochondrial DNA Part B 3, 1259–1260. https://doi.org/10.1080/23802359.2018.1532353.
Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., LeDuc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. Nat. Protoc. 8 https://doi.org/10.1038/nprot.2013.084.
Hahn, M.W., Hibbins, M.S., 2019. A three-sample test for introgression. Mol. Biol. Evol. Doi: 10.1093/molbev/msz178.
Hahn, M.W., Nakhleh, L., 2016. Irrational exuberance for resolved species trees. Evolution 70, 7–17. https://doi.org/10.1111/evo.12832.
Heubl, G., Bringmann, G., Meimberg, H., 2006. Molecular phylogeny and character evolution of carnivorous plant families in Caryophyllales - revisited. Plant Biol. 8, 821–830. https://doi.org/10.1055/s-2006-924460.
Hibbins, M.S., Gibson, M.J.S., Hahn, M.W., 2020. Determining the probability of hemiplasy in the presence of incomplete lineage sorting and introgression. bioRxiv 2020.04.15.043752. Doi: 10.1101/2020.04.15.043752.
Hibbins, M.S., Hahn, M.W., 2019. The Timing and Direction of Introgression Under the Multispecies Network Coalescent. Genetics 211, 1059–1073. https://doi.org/10.1534/genetics.118.301831.
Huerta-Cepas, J., Serra, F., Bork, P., 2016. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. Mol. Biol. Evol. 33, 1635–1638. https://doi.org/10.1093/molbev/msw046.
Huson, D.H., Scornavacca, C., 2012. Dendroscope 3: An Interactive Tool for Rooted Phylogenetic Trees and Networks. Syst. Biol. 61, 1061–1067. https://doi.org/10.1093/sysbio/sys062.
Johnson, M.G., Pokorny, L., Dodsworth, S., Botigué, L.R., Cowan, R.S., Devault, A., Eiserhardt, W.L., Epitawalage, N., Forest, F., Kim, J.T., Leebens-Mack, J.H., Leitch, I.J., Maurin, O., Soltis, D.E., Soltis, P.S., Wong, G.K., Baker, W.J., Wickett, N.J., 2019. A Universal Probe Set for Targeted Sequencing of 353 Nuclear Genes from Any Flowering Plant Designed Using k-Medoids Clustering. Syst. Biol. 68, 594–606. https://doi.org/10.1093/sysbio/syy086.
Joly, S., 2012. JML: testing hybridization from species trees. Mol. Ecol. Resour. 12, 179–184. https://doi.org/10.1111/j.1755-0998.2011.03065.x.
Juniper, B.E., Robins, R.J., Joel, D.M., 1989. The carnivorous plants. Academic Press, London.
Kitching, R.L., 2000. Food webs and container habitats: The natural history and ecology of phytotelmata. Cambridge University Press.
Kozlov, A.M., Aberer, A.J., Stamatakis, A., 2015. ExaML version 3: a tool for phylogenomic analyses on supercomputers. Bioinformatics 31, 2577–2579. https://doi.org/10.1093/bioinformatics/btv184.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054.
Lambert, S.M., Streicher, J.W., Fisher-Reid, M.C., de la Cruz, F.R.M., Martínez-Méndez, N., García-Vázquez, U.O., de Oca, A.N.M., Wiens, J.J., 2019. Inferring introgression using RADseq and DFOIL: Power and pitfalls revealed in a case study of spiny lizards (Sceloporus). Mol. Ecol. Resour. 19, 818–837. https://doi.org/10.1111/1755-0998.12972.
Leaché, A.D., Banbury, B.L., Felsenstein, J., Oca, A. nieto-Montes de, Stamatakis, A., 2015. Short tree, long tree, right tree, wrong tree: New acquisition bias corrections for inferring SNP phylogenies. Syst. Biol. 64, 1032–1047. Doi: 10.1093/sysbio/syv053.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv13033997 Q-Bio.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. Doi: 10.1093/bioinformatics/btp352.
Li, W., Godzik, A., 2006. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22, 1658–1659. https://doi.org/10.1093/bioinformatics/btl158.
Liu, L., Xi, Z., Wu, S., Davis, C.C., Edwards, S.V., 2015. Estimating phylogenetic trees from genome-scale data. Ann. N. Y. Acad. Sci. 1360, 36–53. https://doi.org/10.1111/nyas.12747.
Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536. https://doi.org/10.1093/sysbio/46.3.523.
Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L.L., Hernández-Hernández, T., 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. New Phytol. 207, 437–453. https://doi.org/10.1111/nph.13264.
Malinsky, M., Matschiner, M., Svardal, H., 2020. Dsuite - fast D-statistics and related admixture evidence from VCF files. bioRxiv 634477. Doi: 10.1101/634477.
Malinsky, M., Svardal, H., Tyers, A.M., Miska, E.A., Genner, M.J., Turner, G.F., Durbin, R., 2018. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. Nat. Ecol. Evol. 2, 1940–1955. https://doi.org/10.1038/s41559-018-0717-x.
Marabini, J., 1983. Eine neue *Nepenthes*-Hybride aus Borneo. Mitteilungen Bot. Staatssamml. Münch. 19, 449–452.
Marques, D.A., Meier, J.I., Seehausen, O., 2019. A Combinatorial View on Speciation and Adaptive Radiation. Trends Ecol. Evol. 34, 531–544. https://doi.org/10.1016/j.tree.2019.02.008.
McKain, M.R., Johnson, M.G., Uribe-Convers, S., Eaton, D., Yang, Y., 2018. Practical considerations for plant phylogenomics. Appl. Plant Sci. 6 https://doi.org/10.1002/aps3.1038.
McPherson, S., 2011. New Nepenthes - Volume 1. Redfern Natural History, Poole.
McPherson, S., 2009. Pitcher plants of the Old World. Redfern Natural History, Poole.
McVay, J.D., Hipp, A.L., Manos, P.S., 2017. A genetic legacy of introgression confounds phylogeny and biogeography in oaks. Proc R Soc B 284, 20170300. https://doi.org/10.1098/rspb.2017.0300.
Meier, J.I., Marques, D.A., Mwaiko, S., Wagner, C.E., Excoffier, L., Seehausen, O., 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. Nat. Commun. 8, 14363. https://doi.org/10.1038/ncomms14363.
Meimberg, H., Heubl, G., 2006. Introduction of a nuclear marker for phylogenetic analysis of Nepenthaceae. Plant Biol. 8, 831–840. https://doi.org/10.1055/s-2006-924676.
Meimberg, H., Wistuba, A., Dittrich, P., Heubl, G., 2001. Molecular phylogeny of Nepenthaceae based on cladistic analysis of plastid trnK intron sequence data. Plant Biol. 3, 164–175. https://doi.org/10.1055/s-2001-12897.
Mendes, F.K., Fuentes-González, J.A., Schraiber, J.G., Hahn, M.W., 2018. A multispecies coalescent model for quantitative traits. eLife 7, e36482. https://doi.org/10.7554/eLife.36482.
Merckx, V.S.F.T., Hendriks, K.P., Beentjes, K.K., Mennes, C.B., Becking, L.E., Peijnenburg, K.T.C.A., Afendy, A., Arumugam, N., de Boer, H., Biun, A., Buang, M. M., Chen, P.-P., Chung, A.Y.C., Dow, R., Feijen, F.A.A., Feijen, H., Soest, C.F., Geml, J., Geurts, R., Gravendeel, B., Hovenkamp, P., Imbun, P., Ipor, I., Janssens, S. B., Jocqué, M., Kappes, H., Khoo, E., Koomen, P., Lens, F., Majapun, R.J., Morgado, L.N., Neupane, S., Nieser, N., Pereira, J.T., Rahman, H., Sabran, S., Sawang, A., Schwallier, R.M., Shim, P.-S., Smit, H., Sol, N., Spait, M., Stech, M., Stokvis, F., Sugau, J.B., Suleiman, M., Sumail, S., Thomas, D.C., van Tol, J., Tuh, F.Y. Y., Yahya, B.E., Nais, J., Repin, R., Lakim, M., Schilthuizen, M., 2015. Evolution of endemism on a young tropical mountain. Nature 524, 347–350. https://doi.org/10.1038/nature14949.
Mirarab, S., Bayzid, M.S., Warnow, T., 2016. Evaluating summary methods for multilocus species tree estimation in the presence of incomplete lineage sorting. Syst. Biol. 65, 366–380. https://doi.org/10.1093/sysbio/syu063.
Mirarab, S., Reaz, R., Bayzid, M.S., Zimmermann, T., Swenson, M.S., Warnow, T., 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30, i541–i548. https://doi.org/10.1093/bioinformatics/btu462.
Moran, J.A., Clarke, C.M., 2010. The carnivorous syndrome in *Nepenthes* pitcher plants. Plant Signal. Behav. 5, 644–648.
Morley, R.J., 2011. Cretaceous and Tertiary climate change and the past distribution of megathermal rainforests, in: Tropical Rainforest Responses to Climatic Change, Springer Praxis Books. Springer, Berlin, Heidelberg, pp. 1–34. Doi: 10.1007/978-3-642-05383-2_1.
Mullins, J.T., 2000. Molecular Systematics of Nepenthaceae. PhD Thesis. University of Reading, Reading, UK.
Murphy, B., Forest, F., Barraclough, T., Rosindell, J., Bellot, S., Cowan, R., Golos, M., Jebb, M., Cheek, M., 2019. A phylogenomic analysis of *Nepenthes* (Nepenthaceae). Mol. Phylogenet. Evol. 106668 https://doi.org/10.1016/j.ympev.2019.106668.
Nadeau, N.J., Martin, S.H., Kozak, K.M., Salazar, C., Dasmahapatra, K.K., Davey, J.W., Baxter, S.W., Blaxter, M.L., Mallet, J., Jiggins, C.D., 2013. Genome-wide patterns of divergence and gene flow across a butterfly radiation. Mol. Ecol. 22, 814–826. https://doi.org/10.1111/j.1365-294X.2012.05730.x.
Nauheimer, L., Cui, L., Clarke, C., Crayn, D.M., Bourke, G., Nargar, K., 2019. Genome skimming provides well resolved plastid and nuclear phylogenies, showing patterns of deep reticulate evolution in the tropical carnivorous plant genus *Nepenthes* (Caryophyllales). Aust. Syst. Bot. 32, 243–254. https://doi.org/10.1071/SB18057.
One Thousand Plant Transcriptomes Initiative, 2019. One thousand plant transcriptomes and the phylogenomics of green plants. Nature 574, 679–685. Doi: 10.1038/s41586-019-1693-2.
Paradis, E., Schliep, K., 2019. APE 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35, 526–528. https://doi.org/10.1093/bioinformatics/bty633.
Patterson, N.J., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., Reich, D., 2012. Ancient Admixture in Human History. Genetics genetics.112.145037. Doi: 10.1534/genetics.112.145037.
Pease, J.B., Haak, D.C., Hahn, M.W., Moyle, L.C., 2016. Phylogenomics reveals three sources of adaptive variation during a rapid radiation. PLOS Biol. 14, e1002379 https://doi.org/10.1371/journal.pbio.1002379.
Pease, J.B., Hahn, M.W., 2015. Detection and polarization of introgression in a five-taxon phylogeny. Syst. Biol. 64, 651–662. https://doi.org/10.1093/sysbio/syv023.
Peng, H.S., Clarke, C.M., 2015. Prey capture patterns in *Nepenthes* species and natural hybrids - are the pitchers of hybrids as effective at trapping prey as those of their parents? Carniv. Plant Newsl. 44, 62–79.
Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLOS ONE 7, e37135. https://doi.org/10.1371/journal.pone.0037135.
Pickrell, J.K., Pritchard, J.K., 2012. Inference of population splits and mixtures from genome-wide allele frequency data. PLOS Genet. 8, e1002967 https://doi.org/10.1371/journal.pgen.1002967.
Plagnol, V., Wall, J.D., 2006. Possible Ancestral Structure in Human Populations. PLoS Genet. 2 https://doi.org/10.1371/journal.pgen.0020105.
Puritz, J.B., Hollenbeck, C.M., Gold, J.R., 2014. dDocent: A RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ 2, e431. https://doi.org/10.7717/peerj.431.
Racimo, F., Sankararaman, S., Nielsen, R., Huerta-Sánchez, E., 2015. Evidence for archaic adaptive introgression in humans. Nat. Rev. Genet. 16, 359–371. https://doi.org/10.1038/nrg3936.
Renner, T., Specht, C.D., 2011. A sticky situation: Assessing adaptations for plant carnivory in the Caryophyllales by means of stochastic character mapping. Int. J. Plant Sci. 172, 889–901. https://doi.org/10.1086/660882.
Rieseberg, L.H., Willis, J.H., 2007. Plant speciation. Science 317, 910–914. https://doi.org/10.1126/science.1137729.
Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584. https://doi.org/10.7717/peerj.2584.
Rokas, A., Williams, B.L., King, N., Carroll, S.B., 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425, 798–804. https://doi.org/10.1038/nature02053.
Schwallier, R., Gravendeel, B., de Boer, H., Nylinder, S., van Heuven, B.J., Sieder, A., Sumail, S., van Vugt, R., Lens, F., 2017. Evolution of wood anatomical characters in Nepenthes and close relatives of Caryophyllales. Ann. Bot. Doi: 10.1093/aob/mcx010.
Schwallier, R., Raes, N., de Boer, H.J., Vos, R.A., van Vugt, R.R., Gravendeel, B., 2016. Phylogenetic analysis of niche divergence reveals distinct evolutionary histories and climate change implications for tropical carnivorous pitcher plants. Divers. Distrib. 22, 97–110. https://doi.org/10.1111/ddi.12382.
Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., Peichel, C.L., Saetre, G.-P., Bank, C., Brännström, Å., Brelsford, A., Clarkson, C.S., Eroukhmanoff, F., Feder, J.L., Fischer, M.C., Foote, A.D., Franchini, P., Jiggins, C.D., Jones, F.C., Lindholm, A.K., Lucek, K., Maan, M.E., Marques, D.A., Martin, S.H., Matthews, B., Meier, J.I., Möst, M., Nachman, M.W., Nonaka, E., Rennison, D.J., Schwarzer, J., Watson, E.T., Westram, A.M., Widmer, A., 2014. Genomics and the origin of species. Nat. Rev. Genet. 15, 176–192. https://doi.org/10.1038/nrg3644.
Sickel, W., Grafe, T.U., Meuche, I., Steffan-Dewenter, I., Keller, A., 2016. Bacterial diversity and community structure in two Bornean *Nepenthes* species with differences in nitrogen acquisition strategies. Microb. Ecol. 1–16 https://doi.org/10.1007/s00248-015-0723-3.
Smith, S.A., Donoghue, M.J., 2008. Rates of Molecular Evolution Are Linked to Life History in Flowering Plants. Science 322, 86–89. https://doi.org/10.1126/science.1163197.
Solís-Lemus, C., Ané, C., 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. PLOS Genet. 12, e1005896 https://doi.org/10.1371/journal.pgen.1005896.
Solís-Lemus, C., Bastide, P., Ané, C., 2017. PhyloNetworks: A package for phylogenetic networks. Mol. Biol. Evol. 34, 3292–3298. https://doi.org/10.1093/molbev/msx235.
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.
Sukumaran, J., Holder, M.T., 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics 26, 1569–1571. https://doi.org/10.1093/bioinformatics/btq228.
Takeuchi, Y., Chaffron, S., Salcher, M.M., Shimizu-Inatsugi, R., Kobayashi, M.J., Diway, B., von Mering, C., Pernthaler, J., Shimizu, K.K., 2015. Bacterial diversity and composition in the fluid of pitcher plants of the genus *Nepenthes*. Syst. Appl. Microbiol. 38, 330–339. https://doi.org/10.1016/j.syapm.2015.05.006.
Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A., Kumar, S., 2012. Estimating divergence times in large molecular phylogenies. Proc. Natl. Acad. Sci. 109, 19333–19338. https://doi.org/10.1073/pnas.1213199109.
Tamura, K., Tao, Q., Kumar, S., 2018. Theoretical foundation of the RelTime method for estimating divergence times from variable evolutionary rates. Mol. Biol. Evol. Advance Access. Doi: 10.1093/molbev/msy044.

Tonini, J., Moore, A., Stern, D., Shcheglovitova, M., Ortí, G., 2015. Concatenation and Species Tree Methods Exhibit Statistically Indistinguishable Accuracy under a Range of Simulated Conditions. PLoS Curr. 7 https://doi.org/10.1371/currents.tol.34260cc27551a527b124ec5f6334b6be.

Tsitrone, A., Kirkpatrick, M., Levin, D.A., 2003. A Model for Chloroplast Capture. Evolution 57, 1776–1782. https://doi.org/10.1111/j.0014-3820.2003.tb00585.x.

Twyford, A.D., Ennos, R.A., 2012. Next-generation hybridization and introgression. Heredity 108, 179–189. https://doi.org/10.1038/hdy.2011.68.

van Dongen, S.M., 2001. Graph clustering by flow simulation. PhD thesis. Utrecht University, Utrecht, The Netherlands.

Vernot, B., Akey, J.M., 2014. Resurrecting Surviving Neandertal Lineages from Modern Human Genomes. Science 343, 1017–1021. https://doi.org/10.1126/science.1245938.

Wagner, C.E., Keller, I., Wittwer, S., Selz, O.M., Mwaiko, S., Greuter, L., Sivasundar, A., Seehausen, O., 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Mol. Ecol. 22, 787–798. https://doi.org/10.1111/mec.12023.

Walker, J.F., Yang, Y., Moore, M.J., Mikenas, J., Timoneda, A., Brockington, S.F., Smith, S.A., 2017. Widespread paleopolyploidy, gene tree conflict, and recalcitrant relationships among the carnivorous Caryophyllales. Am. J. Bot. 104, 858–867. https://doi.org/10.3732/ajb.1700083.

Wen, D., Yu, Y., Zhu, J., Nakhleh, L., 2018. Inferring Phylogenetic Networks Using PhyloNet. Syst. Biol. 67, 735–740. https://doi.org/10.1093/sysbio/syy015.

Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., Ayyampalayam, S., Barker, M.S., Burleigh, J.G., Gitzendanner, M.A., Ruhfel, B.R., Wafula, E., Der, J.P., Graham, S.W., Mathews, S., Melkonian, M., Soltis, D.E., Soltis, P.S., Miles, N.W., Rothfels, C.J., Pokorny, L., Shaw, A.J., DeGironimo, L., Stevenson, D.W., Surek, B., Villarreal, J.C., Roure, B., Philippe, H., dePamphilis, C. W., Chen, T., Deyholos, M.K., Baucom, R.S., Kutchan, T.M., Augustin, M.M., Wang, J., Zhang, Y., Tian, Z., Yan, Z., Wu, X., Sun, X., Wong, G.K.-S., Leebens-Mack, J., 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc. Natl. Acad. Sci. 111, E4859–E4868. https://doi.org/10.1073/pnas.1323926111.

Widmer, A., Lexer, C., Cozzolino, S., 2009. Evolution of reproductive isolation in plants. Heredity 102, 31–38. https://doi.org/10.1038/hdy.2008.69.

Wu, M., Kostyun, J.L., Hahn, M.W., Moyle, L.C., 2018. Dissecting the basis of novel trait evolution in a radiation with widespread phylogenetic discordance. Mol. Ecol. 27, 3301–3316. https://doi.org/10.1111/mec.14780.

Yang, Y., Moore, M.J., Brockington, S.F., Soltis, D.E., Wong, G.K.-S., Carpenter, E.J., Zhang, Y., Chen, L., Yan, Z., Xie, Y., Sage, R.F., Covshoff, S., Hibberd, J.M., Nelson, M.N., Smith, S.A., 2015. Dissecting molecular evolution in the highly diverse plant clade Caryophyllales using transcriptome sequencing. Mol. Biol. Evol. 32, 2001–2014. https://doi.org/10.1093/molbev/msv081.

Yang, Y., Smith, S.A., 2014. Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: improving accuracy and matrix occupancy for phylogenomics. Mol. Biol. Evol. 31, 3081–3092. https://doi.org/10.1093/molbev/msu245.

Yao, G., Jin, J.-J., Li, H.-T., Yang, J.-B., Mandala, V.S., Croley, M., Mostow, R., Douglas, N.A., Chase, M.W., Christenhusz, M.J.M., Soltis, D.E., Soltis, P.S., Smith, S. A., Brockington, S.F., Moore, M.J., Yi, T.-S., Li, D.-Z., 2019. Plastid phylogenomic insights into the evolution of Caryophyllales. Mol. Phylogenet. Evol. 134, 74–86. https://doi.org/10.1016/j.ympev.2018.12.023.

Zhang, Chi, Ogilvie, H.A., Drummond, A.J., Stadler, T., 2018a. Bayesian inference of species networks from multilocus sequence data. Mol. Biol. Evol. 35, 504–517. https://doi.org/10.1093/molbev/msx307.

Zhang, Chao, Rabiee, M., Sayyari, E., Mirarab, S., 2018b. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19, 153. https://doi.org/10.1186/s12859-018-2129-y.