Do Reticulate Relationships Between Tropical Trees Drive Diversification? Insights from *Inga* (Fabaceae)

Rowan J. Schley[1]*, Rosalía Piñeiro[2], James A. Nicholls[3], Flávia Fonseca Pezzini[4], Catherine Kidner[4,5], Audrey Farbos[6], Karen Moore[6], Jens J. Ringelberg[7], Alex D. Twyford[4,8], Kyle G. Dexter[4,7], R. Toby Pennington[1,4]

[1] *Department of Geography, University of Exeter, Laver Building, North Park Road, Exeter, Devon, UK*

[2] *Grupo de Investigación en Biología Evolutiva, CICA, Departamento de Biología, Universidade da Coruña, A Coruña, Spain*

[3] *Australian National Insect Collection, GPO Box 1700, Canberra, ACT 2601, Australia.*

[4] *Tropical Diversity Section, Royal Botanic Garden Edinburgh, Edinburgh, UK*

[5] *Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh, UK*

[6] *University of Exeter Sequencing Service, Geoffrey Pope Building, Stocker Road, Exeter, Devon, UK*

[7] *School of Geosciences, University of Edinburgh, Edinburgh, UK*

[8] *Institute of Ecology and Evolution, University of Edinburgh, Edinburgh, UK*

* *Corresponding author: rowan.schley (at) gmail.com*

ABSTRACT

Evolutionary radiations underlie much of the species diversity of life on Earth, particularly within the world's most species-rich tree flora – that of the Amazon rainforest. Hybridisation catalyses many radiations by generating genetic and phenotypic novelty that promote rapid speciation, but the influence of hybridisation on Amazonian tree radiations has been little studied. We address this using the ubiquitous, species-rich neotropical tree genus *Inga*, which typifies rapid radiations of rainforest trees. We assess patterns of gene tree incongruence to ascertain whether hybridisation catalysed rapid radiation in *Inga*. Given the importance of insect herbivory in structuring rainforest tree communities (and hence the potential for hybridisation to promote adaptation through admixture of defence traits), we

also test whether introgression of loci underlying chemical defences against herbivory facilitated rapid speciation in *Inga*. Our phylogenomic analyses of 189/288 *Inga* species using >1300 target capture loci showed widespread introgression in *Inga* and closely related genera. Specifically, we found widespread phylogenetic incongruence explained by introgression, with phylogenetic networks recovering multiple introgression events across *Inga* and related genera. In addition, most defence chemistry loci showed evidence of positive selection and marginally higher levels of introgression. Interestingly, we recovered around 20% admixed variation between multiple *Inga* species, a similar proportion to that found in other radiations catalysed by 'ancient' hybridisation. Overall our results suggest that introgression has occurred widely over the course of *Inga's* history, likely facilitated by extensive dispersal across Amazonia, and that in some cases introgression of chemical defence loci may facilitate adaptation in *Inga*.

Keywords: Hybridisation; Introgression; Radiation; Diversification; Phylogenomics; Amazon; Rainforest; Trees

Rapid evolutionary radiations that generate exceptionally species-rich groups are a fundamental component of biodiversity (Hughes et al. 2015). Hybridisation (interbreeding between species) underlies rapid radiation in many taxa (e.g. Barrier et al. 1999; Meier et al. 2017; Lamichhaney et al. 2018) because it allows admixture of existing genetic variation, which can generate genomic and phenotypic novelty more rapidly than *de novo* mutation alone (Rieseberg and Willis 2007). Under this model, described as 'combinatorial' speciation (Marques et al. 2019), genetic and phenotypic variation generated by hybridisation may confer adaptation to new environments (e.g. adaptive introgression and/or transgressive segregation (Rieseberg et al. 1999; Suarez-Gonzalez et al. 2018)) or lead to re-sorting of

intrinsic incompatibilities (e.g. Bateson-Dobzhansky-Muller incompatibilities) that promote reproductive isolation and hence rapid speciation (Schumer et al. 2015). A relatively small number of hybridisation events at the 'base' of a clade can catalyse the rapid diversification of many species from a few initial lineages (Seehausen 2004), as demonstrated in Lake Victoria's cichlid fishes (Meier et al. 2017) and Hawaiian silverswords (Barrier et al. 1999).

It is possible to detect genetic admixture and infer past reticulation events in a clade through examining gene tree conflict (Doyle 1992; Naciri and Linder 2015). Specifically, the proportions of different conflicting gene tree topologies can provide evidence for whether introgression (interspecific transfer of genetic material following persistent hybridisation) or incomplete lineage sorting (ILS) underlie observed incongruence (e.g. Green et al. 2010; Durand et al. 2011; Pease et al. 2018). This incongruence can also help estimate the relative genetic contributions of progenitor lineages to introgressant descendants (Patterson et al. 2012). There is a growing body of work that explores incongruence to better understand introgression across the tree of life, particularly in plants (e.g. in oaks (McVay et al. 2017) and willows (Wagner et al. 2020)), but only recently have such studies been undertaken in the most species-rich flora on Earth - that of the tropical Americas (e.g. Morales Briones et al. 2018; Schley et al. 2020; Larson et al. 2021; reviewed in Schley et al. 2022).

The flora of neotropical rainforests is remarkable in its species diversity (Antonelli and Sanmartín 2011; Ulloa Ulloa et al. 2017; Raven et al. 2020), particularly at local scales – there are more tree species in a single hectare of the Amazon (c. 655 spp. (Valencia et al. 2004)) than in all of Europe (c. 454 spp. (Rivers et al. 2019)). Many species-rich neotropical plant groups arose through recent, rapid radiations (e.g. Erkens et al. 2007; Koenen et al. 2015), but the influence of hybridisation on tropical plant radiations has been little studied, and virtually not at all in tropical rainforest trees (reviewed in Abbott (2017)). The prevailing view, based largely on morphological patterns, has been that hybrids between tropical tree

species are exceptionally rare (e.g. Ashton 1969), but this is challenged by recent genomic data for a few Amazonian tree genera (e.g. *Brownea* (Schley et al. 2020); *Eschweilera* (Larson et al. 2021)).

The tree genus *Inga* is widespread and abundant in neotropical rainforests, and was the first documented example of recent, rapid radiation in the Amazonian tree flora (Richardson et al. 2001). *Inga* exhibits the highest diversification rate of any Amazonian tree genus (Baker et al. 2014), with *ca.* 300 species arising in the last *ca.* 10 Ma (Ringelberg et al. 2023). Similar recent, rapid radiation events in other genera gave rise to a large portion of the Amazonian tree flora, with over half the tree species in the Amazon belonging to genera with at least 100 species (Dexter and Chave 2016). *Inga* is an ideal study system with which to test the combinatorial speciation model as a catalyst of radiations in the Amazon due to the large volume of previous work examining the diversification, ecology and biogeography of the group (Richardson et al. 2001; Kursar et al. 2009; Dexter et al. 2017; Forrister et al. 2019) coupled with its ubiquity, abundance and species diversity in the Amazon (Pennington, T. D. 1997). Previous phylogenetic work on *Inga,* based on Sanger sequencing of relatively few species, revealed low resolution of species-level relationships (Richardson et al. 2001; Kursar et al. 2009; Dexter et al. 2010), with resolution improving in later phylogenomic studies using 22 *Inga* species (Nicholls et al. 2015). Here we generate the most comprehensive phylogenetic tree of *Inga* to date, comprising >1300 loci and 189 species, to help understand how hybridisation contributed to diversification of the genus.

One factor that could promote hybridisation more widely amongst *Inga* species is their remarkable level of co-occurrence in local rainforest communities. Up to 19 species of *Inga* can coexist in 1ha, and such high local diversity is also typical of many other species-rich Amazonian tree genera (e.g. *Protium* and *Eschweilera* (Valencia et al. 1994)). These dispersal-assembled local communities of *Inga* (Dexter et al. 2017) are largely structured by

insect herbivore pressure, such that co-occurring *Inga* species differ more in their chemical defences against herbivores than would be expected by chance (Kursar et al. 2009; Endara et al. 2022). This is because growing in proximity to congeners with similar chemical defences results in higher herbivore pressure and hence lower survival probability of individuals (Forrister et al. 2019). This suggests that herbivore pressure plays a role in divergence and coexistence between *Inga* species.

Divergent adaptation in chemical defence traits is well documented in *Inga* over evolutionary timescales (Forrister et al. 2023), likely being driven by the negative density-dependent processes mentioned above over ecological timescales (Forrister et al. 2019). Higher densities of conspecifics with the same defences leads to increased mortality from herbivores (Janzen 1970; Connell 1971), suggesting that it is adaptive to possess rare defence chemistry phenotypes, as fewer local herbivores can overcome them. Therefore, the transfer of defence chemistry loci between *Inga* species via hybridisation, followed by positive selection on those loci, may be predicted to facilitate colonisation of new ecological spaces and eventually lead to speciation. To our knowledge this has never been investigated from a genomic perspective. Therefore, here we aim to test the significance of hybridisation in *Inga*'s diversification and we hypothesise that:

1) Reticulation occurred at the base of the *Inga* radiation, as well as throughout it, and hence may have contributed to the rapid diversification of the genus throughout its history
2) Genomic loci with the strongest evidence for introgression and natural selection will be related to chemical defences against herbivores

MATERIALS & METHODS

Taxon Sampling and DNA Sequencing

We performed phylogenomic analyses on target capture sequencing data from 189 of the 282 accepted *Inga* species (67%), 109 of which were sequenced for this study, and 80 of which were taken from previous work (one from Nicholls et al. 2015; 79 from Ringelberg et al. 2023). Our sampling was based on a list of all accepted *Inga* species compiled using the World Checklist of Vascular Plants (WCVP 2020) (as of June 2021) and a monograph of *Inga* (Pennington, T. D. 1997). Preliminary analyses identified three nested clades within *Inga*, which reflect major transitions in substitution rate (the ‘fast clade’), shifts in the trait of indumentum as a defence for young shoots (the ‘hairy clade’) and shifts in ecology and defence chemistry (the 'Red hair' clade) (Coley et al. 2018; Ringelberg et al. 2023), and so downstream analyses are divided into these three nested, monophyletic groups.

To contextualise our *Inga* analyses at broader phylogenetic scales, we collated sequence data from many outgroup species generated by previous phylogenomic studies (Koenen et al. 2020; Ringelberg et al. 2022). We included all seven genera and 32 other species from the ‘Inga clade’ (*sensu* Koenen et al. 2020) within which *Inga* is nested. In addition, we included a further 73 species comprising all 42 genera from the ‘Ingoid clade’ (sensu Koenen et al. 2020) within which the Inga clade is nested. Finally, we included 22 species from the broader Mimosoid legume clade, giving a total of 127 outgroup species. A list of accessions sampled including species, sampling location, data source and voucher information is detailed in Supplementary Table S1, available on Dryad.

DNA was extracted from 20 mg of dried leaf material with the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) using modifications described in Nicholls et al. (2015). DNA library preparation, enrichment and sequencing were carried out either by Arbor BioSciences

(Ann Arbor, MI, USA) or the University of Exeter sequencing service (Exeter, UK) following the NEBnext Ultra II FS protocol (New England Biolabs, Ipswich, MA, USA). Targeted bait capture was performed with a subfamily-specific 'Mimobaits' bait set (Nicholls et al. 2015; Koenen et al. 2020) using the MyBaits protocol v.2 and 3 (Arbor Biosciences, Ann Arbor, MI, USA). The Mimobaits set targets 1320 loci, including 113 genes associated with anti-herbivore defence chemistry in *Inga* (hereafter 'defence chemistry' loci). Other loci targeted by the Mimobaits set are 'single-copy phylogenetically informative' loci (1044), 'differentially expressed' loci (109) and 'miscellaneous' loci (54), that latter of which are unannotated but contain phylogenetic signal. Enriched libraries were sequenced using the NovaSeq 6000 platform with a paired-end 150bp run on two S1 flow cell lanes.

Sequence Assembly, Trimming and Alignment

All analyses were conducted on the UK Crop Diversity Bioinformatics HPC Resource. DNA sequencing reads were quality-checked with FASTQC 0.11.3 (Andrews 2010) and were trimmed using TRIMMOMATIC 0.3.6 (Bolger et al. 2014) to remove adapter sequences and to quality-filter reads. TRIMMOMATIC settings permitted < 2 mismatches, a palindrome clip threshold of 30 and a simple clip threshold of 10. Bases with a quality score < 28 and reads shorter than 36 bases long were removed from the dataset. Following quality-filtering, reads were mapped to target loci using BWA 0.7.17 (Li and Durbin 2009), loci were assembled using SPADES 3.11.1 (Bankevich et al. 2012) with a coverage cut-off of 5× and exons were extracted with EXONERATE (Slater and Birney 2005), all of which are implemented in the HYBPIPER pipeline 1.2 (Johnson et al. 2016). Recovery of loci was visualised using the 'gene_recovery_heatmap.R' script distributed with HYBPIPER. To improve the signal:noise ratio and remove possible paralogues from our *Inga* dataset we used the Putative Paralogs

Detection pipeline 1.0.1, (PPD; https://github.com/Bean061/putative_paralog; Zhou et al. (2022)), described in Supplementary Methods on Dryad.

We conducted all subsequent analyses on three data subsets to include a broad range of phylogenetic scales and to assess the influence of paralogy on our analyses. The first dataset included one accession for each of 189 *Inga* species with *Zygia* sp. '*mediana*' as the outgroup (dataset 'Singlesp'). The second dataset included one accession from each of 127 outgroup species and proportional per-clade sampling of 50 *Inga* species (dataset 'Outgroup'), with species chosen based on locus recovery. The third and final dataset comprised the same 189 accessions as the 'Singlesp' *Inga* dataset, but loci were assembled and cleaned using the PPD pipeline, with paralogous loci removed (dataset 'PPD').

Targeted loci for each of the three datasets were aligned by gene region using 1000 iterations in MAFFT 7.453 (Katoh and Standley 2013) with the '*—adjustdirectionaccurately*' flag to incorporate reversed sequences. These alignments were then cleaned using the '*-automated1'* flag in trimAl 1.3 (Capella-Gutiérrez et al. 2009), and realigned with MAFFT using the '*—auto'* flag. This resulted in 1305 refined alignments for the 'Singlesp' dataset, 1044 for the 'Outgroup' dataset and 1267 loci for the 'PPD' dataset. Alignment summaries were then generated with AMAS (Borowiec 2016).

Phylogenomic Analyses

Gene trees were inferred for each locus alignment across the three datasets using IQ-TREE (Nguyen et al. 2015) by selecting the best-fit substitution model (*-MFP*) while reducing the impact of severe model violations (*-bnni*) with 1000 ultrafast bootstrap replicates. Following this, a 'species tree' was generated based on the best-scoring IQtrees using ASTRALMP 5.15.5 under the default parameters (Zhang et al. 2018) for all three

datasets. Following concatenation of all locus alignments within the 'Singlesp' data subset using AMAS, we used the same parameters as above to infer a phylogenetic tree from the concatenated supermatrix using IQ-TREE. Finally, we visualised shared genetic variation among species by building neighbour net plots with uncorrected P-distances in SPLITSTREE v4.14.6 (Huson and Bryant 2005) for each of the three datasets.

Analysing Incongruence and Reticulation

To assess incongruence among our gene trees, we estimated three metrics implemented in the QUARTET SAMPLING method 1.3.1 (https://www.github.com/fephyfofum/quartetsampling; Pease et al. (2018)) based on each dataset's ASTRAL species tree, using 100 replicate runs. For each node, we estimated Quartet Concordance (QC) to assess whether there was incongruence, Quartet Differential (QD) to assess whether one incongruent topology was favoured and Quartet Informativeness (QI) to assess whether data were sufficiently informative to distinguish well-supported incongruence from lack of signal (QI not shown as all datasets were informative).

Having assessed incongruence across our three datasets, we then visually investigated gene tree conflict of certain higher-level relationships highlighted in the Quartet Sampling analysis with DISCOVISTA (Sayyari et al. 2018) (clades defined in Supplementary Table S2, available on Dryad). We interpreted gene tree conflict as "low" if the proportion of gene trees supporting the species tree topology was >30 percent higher than for both alternative topologies after Kuhnhäuser et al. (2021). We interpreted gene tree conflict as suggesting introgression if one conflicting topology was above the 33% 'equal frequency' threshold.

We then assessed whether the incongruence we inferred was caused by introgression or ILS using the *Dtrios* function in DSUITE (Malinsky et al. 2021). We used Patterson's D-

statistic (i.e. the 'ABBA-BABA' test; Green et al. 2010; Durand et al. 2011) and estimated the proportion of shared variation between species using the F_4 ratio (Patterson et al. 2012), where for both metrics values closer to 1 indicate more introgression. We generated an input VCF file for *Dtrios* for each of the three datasets by calling SNPs *de-novo* using BWA 0.7.17, SAMTOOLS 1.13 (Danecek et al. 2021) and BCFTOOLS 1.13 (Li 2011) as in the 'Speciation Genomics' github (https://speciationgenomics.github.io), using our target bait set sequences as the reference. The resulting VCFs were filtered to contain sites with > 8x coverage. We used the ASTRAL species tree topology from each of the three datasets to inform taxon relationships for each set of tests and assessed significance of each test using 20 jackknife resampling runs. Finally, we visualized our D-statistic and F_4 ratio estimates with Ruby scripts available from https://github.com/mmatschiner. Since F_4 ratios may be correlated among closely-related species, we additionally estimated the F_{branch} statistic (Malinsky et al. 2018) for each combination of taxa to control for this, and plotted them using the dtools.py utility distributed with DSUITE.

To investigate species relationships and model historical migration (i.e. introgression) events, we used TREEMIX v1.13 (Pickrell and Pritchard 2012a) on representative subsets of our three datasets, down-sampled due to computational limitations. Input VCFs were filtered to remove all sites with missing data using VCFTOOLS *--max missing* (Danecek et al. 2011) as suggested for TREEMIX (Pickrell and Pritchard 2012b). We ran TREEMIX with between 0 and 10 migration edges (M), using 5 separate iterations per M value. Since TREEMIX was designed for population-level sampling, but was the only tractable network method for a dataset of this size, we used the *-noss* flag to account for having one accession per species. The best-fit number of migration edges was chosen based on the highest Δm value with the package *OptM* (Fitak 2021) in R v 4.2.1 (R Development Core Team 2013), following which the iteration with the highest log likelihood was plotted using the R scripts in the Speciation

Genomics github (https://speciationgenomics.github.io/Treemix/). Details of SNP calling and accession selection for TREEMIX are found in Supplementary Methods on Dryad.

Assessing Per-Locus Introgression and Selection

We assessed whether defence chemistry loci experienced elevated introgression and selection relative to other loci in our target capture dataset. To do this we first estimated the per-locus proportion of introgression between subclades in our target-capture data using the f_{dM} statistic (Malinsky et al. 2015) implemented in DSUITE. Given that we observed the most introgression in the 'hairy clade' of *Inga*, we performed the analysis on five subsets of our 'Singlesp' dataset, containing five different combinations of taxa to represent all subclades within the Hairy clade (selected subclades shown in Supplementary Fig. S1; subclade selection described in Supplementary Methods, available on Dryad). For each subset, we calculated f_{dM} using nonoverlapping windows of 50 informative SNPs with a rolling mean of one window. We took the absolute values of all f_{dM} scores, as we were only interested in comparing the magnitude of introgression across loci, and calculated a mean f_{dM} score per-locus for downstream analyses.

We then assessed whether each of our target-capture loci experienced positive selection on at least one branch and at least one site using BUSTED (Murrell et al. 2015). We prepared our target capture alignments for BUSTED analysis by trimming non-homologous sequence fragments (i.e. intronic regions captured either side of exons), masking misaligned amino acid residues and producing codon-aware alignments using OMM_MACSE (Ranwez et al. 2011; Ranwez et al. 2021).

Using these codon-aware alignments, we first tested for the presence of selection in each locus across all taxa in our *Inga* 'Singlesp' dataset. We accounted for false positives by

adjusting the selection test *P*-values output by BUSTED with a 5% FDR (false discovery rate) in R (Benjamini and Hochberg 1995). We assessed whether there were associations between locus selection result ('under selection' if the BUSTED FDR *P*-value < 0.05) and locus annotation ('defence chemistry', 'differentially expressed', 'single-copy phylogenetically informative' and 'miscellaneous') using χ^2 tests in R. We also visualised associations between selection result and locus annotation using the R package *corrplot* (Wei et al. 2017). We repeated the BUSTED analysis using identical parameters for five down-sampled alignments comprising the same 'hairy clade' taxon subsets for which we estimated f_{dM} (subclades defined in Supplementary Fig. S1, available on Dryad).

Finally, we used analysis of covariance (ANCOVA) to assess whether variation in our per-locus f_{dM} estimates was explained by interactions between three variables. The variables were locus annotation, selection result and the length of the locus alignment (to control for differences in the number of sites between loci). Response variables were square-root transformed to improve normalcy, and the heteroscedasticity of residuals was examined using the *plot()* function in R. Box plots of per-locus f_{dM} estimates were generated using *ggplot2* (Wickham 2016) in R, with locus annotation and selection result as grouping variables. In total, we estimated both the proportion of introgression and selection for 1167 loci.

RESULTS

Phylogenomic Analyses

We achieved a mean of 72.05% reference length recovery onto the *Mimobaits* bait sequences (78.99%, excluding outgroups, across 1305 loci). Overall, the 'Singlesp' dataset had 1.53×10^6 sites, with a mean of 0.31% variable sites and 2.30% missing data across all loci. The PPD dataset had 1.89×10^6 sites, with a mean of 0.30% variable sites and 3.02%

missing data. Finally, the outgroup dataset had $1.44x10^6$ sites, with a mean of 0.66% variable sites and 12.45% missing data. The 'Outgroup' dataset comprised only of the 1044 'single-copy phylogenetically informative' loci sequenced by previous studies (Koenen et al. 2020; Ringelberg et al. 2022). Heatmaps showing % recovery per locus are available in Supplementary Fig. S2, along with summaries of sites, variability and missing data per locus in Table S3, both available on Dryad.

Our ASTRAL analyses indicated that most bipartitions were well supported across the three datasets, with local posterior probabilities (LPP) >0.8 (Supplementary Fig. S3ai; Fig. S3b; Fig. S3c, available on Dryad) and a quartet score of 0.47, 0.69 and 1.36 for the single-accession-per-species, PPD and outgroup datasets, respectively. Within *Inga*, relationships between 'Red hair' clade species were less well supported than the rest of the tree, and the concatenated IQTREE analysis of the Singlesp dataset recovered a nearly identical topology to the ASTRAL tree with high support (BS >90, Supplementary Fig. S3ai; Fig. S3aii, available on Dryad). The removal of paralogs with PPD did not materially influence the resolution or topology of the ASTRAL *Inga* tree (Supplementary Fig. 3c, available on Dryad). Within the outgroup tree, most generic splits are well supported (LPP >0.9), and *Inga* was monophyletic. However, three *Zygia* species clustered with other genera (*Z. inundata* and *Z. sabatieri* with *Inga; Z. ocumarensis* with *Macrosamanea* (Supplementary Fig. S3b, available on Dryad)).

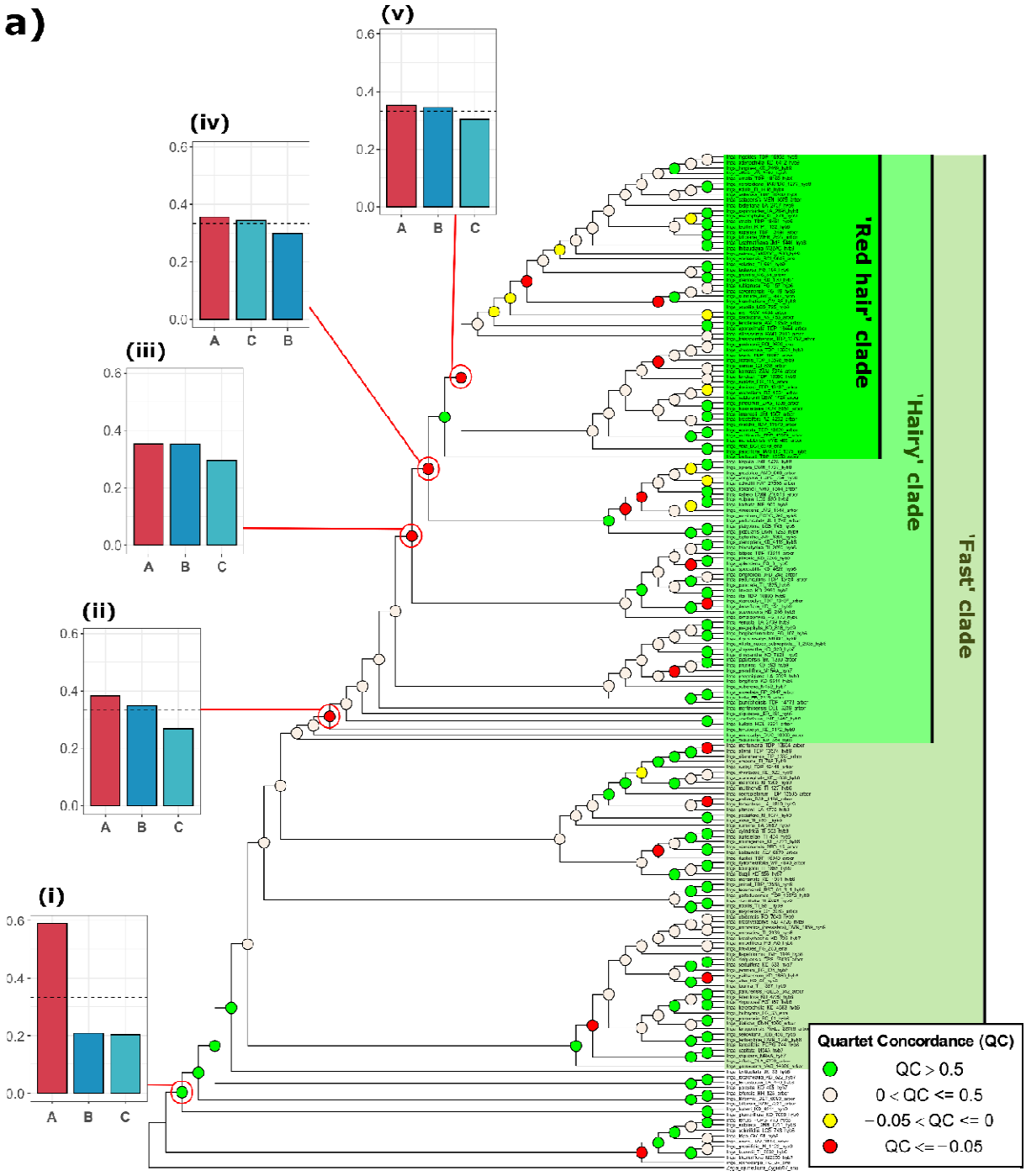
Genetic variation visualized with SPLITSTREE showed many shared splits within *Inga, Zygia* and *Macrosamanea,* as well as between genera (Supplementary Fig. S4a, Fig. S4b available on Dryad). Shared splits are particularly evident within the 'Hairy' clade of *Inga*, along with the 'Red hair' clade nested within it (Supplementary Fig. S4a, available on Dryad). SPLITSTREE analysis of the Outgroup dataset showed that *Zygia inundata* and *Z. sabatieri* clustered with *Inga*, while *Z. ocumarensis* clustered with *Macrosamanea*

(Supplementary Fig. S4b, available on Dryad). Removal of 38 paralogous loci with PPD resulted in a different clustering of *Inga* species, with the long branch leading to *Inga gereauana* bisecting the Red hair and Hairy clades (Supplementary Fig. S4c, available on Dryad).

Incongruence is Common Within and Among Genera

Our analyses recovered evidence of phylogenetic incongruence both within and between Ingoid clade genera (Fig.1a; Fig. 1b). Negative Quartet Concordance (QC) scores and DiscoVista analyses suggested that a single conflicting topology was disproportionately represented at several incongruent nodes, particularly within the ‘Hairy’ and 'Red hair' clades of *Inga* (Fig. 1a nodes ii-v; see also QD trees in Supplementary Fig. S6a, available on Dryad). However, most nodes in the *Inga* ASTRAL tree showed multiple conflicting topologies (i.e., QC scores between 0 and 0.5; Fig. 1a). Removal of paralogs with PPD did not materially influence the results (Supplementary Fig. S5; Fig. S6c). All nodes were shown to be informative with QI scores.

Most nodes in the outgroup tree recovered much higher QC scores, indicative of more phylogenetic concordance (Fig. 1b). This was with the exception of a few nodes involving *Zygia*, *Macrosamanea*, *Abarema* and the clade containing *Jupunba*, that had negative QC scores (Fig. 2a nodes ii-iv; QD trees in Supplementary Fig. S6b, available on Dryad) or showed multiple conflicting topologies (QC <= 0 => 0.5). For both trees, the deepest divergences revealed higher proportions of concordant topologies (Fig. 1a node i; Fig. 1b node i).


a)
(v)
(iv)
(iii)
(ii)
(i)
A B C
'Red hair' clade
'Hairy' clade
'Fast' clade
Quartet Concordance (QC)
QC > 0.5
0 < QC <= 0.5
−0.05 < QC <= 0
QC <= −0.05

342

343

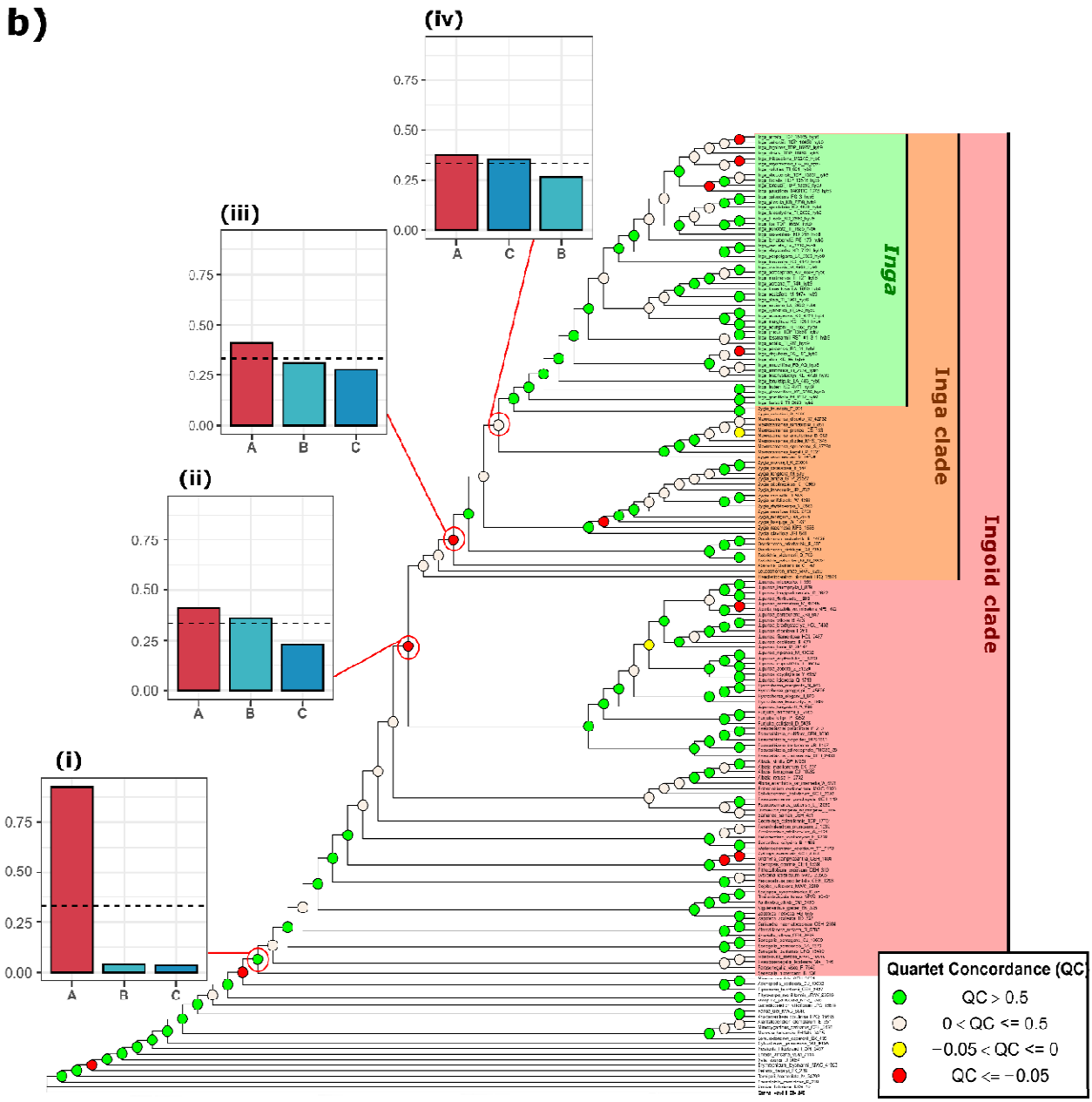


Figure 1

a: *Inga* single accession per species ASTRAL tree with QC values plotted on each node. Nodes of interest are additionally annotated with DiscoVista plots, showing the relative proportions of different discordant topologies. Quartet frequencies are represented as bar graphs, with red bars (left) representing the main topology from the ASTRAL analysis, and with blue and turquoise bars (middle and right) representing alternative topologies. Dashed horizontal lines mark the expectation for equal frequencies of the three possible topologies (Y = 0.333), i.e. maximal gene tree conflict. Node i indicates low proportions of both conflicting alternative topologies. Nodes ii-v indicate one major conflicting topology.

b: *Inga* outgroup ASTRAL tree with QC values plotted on each node. Nodes of interest are additionally annotated with DiscoVista plots, showing the relative proportions of different discordant topologies. Quartet frequencies are represented as bar graphs, with red bars (left) representing the main topology from the ASTRAL analysis, and with blue and turquoise bars (middle and right) representing alternative topologies. Dashed

horizontal lines mark the expectation for equal frequencies of the three possible topologies (Y = 0.333), i.e. maximal gene tree conflict. Node i indicates low proportions of both conflicting alternative topologies. Nodes ii - iv indicate one major conflicting topology.

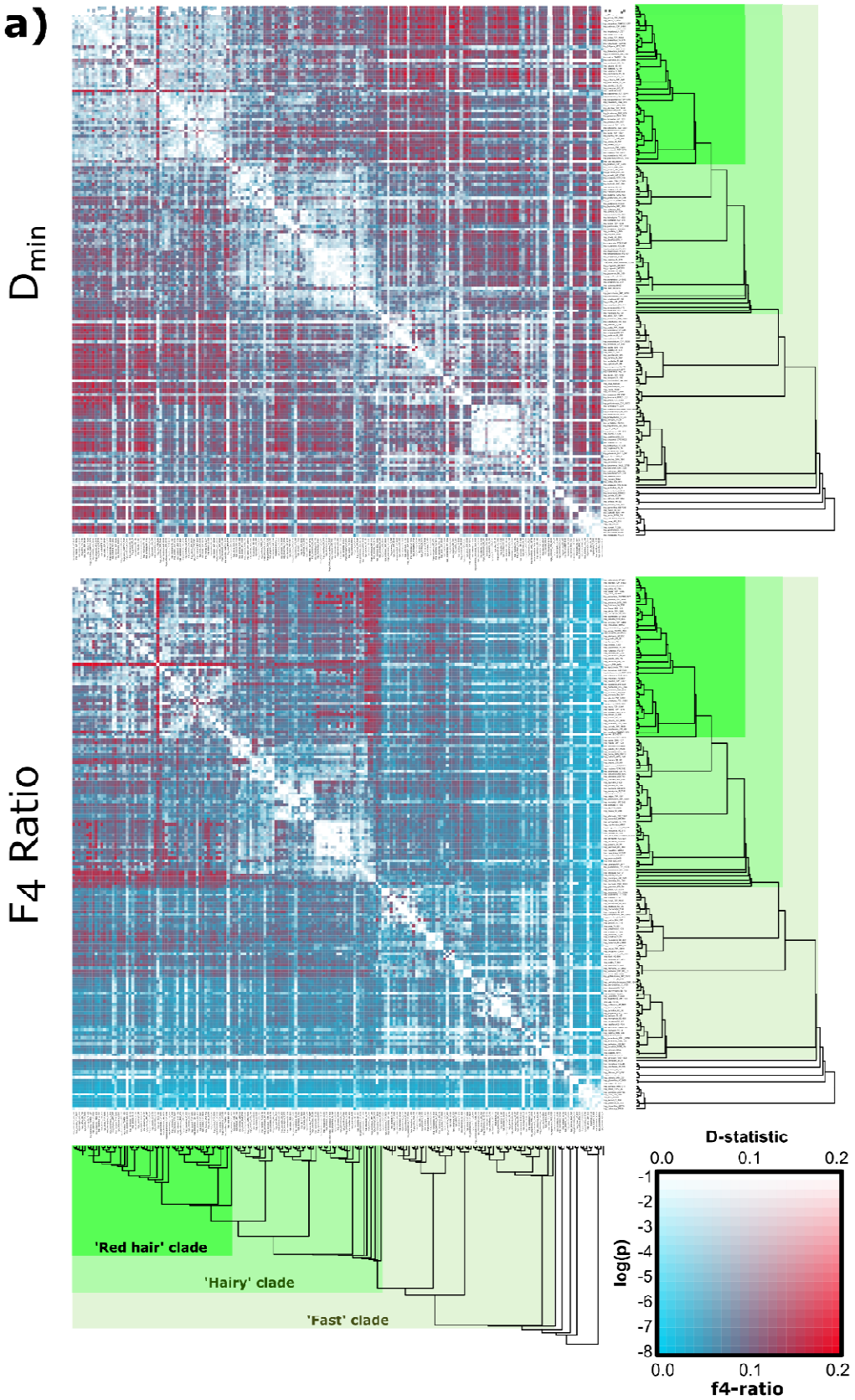
Reticulation Occurs at Multiple Phylogenetic Scales

The overrepresentation of one incongruent topology that we inferred for several nodes (Fig. 1aii-v; Fig. 1bii-iv) was reinforced by the high *D*-statistics, F_4 ratios and F_{branch} scores that we calculated for all three datasets. This suggests reticulation contributed to incongruence at these nodes. Across all methods (D statistics, F_4 ratios and F_{branch}) introgression is evident across entire clades for all three datasets.

Within *Inga*, *D*-statistics up to 0.2 were observed most frequently in the 'Red hair' clade and the sister group to the 'Hairy' clade (Fig. 2a) although significant *D*-statistics were evident across the *Inga* tree, particularly within some early-diverging *Inga* lineages. Similarly, the highest F_4 ratios we recovered suggested that species in the 'Red hair' clade shared up to 20% of their sequence variation with the earliest diverging lineages in the 'Hairy' clade. Only the F_4 ratio varied significantly between ‘Singlesp’ and ‘PPD’ runs, and revealed higher proportions of introgression in the ‘PPD’ run between the ‘Red hair’ and early-diverging ‘Hairy’ subclades (Supplementary Fig. S7, available on Dryad). Estimates of shared ancestry made using F_{branch}, which controls for correlated introgression in closely-related lineages, showed again that species in the ‘Red hair’ clade experienced the most introgression with other lineages, both within and outside the ‘Red hair’ clade – particularly *Inga cardozana* and *Inga pauciflora* (Supplementary Fig. S8a, available on Dryad).

In the broader Ingoid clade, both *D*-statistics and F_4 ratios suggested that most significant introgression occurred within *Zygia*, as well as between *Zygia*, *Inga* and other Ingoid clade genera (D = 0.4, F_4 ratio = ca. 20%; Fig. 2b). Both *D*-statistics and F_4 ratios showed that significant introgression also occurred between *Jupunba* and *Hydrochorea*, as

well as within *Jupunba* and *Osodendron* (Fig. 2b). F_{branch} scores recovered more limited evidence of introgression, with the highest scores occurring within closely related species pairs in *Macrosamanea*, *Robrichia*, *Zygia* and *Jupunba* (Supplementary Fig. S8b, available on Dryad).


a)
Dmin
F4 Ratio
'Red hair' clade
'Hairy' clade
'Fast' clade
D-statistic
0.0
0.1
0.2
log(p)
-1
-2
-3
-4
-5
-6
-7
-8
f4-ratio

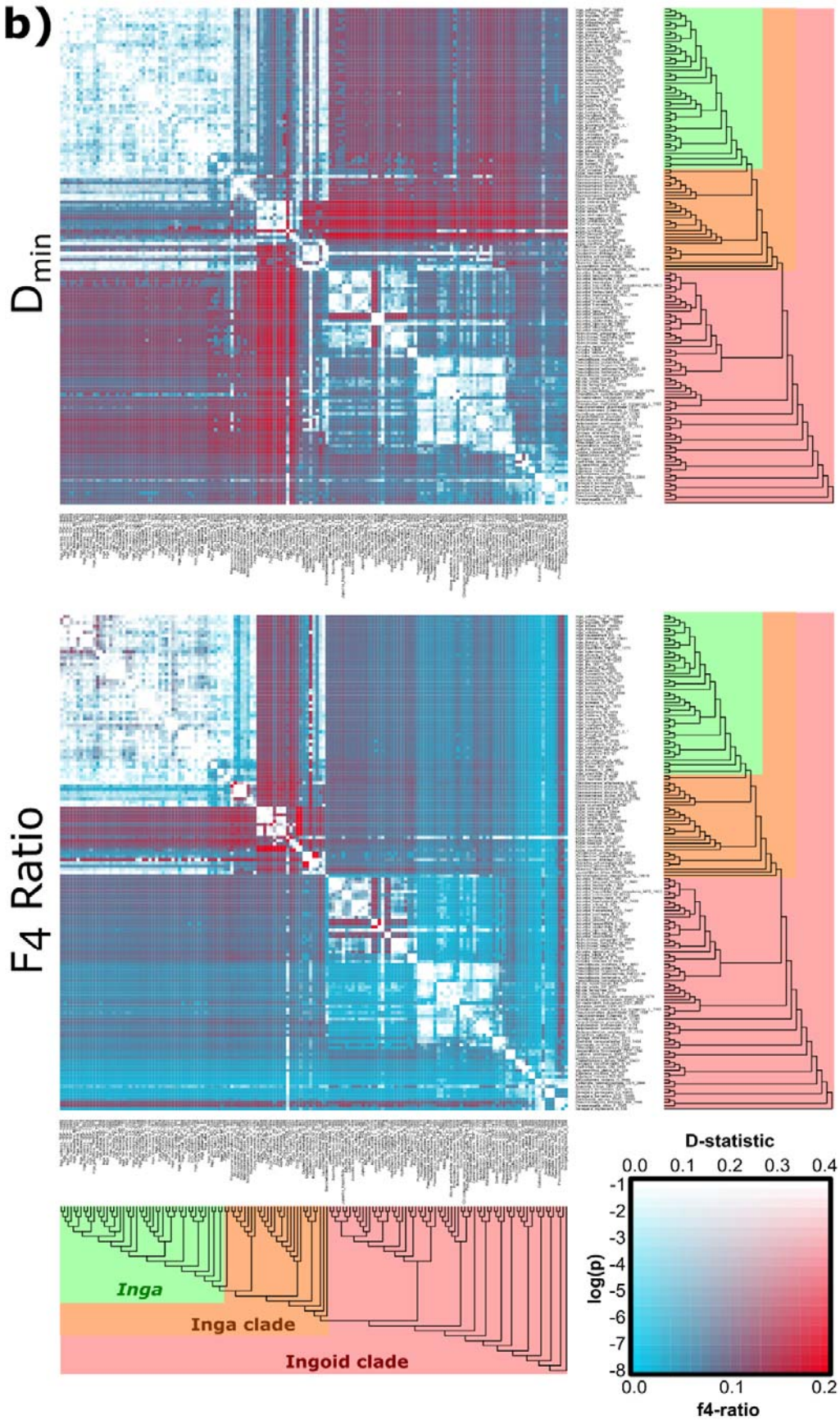

b)
Dmin
F4 Ratio
Inga
Inga clade
Ingoid clade
D-statistic
0.0 0.1 0.2 0.3 0.4
log(p)
-1 -2 -3 -4 -5 -6 -7 -8
0.0 0.1 0.2
f4-ratio

Figure 2

a: Heatmaps of minimum per-triplet D statistic (D_{min}, above) and F_4 ratios (below) plotted for the single-accession-per-species *Inga* dataset. Taxa P2 and P3 are displayed on the *x*- and *y*-axes in the same order as in Figure 1a. The colour of each square signifies the *D*-statistic or F_4 ratio estimate (blue = low estimate; red = high estimate). The saturation of these colours represents the significance for that test (see log(P) in inset box, bottom right). Clades are marked in the same colours as Fig. 1a, representing the 'Fast', 'Hairy' and 'Red hair' clades.

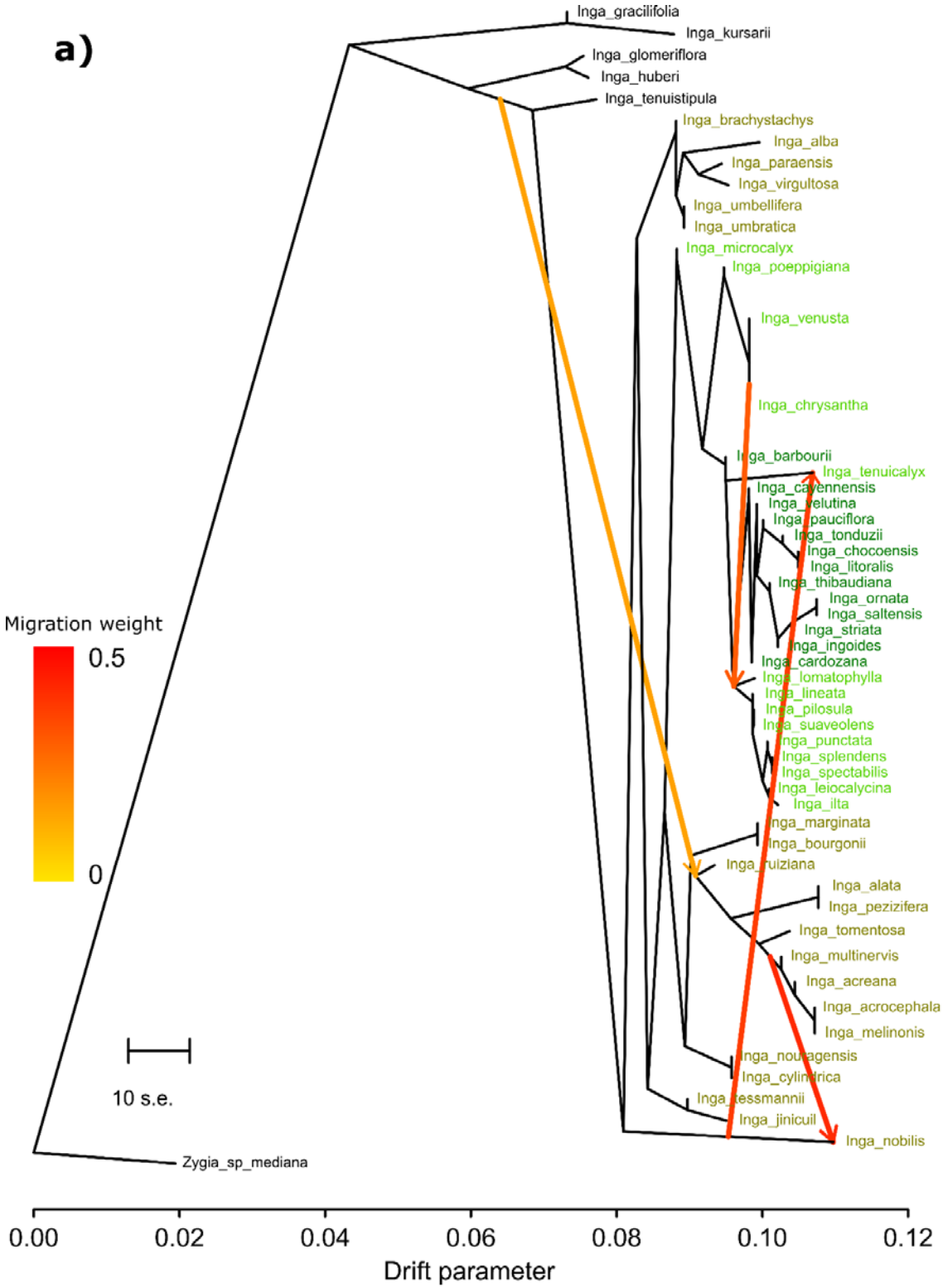
b: Heatmaps of minimum per-triplet D statistic (D_{min}, above) and F_4 ratios (below) plotted for the Outgroup dataset. Taxa P2 and P3 are displayed on the *x*- and *y*-axes in the same order as in Figure 1b. The colour of each square signifies the *D*-statistic or F_4 ratio estimate (blue = low estimate; red = high estimate). The saturation of these colours represents the significance for that test (see log(P) in inset box, bottom right). Clades are marked in the same colours as Fig. 1b, representing the 'Ingoid clade', 'Inga clade' and the genus *Inga*.

Our TREEMIX analyses largely recovered the same clades as our ASTRAL trees (Fig. 1a; Fig. 1b) for both the 'Singlesp' and 'Outgroup' datasets (Fig. 3a; Fig 3b). Within *Inga*, around 96% of variance was explained when no migration (i.e. introgression) was allowed (Supplementary Fig. S9a; Table S4, available on Dryad). Adding migration events to the tree produced better fit, with four edges providing the largest improvement in likelihood (Δm; Fig. 3a). We inferred moderate migration from early-diverging *Inga* lineages into the 'Fast' clade (from the first branch after *I. huberi/glomeriflora* into the base of *I. ruiziana*). We also inferred more significant migration within the 'Fast' clade (the group subtended by *I. tomentosa* into *I. nobilis*), from the 'Fast' clade into the 'Hairy' clade (*I. nobilis* into *I. tenuicalyx*) and from the 'Hairy' clade into the base of the 'Red hair' clade (*I. chrysantha* into the group subtended by *I. cayennensis*).

Removal of paralogs from the *Inga* dataset with PPD resulted in a similar amount of variance being explained with no migration edges (ca. 96%), however 6 migration edges were best supported by TREEMIX for the paralog-free runs (Supplementary Fig. S9c; Supplementary Table S4, available on Dryad). This run inferred introgression events involving some of the same branches as the Singlesp dataset (*Inga nobilis*; the base of the subclade containing *I. leiocalycina;* the base of the Hairy clade; the stem lineage of *I. tenuistipula/huberi/glomeriflora*) as well as two others (the base of the *Inga umbellifera*

subclade and the split between *I. tessmanii/I. jinicuil*) shown in Supplementary Fig. S10 (available on Dryad).

Among Ingoid clade ’Outgroup’ species a high proportion of variance was explained (99.6%) when no migration was allowed (Supplementary Table S4; Fig. S9b, available on Dryad). Adding migration events to the tree produced slightly better fit, with two edges providing the largest improvement in likelihood (Δm) (Fig. 3b). The inferred migration events occurred from *Jupunba oxyphyllidia* into *Jupunba abbottii* and from *Zygia obolingoides* into *Zygia morongii*.


a)
Inga_gracilifolia
Inga_kursarii
Inga_glomeriflora
Inga_huberi
Inga_tenuistipula
Inga_brachystachys
Inga_alba
Inga_paraensis
Inga_virgultosa
Inga_umbellifera
Inga_umbratica
Inga_microcalyx
Inga_poeppigiana
Inga_venusta
Inga_chrysantha
Inga_barbourii
Inga_tenuicalyx
Inga_cayennensis
Inga_velutina
Inga_pauciflora
Inga_tonduzii
Inga_chocoensis
Inga_litoralis
Inga_thibaudiana
Inga_ornata
Inga_saltensis
Inga_striata
Inga_ingoides
Inga_cardozana
Inga_lomatophylla
Inga_lineata
Inga_pilosula
Inga_suaveolens
Inga_punctata
Inga_splendens
Inga_spectabilis
Inga_leiocalycina
Inga_ilta
Inga_marginata
Inga_bourgonii
Inga_ruizlana
Inga_alata
Inga_pezizifera
Inga_tomentosa
Inga_multinervis
Inga_acreana
Inga_acrocephala
Inga_melinonis
Inga_nouragensis
Inga_cylindrica
Inga_tessmannii
Inga_jinicuil
Inga_nobilis
Zygia_sp_mediana
Migration weight
0.5
0
10 s.e.
0.00
0.02
0.04
0.06
0.08
0.10
0.12
Drift parameter

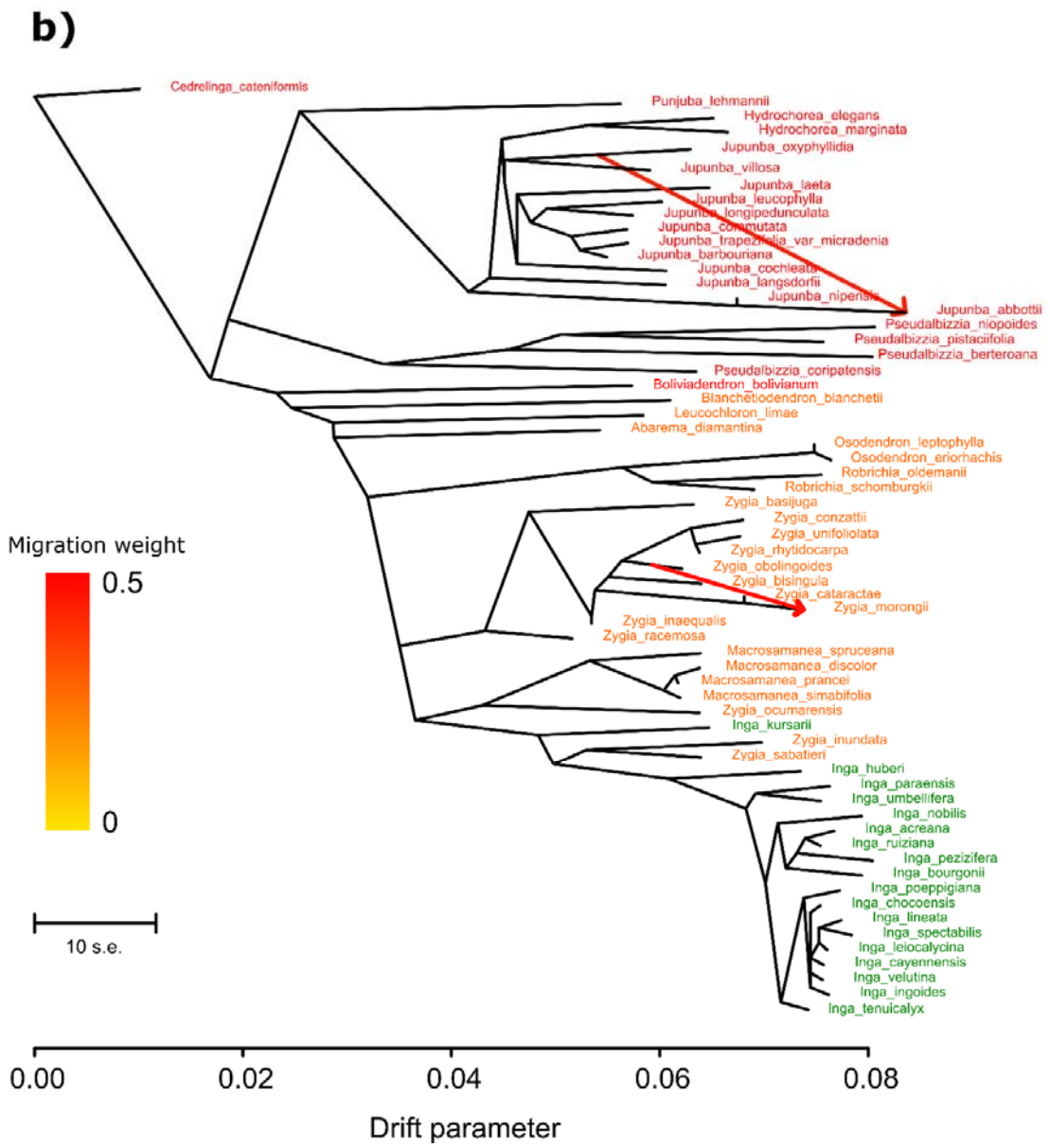


Figure 3

a: *Treemix* plot for the single-accession-per-species *Inga* dataset with M=4 migration edges. Text colours of taxa reflect the clade that they belong to (Brown = 'Fast' clade, light green = 'Hairy' clade, dark green = 'Red hair' clade). Arrows indicate ‘migration’ (i.e. introgression) events between different branches, and the colour of the arrow depicts the ‘migration weight’ (i.e. the proportion of shared variation). The drift parameter corresponds to the amount of genetic drift along each population. The scale bar shows ten times the average standard error of the entries in the sample covariance matrix.

b: *Treemix* plot for the outgroup dataset with M=2 migration edges. Text colours of taxa reflect the clade that they belong to (Red = Ingoid clade, orange = Inga clade, green = *Inga*). Arrows indicate ‘migration’ (i.e. introgression) events between different branches, and the colour of the arrow depicts the ‘migration weight’ (i.e. the proportion of shared variation). The drift parameter corresponds to the amount of genetic drift along each population. The scale bar shows ten times the average standard error of the entries in the sample covariance matrix.

Levels of Introgression and Selection Differ Between Loci in Inga

Our estimates of introgression and selection varied widely across the 1167 target capture loci we analysed. Per-locus introgression (f_{dM}) varied between 0.0003 – 0.77 across all five data subsets, approaching the maximum f_{dM} score of 1 in some datasets (Table S5; Supplementary Fig. S11, available on Dryad). The 'poeppigiana + hairy clade small + leiocalycina' subset produced the highest f_{dM} scores overall (Table S5; all clade subsets illustrated in Supplementary Fig. S1, available on Dryad). Across the analyses, the highest f_{dM} scores were observed in the single-copy phylogenetically informative loci, which were the most numerous (Supplementary Fig. S11, available on Dryad).

In total, BUSTED inferred that 768 out of 1167 loci (65.7%) showed evidence of positive selection in at least one taxon after accounting for false discovery rate. Across all taxa, the χ^2 test revealed that there was a significant association between locus annotation and selection result ($\chi^2 = 9.575$, df = 3, *P*-value = 0.022). Furthermore, across all taxa a greater number of 'defence chemistry' loci showed evidence of selection than the null expectation, whereas fewer 'differentially expressed' and 'miscellaneous' loci than the null expectation showed evidence of selection (Supplementary Fig. S12, available on Dryad).

Similarly, in the data subset containing the 'leiocalycina + vulpina + red hair' subclades, more 'defence chemistry' loci showed evidence of selection than the null expectation, whereas fewer 'miscellaneous' loci showed evidence of selection than the null expectation (Fig. 4a). However, none of the five data subsets for which we estimated selection returned significant χ^2 results when this association was tested statistically (Supplementary Table S6, available on Dryad).

Our ANCOVA analyses showed significant differences in f_{dM} score means between different locus annotations in the 'leiocalycina + vulpina + red hair' data subset (P=<0.01, d.f. = 3, F =

6.76) (Fig. 4b; 4c showing clade subset tested). Moreover, ANCOVA revealed a marginal difference in f_{dM} means between locus annotations when loci experienced selection (P= 0.05, df = 3, F = 2.60) (Fig. 4b). For both ‘defence chemistry’ and ‘differentially expressed loci’, the mean f_{dM} score was markedly higher in loci inferred to be under selection – i.e., loci with more evidence of introgression tended to show more evidence of selection (Fig. 4b). By contrast, ‘miscellaneous’ loci inferred to be under selection had markedly lower f_{dM} scores (Fig. 4b). ANCOVA did not recover significant differences in f_{dM} scores for any of the other four data subsets (Supplementary Table S7, available on Dryad).

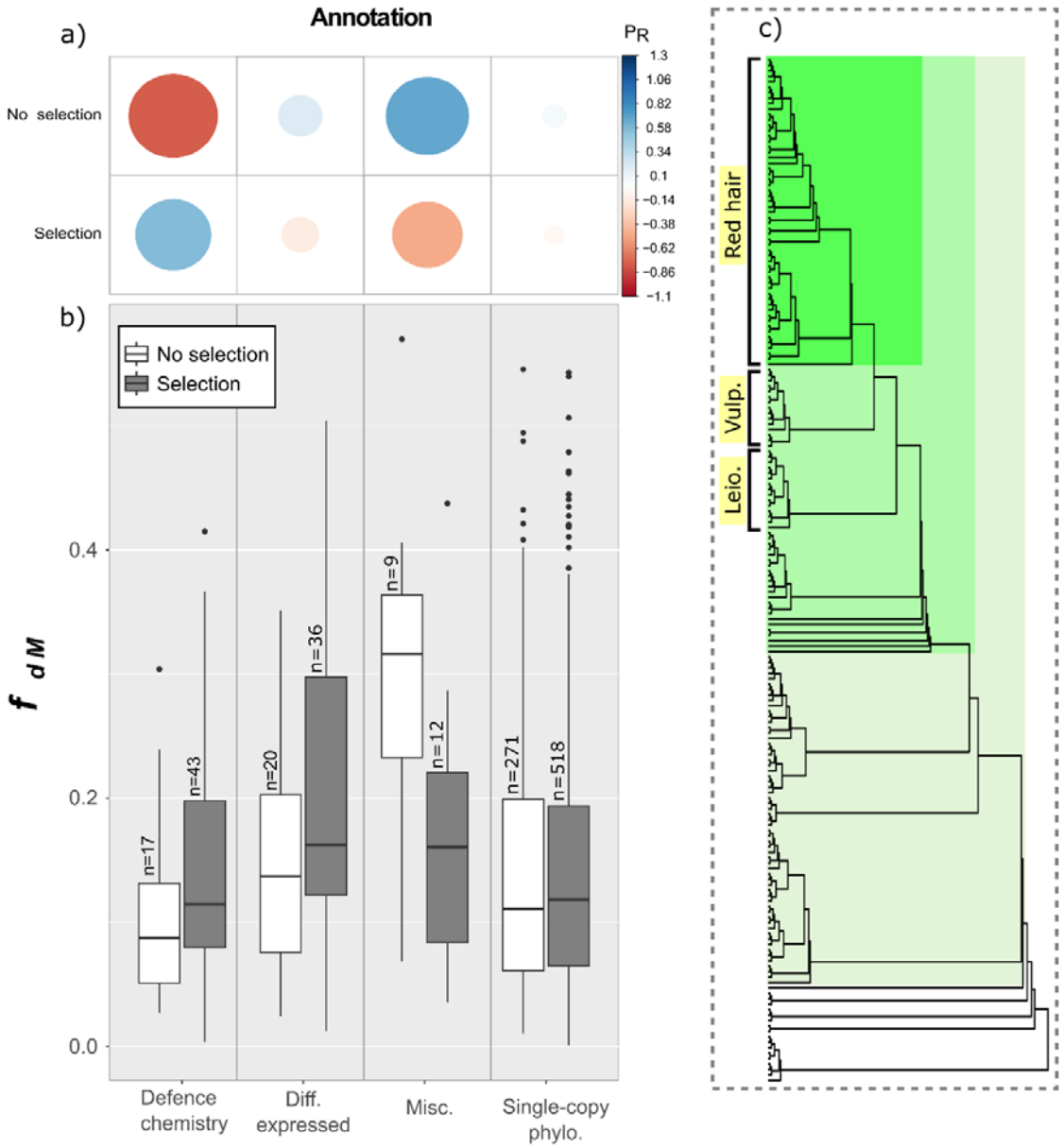


Figure 4: Per-locus introgression (f_{dM}) and selection (BUSTED) results summarised for the data subset containing the 'leiocalycina + vulpina + red hair' subclades.

a) Correlation plot showing associations between selection ('under selection' if FDR-corrected *P*-value of BUSTED analysis <0.05) and locus annotation (Defence chemistry; Differentially expressed; Miscellaneous; Single-copy phylogenetically informative). The size of each circle represents the Pearson's residual from χ^2 analysis (P_R in the figure legend). The colour of each circle indicates whether a proportion is above (positive = blue) or below (negative = red) the null expectation of χ^2.

b) Box plots of per-locus proportion of introgression (f_{dM}), grouped by selection result ('under selection' if the *P*-value of BUSTED analysis <0.05) and locus annotation (*x*-axis). Sample sizes are shown vertically above each corresponding box. Medians are shown as horizontal lines in each box.

c) The ASTRAL phylogenetic tree of *Inga* indicating which subclades were used for these analyses – the leiocalycina, vulpina and red hair subclades and highlighted and labelled.

DISCUSSION

Diversification of Inga *and the Ingoid Clade*

Our analyses recovered well-supported phylogenetic trees for *Inga* as well as the broader Ingoid clade within which it is nested (LPP >0.8, BS >90; Supplementary Fig. S3ai-ii; Fig. S3b; Fig. S3c available on Dryad). The phylogenetic tree of *Inga* inferred by this study marks a great advance in the resolution of inter-species relationships when compared to previous phylogenetic work using fewer loci and species (e.g. Richardson et al. 2001; Kursar et al. 2009; Dexter et al. 2010; Nicholls et al. 2015). Thus, our phylogenetic tree provides the best available framework to investigate the role of hybridisation in this diverse group.

Our analyses revealed three nested clades within *Inga* (Fig. 1a). The deepest-level group is the ‘fast’ clade, at the base of which there was a substitution rate shift inferred by other studies (Ringelberg et al. 2023). Nested within the ‘Fast’ clade is the ‘Hairy’ clade, in which many species possess indumentum on young leaves as a defence against herbivores (Agrawal 1999; Coley et al. 2018). Finally, nested within the ‘Hairy’ clade is the 'Red hair' clade, containing species that possess dense indumentum on young leaves and highly diverse defence chemistry (Coley et al. 2018). SPLITSTREE inferred a high degree of shared genetic variation within the rapidly diversifying 'Hairy' and 'Red hair' clades of *Inga*, as well as within and between other Ingoid clade genera (*Macrosamanea*, *Zygia* and *Jupunba*) (Supplementary Fig. S4a; S4b, available on Dryad). Of particular interest was the grouping of multiple *Zygia* species with other South American genera (e.g. *Zygia ocumarensis* with *Macrosamanea; Zygia sabatieri* and *Z. inundata* with *Inga*). The non-monophyly of *Zygia* was first described by Ferm et al. (2019), and further cases of generic non-monophyly within the Ingoid clade have been highlighted by Ringelberg et al. (2022).

Phylogenetic Incongruence is Widespread and Influenced by Introgression

We found phylogenetic incongruence both within and between Ingoid clade genera using Quartet Concordance (QC) scores and DISCOVISTA (Fig.1a; Fig. 1b). Several nodes of the phylogenetic trees inferred for the Singlesp and Outgroup datasets showed that one conflicting topology was overrepresented at several incongruent nodes (Fig. 1a nodes ii-v; Fig 1b nodes ii-iv; see also QD trees in Supplementary Fig. S6a; Fig. S6b, available on Dryad), which is indicative of introgression at these nodes (Wendel and Doyle 1998). However, many deeper nodes showed that several conflicting topologies were equally frequent (QC scores 0 < 0.5), which is indicative of incomplete lineage sorting (ILS) as is common in rapidly-diversifying clades such as *Inga* (Degnan and Rosenberg 2009).

Our explicit tests for introgression using *D*-statistic, F_4 ratios and F_{branch} (Fig. 2a; Fig. 2b; Supplementary Fig. S8a, S8b, available on Dryad) inferred widespread introgression across *Inga* as well as other rapidly-diversifying Amazonian genera (namely *Zygia* and *Jupunba*). Our F_4 ratio estimates suggested that up to 20% of genetic variation was shared between some *Inga* species, with up to 30% of shared variation being evident in the more conservative F_{branch} analyses (Supplementary Fig. S8a, available on Dryad). Interestingly, this is a similar proportion of shared variation as reported in other radiations that were catalysed by 'ancient' hybridisation (e.g., Lake Victoria cichlid fish (Meier et al. 2017)). TREEMIX inferred at least four migration events in *Inga* (Fig. 3a) and at least two migration events between smaller Amazonian tree genera (*Zygia* and *Macrosamanea*) (Fig. 3b; Supplementary Table S4, available on Dryad). These events represent minimum estimates since not all species were included in the TREEMIX analyses, and many more such events are likely to have occurred over *Inga's* evolutionary history given the extensive evidence of introgression we recovered.

Inga may be representative of other large genera in neotropical rainforests, which account for half of all Amazonian tree diversity and show high sympatry in local communities (e.g., *Eschweilera*, *Protium*, (Bermingham and Dick 2001)). Assuming *Inga* is representative of these other groups, our analyses suggest that introgression is more widespread than previously thought in species-rich Amazonian tree genera (Ashton 1969), particularly in the most widespread, abundant and rapidly-diversifying genera. While introgression is particularly prominent in certain taxa within these radiations (e.g. *Inga cardozana* and *Inga tenuicalyx*) we also inferred introgression across *Inga*, between species that are distantly related (e.g. Fig. 3a; Fig. 3b). Our findings are best described by the syngameon hypothesis of adaptive radiation (Seehausen 2004), which posits that admixture between species in geographical proximity occurs continually in an ongoing radiation, often generating functional diversity and promoting speciation. Rapid radiations may also arise when admixture between recently divergent lineages at the base of a clade triggers rapid radiation (Seehausen 2004), but we found comparatively little evidence for this hypothesis since introgression was not concentrated in early-diverging *Inga* lineages.

Periodical hybridisation within a syngameon may be adaptive for Amazonian tree species, which are typified by large, disparate populations (ter Steege et al. 2013). This is because periodical hybridisation can elevate genetic diversity and prevent Allee effects, such as inbreeding, at low population densities (Cannon and Lerdau 2015; Cannon and Lerdau 2019). *Inga* species are known to be highly dispersible, with the entirety of Amazonia acting as a species pool for the assembly of local *Inga* communities (Dexter et al. 2017). As such, the introgression we inferred in *Inga* (as well as other in tree genera such as *Zygia* and *Jupunba*) likely results from hybridisation between co-occurring, interfertile species (i.e. syngameons). Indeed, recent work on Amazonian trees has documented putative syngameons

in other genera, such as between several *Brownea* species (Schley et al. 2020) and amongst three species of *Eschweilera* around Manaus in the Brazilian Amazon (Larson et al. 2021).

The multiple introgression events that we inferred could have influenced 'sub-radiations' within *Inga*, *Zygia* and *Jupunba*, which would also explain why we found introgression across whole clades (Fig. 2a; Fig. 2b; Supplementary Fig. S8a; Fig. S8b available on Dryad). Shared variation across entire clades would be expected following introgression events within ancestral syngameons, as introgressant variants are inherited by descendent species (e.g. Meyer et al. 2017; Meier et al. 2017; Schley et al. 2020). Such an effect may also result from methodological artefacts which we could not fully account for in the D-statistic, F_4 ratio and F_{branch} analyses, such as an incorrectly placed lineage within the trees or changes in substitution rates (Patterson et al. 2012). However, we were highly stringent with our accession filtering and selection, and substitution rates only change materially at the base of the 'fast' clade (Ringelberg et al. 2023). In addition, we inferred introgression into the base of several clades in the TREEMIX analyses, again suggesting introgressed loci were inherited by descendent species.

Incomplete lineage sorting, i.e. the retention of ancestral polymorphisms in descendent lineages (Doyle 1992), is also pervasive in rapid Amazonian tree radiations. This was shown both by our results (Fig. 1a-b; Fig. 2a-2b; Supplementary Fig. 6a; Fig. S6b available on Dryad) and previously demonstrated extensively in the Mimosoid legumes, to which the Ingoid clade belongs (Koenen et al. 2020). ILS arises in these groups because the probability of coalescence (i.e. sorting of derived alleles into descendent lineages reflecting speciation history) in *t* generations decreases with increasing effective population size (Fisher 1930; Wright 1931; Kingman 2000). Most rainforest trees have large, widespread populations (ter Steege et al. 2013), so coalescence and sorting of alleles is unlikely to have yet occurred

across all loci in rapidly-diversifying rainforest tree genera such as *Inga* (discussed in Pennington, R. T. and Lavin 2016).

Introgression and Selection Influence the Evolution of Defence Chemistry Loci in Inga

We detected multiple introgression events across *Inga*, suggesting that some loci transferred by introgression are retained over time. It is likely that these loci were not immediately deleterious and were not subject to purifying selection, perhaps residing in areas of the genome that are distant from incompatibility loci, allowing them to recombine freely (Edelman et al. 2019). It is also possible that these regions are adaptive and so are maintained by positive selection, as shown in temperate tree species (e.g. Rendón-Anaya et al. 2021). This is particularly interesting in the context of chemical defences against insect herbivores, since these defences are known to be important for survival, co-existence and ecological divergence in *Inga* (Kursar et al. 2009; Coley et al. 2018; Forrister et al. 2023).

We found differences in the proportion of loci under selection between locus annotation classes, with more defence chemistry loci under selection than would be expected by chance. This was evident across all subsampled species ($\chi^2 = 9.575$, df = 3, p-value = 0.022, Fig. 4a; Supplementary Table S6; Fig. S12, available on Dryad). Previous work using phylogenetic comparative methods demonstrated divergent evolution in defence chemical profiles among sister species of *Inga* (Forrister et al. 2023). Our results suggest a potential mechanism underlying this divergent evolution, given the molecular evidence of positive selection in defence chemistry loci we observed. This provides an important exemplar for more broadly understanding the assembly of diverse rainforest tree communities, since herbivore pressure structures tree communities and divergent defence chemistry facilitates

coexistence at the community scale for rainforest trees like *Inga* (Kursar et al. 2009; Forrister et al. 2019).

Our estimates of per-locus introgression (f_{dM}) varied widely across the 1167 target capture loci we analysed and across data subsets (Table S5; Supplementary Fig. S11, available on Dryad). Locus annotation best explained variance in introgression across loci in the 'leiocalcycina + vulpina + red hair' data subset (P=<0.01, d.f. = 3, F = 6.76; Fig. 4b; Fig. 4c; Supplementary Table S7, available on Dryad). Moreover, loci inferred to be under selection had marginally significant differences in the proportion of introgressed sites (P= 0.05, df = 3, F = 2.60; Supplementary Table S7, available on Dryad).

Altogether, these analyses suggest a plausible role of introgression in generating adaptive defence chemistry phenotypes in *Inga*, given that f_{dM} scores were marginally higher in defence chemistry loci that were under selection (Fig. 4b). Moreover, novel defence chemicals in *Inga* are hypothesised to arise from combinations of different chemical precursors (Coley et al. 2018), suggesting a potential role for admixture. Thus, assuming introgression generates a novel defence chemical combination that is effective against herbivores, this phenotype may rise in frequency in a population and facilitate adaptation where the abundance of individuals with similar defences is low. In other words, introgression of rare defence loci that confer resistance to different herbivore communities may facilitate colonisation of new areas with different suites of herbivores. This is similar to how introgression of wing pattern mimicry genes facilitates adaptation to local mimicry rings in *Heliconius* butterflies (The Heliconius Genome Consortium 2012).

In light of the Janzen-Connell hypothesis, where higher densities of conspecifics with the same defences leads to increased mortality from herbivores (Janzen 1970; Connell 1971), possession of a rare, introgressant defence chemistry phenotype is likely to be adaptive, as

fewer herbivores in the new area can overcome it. Adaptive introgression facilitating colonisation of new habitats is well known in plants (Suarez-Gonzalez et al. 2016), particularly in the context of defence against herbivores (Whitney et al. 2006), and may have influenced the rapid radiation of *Inga*.

CONCLUSIONS

Our analyses indicate that rapid Amazonian tree radiations (e.g. *Inga*, *Zygia* and *Jupunba*) display evidence of widespread introgression, in addition to incomplete lineage sorting. The introgression we inferred may be evidence of 'syngameons' of co-occurring interfertile species, which are created by dispersal-assembled local tree communities in neotropical rainforests (Dexter et al. 2017). This introgression has likely influenced concurrent 'sub-radiations', rather than an 'early burst' resulting from a single hybrid swarm underlying the entire *Inga* radiation. Functional loci relating to defence chemistry additionally show evidence of selection, and those loci under selection have slightly higher proportions of introgression, indicating that introgression likely facilitates adaptation, local coexistence and diversification in Amazonian trees.

SUPPLEMENTARY MATERIAL

Supplementary material is available from the Dryad Digital Repository: https://doi.org/xxxxxx

FUNDING

This work was supported by a Natural Environment Research Council standard grant (grant number NE/V012258/1). Sequencing was funded partly by the BBSRC, grant number BB/P022898/1. J.J.R. would like to thank the Swiss National Science Foundation Postdoc. mobility grant number P500PB_211111 for support. J.N.'s work was funded by NSF Standard and Dimensions of Biodiversity grants, numbers DEB-0640630 and DEB-1135733. This project utilised equipment funded by the Wellcome Trust (Multi-User Equipment Grant award number 218247/Z/19/Z)

ACKNOWLEDGEMENTS

The authors acknowledge the Research/Scientific Computing teams at The James Hutton Institute and NIAB for providing computational resources and technical support for the 'UK's Crop Diversity Bioinformatics HPC' (BBSRC grant BB/S019669/1), use of which has contributed to the results reported within this paper. Thanks also to Colin Hughes and Erik Koenen for their help in generating the sequencing data for the outgroup species.

DATA AVAILABILITY

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.69p8cz92v. The accession numbers for all data collated from previous studies are found in Supplementary Table S1 on Dryad. All nucleotide sequence data produced by this study are available on NCBI GenBank under the accession numbers **XXXX**. In addition, all phylogenetic trees we produced are available on TreeBASE under the accession numbers **XXXX.**

REFERENCES

Abbott R.J. 2017. Plant speciation across environmental gradients and the occurrence and nature of hybrid zones. *J. Syst. Evol*. 55:238-258.

Agrawal A.A. 1999. Induced responses to herbivory in wild radish: effects on several herbivores and plant fitness. *Ecology*. 80:1713-1723.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available online at http://www.bioinformatics.babraham.ac.uk/projects/fastqc. 0.11.9.

Antonelli A., Sanmartín I. 2011. Why are there so many plant species in the Neotropics? *Taxon*. 60: 403-414.

Ashton P.S. 1969. Speciation among tropical forest trees: some deductions in the light of recent evidence. *Biol. J. Linn. Soc.* 1:155-196.

Baker T.R., Pennington R.T., Magallon S., Gloor E., Laurance W.F., Alexiades M., Alvarez E., Araujo A., Arets E.J., Aymard G. 2014. Fast demographic traits promote high diversification rates of Amazonian trees. *Ecol. Lett*. 17:527-536.

Bankevich A., Nurk S., Antipov D., Gurevich A.A., Dvorkin M., Kulikov A.S., Lesin V.M., Nikolenko S.I., Pham S., Prjibelski A.D. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol*. 19: 455-477.

Barrier M., Baldwin B.G., Robichaux R.H., Purugganan M.D. 1999. Interspecific hybrid ancestry of a plant adaptive radiation: allopolyploidy of the Hawaiian silversword alliance (Asteraceae) inferred from floral homeotic gene duplications. *Mol. Biol. Evol*. 16:1105-1113.

Benjamini Y., Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B*: 289-300.

Bermingham E., Dick C. 2001. The *Inga* - newcomer or museum antiquity? *Science*. 293:2214-2216.

Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30:2114-2120.

Borowiec M.L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ*. 4: e1660.

Cannon C.H., Lerdau M.T. 2019. Demography and destiny: The syngameon in hyperdiverse systems. *Proc. Natl. Acad. Sci. U.S.A.* 116: 8105.

Cannon C.H., Lerdau M.T. 2015. Variable mating behaviors and the maintenance of tropical biodiversity. *Front. Genet*. 6:183.

Capella-Gutiérrez S., Silla-Martínez J.M., Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*. 25: 1972-1973.

Coley P.D., Endara M., Kursar T.A. 2018. Consequences of interspecific variation in defenses and herbivore host choice for the ecology and evolution of *Inga*, a speciose rainforest tree. *Oecologia*. 187:361-376.

Connell J.H. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. In: Den Boer PJ, Gradwell GR, editors. Dynamics of Populations. Wageningen, The Netherlands: Centre for Agricultural Publishing and Documentation

Danecek P., Auton A., Abecasis G., Albers C.A., Banks E., DePristo M.A., Handsaker R.E., Lunter G., Marth G.T., Sherry S.T. 2011. The variant call format and VCFtools. *Bioinformatics*. 27:2156-2158.

Danecek P., Bonfield J.K., Liddle J., Marshall J., Ohan V., Pollard M.O., Whitwham A., Keane T., McCarthy S.A., Davies R.M. 2021. Twelve years of SAMtools and BCFtools. *Gigascience*. 10:giab008.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol*. 24:332-340.

Dexter K.G., Pennington T.D., Cunningham C.W. 2010. Using DNA to assess errors in tropical tree identifications: How often are ecologists wrong and when does it matter? *Ecol. Monogr*. 80: 267-286.

Dexter K.G., Chave J. 2016. Evolutionary patterns of range size, abundance and species richness in Amazonian angiosperm trees. *PeerJ*. 4:e2402.

Dexter K.G., Lavin M., Torke B.M., Twyford A.D., Kursar T.A., Coley P.D., Drake C., Hollands R., Pennington R.T. 2017. Dispersal assembly of rain forest tree communities across the Amazon basin. *Proc. Natl. Acad. Sci. U.S.A.* 114: 2645-2650.

Doyle J.J. 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. *Syst. Bot.*: 144-163.

Durand E.Y., Patterson N., Reich D., Slatkin M. 2011. Testing for ancient admixture between closely related populations. *Mol. Biol. Evol*. 28: 2239-2252.

Edelman N.B., Frandsen P.B., Miyagi M., Clavijo B., Davey J., Dikow R.B., García-Accinelli G., Van Belleghem S.M., Patterson N., Neafsey D.E. 2019. Genomic architecture and introgression shape a butterfly radiation. *Science*. 366:594-599.

Endara M., Soule A.J., Forrister D.L., Dexter K.G., Pennington R.T., Nicholls J.A., Loiseau O., Kursar T.A., Coley P.D. 2022. The role of plant secondary metabolites in shaping regional and local plant community assembly. *J. Ecol*. 110:34-45.

Erkens R.H., Chatrou L.W., Maas J.W., van der Niet T., Savolainen V. 2007. A rapid diversification of rainforest trees (*Guatteria*; Annonaceae) following dispersal from Central into South America. *Mol. Phylogenet. Evol*. 44: 399-411.

Ferm J., Korall P., Lewis G.P., Ståhl B. 2019. Phylogeny of the Neotropical legume genera *Zygia* and *Marmaroxylon* and close relatives. *Taxon*. 68:661-672.

Fisher R.A. 1930. The genetical theory of natural selection. Oxford, UK.: The Clarendon Press.

Fitak R.R. 2021. OptM: estimating the optimal number of migration edges on population trees using Treemix. *Biol. Methods Protoc.* 6:bpab017.

Forrister D.L., Endara M., Soule A.J., Younkin G.C., Mills A.G., Lokvam J., Dexter K.G., Pennington R.T., Kidner C.A., Nicholls J.A. 2023. Diversity and divergence: evolution of secondary metabolism in the tropical tree genus *Inga*. *New Phytol*. 237:631-642.

Forrister D.L., Endara M., Younkin G.C., Coley P.D., Kursar T.A. 2019. Herbivores as drivers of negative density dependence in tropical forest saplings. *Science*. 363:1213-1216.

Green R.E., Krause J., Briggs A.W., Maricic T., Stenzel U., Kircher M., Patterson N., Li H., Zhai W., Fritz M.H., Hansen N.F., Durand E.Y., Malaspinas A.S., Jensen J.D., Marques-Bonet T., Alkan C., Prufer K., Meyer M., Burbano H.A., Good J.M., Schultz R., Aximu-Petri A., Butthof A., Hober B., Hoffner 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 Z., Gusic 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., Paabo S. 2010. A draft sequence of the Neandertal genome. *Science*. 328: 710-722.

Hughes C.E., Nyffeler R., Linder H.P.. 2015. Evolutionary plant radiations: where, when, why and how? *New Phytol*. 207:249-253.

Huson D.H., Bryant D. 2005. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol*. 23: 254-267.

Janzen D.H. 1970. Herbivores and the number of tree species in tropical forests. *Am. Nat*. 104:501-528.

Johnson M.G., Gardner E.M., Liu Y., Medina R., Goffinet B., Shaw A.J., Zerega N.J., Wickett N.J.. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *App. Plant. Sci*. 4: 1600016.

Katoh K., Standley D.M.. 2013. MAFFT: Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Mol. Biol. Evol*. 30:772-780.

Kingman J.F. 2000. Origins of the coalescent. 1974-1982. *Genetics*. 156: 1461-1463.

Koenen E.J., Clarkson J.J., Pennington T.D., Chatrou L.W. 2015. Recently evolved diversity and convergent radiations of rainforest mahoganies (Meliaceae) shed new light on the origins of rainforest hyperdiversity. *New Phytol*. 207: 327-339.

Koenen E.J., Kidner C., de Souza É.R., Simon M.F., Iganci J.R., Nicholls J.A., Brown G.K., De Queiroz L.P., Luckow M., Lewis G.P. 2020. Hybrid capture of 964 nuclear genes resolves evolutionary relationships in the mimosoid legumes and reveals the polytomous origins of a large pantropical radiation. *Am. J. Bot*. 107:1710-1735.

Kuhnhäuser B.G., Bellot S., Couvreur T.L., Dransfield J., Henderson A., Schley R., Chomicki G., Eiserhardt W.L., Hiscock S.J., Baker W.J. 2021. A robust phylogenomic framework for the calamoid palms. *Mol. Phylogenet. Evol*. 157:107067.

Kursar T.A., Dexter K.G., Lokvam J., Pennington R.T., Richardson J.E., Weber M.G., Murakami E.T., Drake C., McGregor R., Coley P.D. 2009. The evolution of antiherbivore defenses and their contribution to species coexistence in the tropical tree genus *Inga*. *Proc. Natl. Acad. Sci. U.S.A.* 106: 18073-18078.

Lamichhaney S., Han F., Webster M.T., Andersson L., Grant B.R., Grant P.R. 2018. Rapid hybrid speciation in Darwin's finches. *Science*. 359: 224-228.

Larson D.A., Vargas O.M., Vicentini A., Dick C.W. 2021. Admixture may be extensive among hyperdominant Amazon rainforest tree species. *New Phytol*. 232:2520-2534.

Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 27:2987-2993.

Li H., Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 25: 1754-1760.

Malinsky M., Challis R.J., Tyers A.M., Schiffels S., Terai Y., Ngatunga B.P., Miska E.A., Durbin R., Genner M.J., Turner G.F. 2015. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science*. 350:1493-1498.

Malinsky M., Matschiner M., Svardal H. 2021. Dsuite□fast D□statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour*. 21:584-595.

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.

Marques D.A., Meier J.I., Seehausen O. 2019. A combinatorial view on speciation and adaptive radiation. *Trend. Ecol. Evol*. 34: 531-544.

McVay J.D., Hauser D., Hipp A.L., Manos P.S. 2017. Phylogenomics reveals a complex evolutionary history of lobed-leaf white oaks in western North America. *Genome*. 60:733-742.

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. Comms*. 8: 14363.

Meyer B.S., Matschiner M., Salzburger W. 2017. Disentangling incomplete lineage sorting and introgression to refine species-tree estimates for Lake Tanganyika cichlid fishes. *Syst. Biol*. 66:531-550.

Morales Briones D.F., Liston A., Tank D.C. 2018. Phylogenomic analyses reveal a deep history of hybridization and polyploidy in the Neotropical genus *Lachemilla* (Rosaceae). *New Phytol*. 218: 1668-1684.

Murrell B., Weaver S., Smith M.D., Wertheim J.O., Murrell S., Aylward A., Eren K., Pollner T., Martin D.P., Smith D.M. 2015. Gene-wide identification of episodic selection. *Mol. Biol. Evol*. 32:1365-1371.

Naciri Y., Linder H.P. 2015. Species delimitation and relationships: the dance of the seven veils. *Taxon*. 64: 3-16.

Nguyen L., Schmidt H.A., Von Haeseler A., Minh B.Q. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol*. 32:268-274.

Nicholls J.A., Pennington R.T., Koenen E.J., Hughes C.E., Hearn J., Bunnefeld L., Dexter K.G., Stone G.N., Kidner C.A. 2015. Using targeted enrichment of nuclear genes to increase phylogenetic resolution in the neotropical rain forest genus *Inga* (Leguminosae: Mimosoideae). *Front. Plant Sci*. 6:710.

Patterson N., Moorjani P., Luo Y., Mallick S., Rohland N., Zhan Y., Genschoreck T., Webster T., Reich D. 2012. Ancient admixture in human history. *Genetics*. 192:1065-1093.

Pease J.B., Brown J.W., Walker J.F., Hinchliff C.E., Smith S.A. 2018. Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. *Am. J. Bot*. 105:385-403.

Pennington R.T., Lavin M. 2016. The contrasting nature of woody plant species in different neotropical forest biomes reflects differences in ecological stability. *New Phytol*. 210: 25-37.

Pennington T.D. 1997. The Genus *Inga*: Botany. Royal Botanic Gardens, Kew, London, UK: Kew Publishing

Pickrell J.K., Pritchard J.K. 2012a. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet*. 8: e1002967.

Pickrell J.K., Pritchard J.K. 2012b. User Manual for TreeMix v1. Available at: https://docplayer.net/31797577-User-manual-for-treemix-v1-0-joseph-k-pickrell-jonathan-k-pritchard.html

R Development Core Team. 2013. R: A language and environment for statistical computing. Available at: http://www.R-project.org/. 3.6.

Ranwez V., Chantret N., Delsuc F. 2021. Aligning Protein-Coding nucleotide sequences with MACSE. In: Katoh K, editor. Multiple Sequence Alignment: Methods and Protocols. New York, NY, USA: Springer p. 51-70.

Ranwez V., Harispe S., Delsuc F., Douzery E.J. 2011. MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. *PloS one*. 6:e22594.

Raven P.H., Gereau R.E., Phillipson P.B., Chatelain C., Jenkins C.N., Ulloa Ulloa C. 2020. The distribution of biodiversity richness in the tropics. *Sci. Adv*. 6:eabc6228.

Rendón-Anaya M., Wilson J., Sveinsson S., Fedorkov A., Cottrell J., Bailey M.E., Ruņ is D., Lexer C., Jansson S., Robinson K.M. 2021. Adaptive introgression facilitates adaptation to high latitudes in European aspen (*Populus tremula* L.). *Mol. Biol. Evol*. 38:5034-5050.

Richardson J.E., Pennington R.T., Pennington T.D., Hollingsworth P.M. 2001. Rapid diversification of a species-rich genus of neotropical rain forest trees. *Science*. 293: 2242-2245.

Rieseberg L.H., Archer M.A., Wayne R.K. 1999. Transgressive segregation, adaptation and speciation. *Heredity*. 83:363-372.

Rieseberg L.H., Willis J.H. 2007. Plant speciation. *Science*. 317:910-914.

Ringelberg J.J., Koenen E.J., Iganci J.R., de Queiroz L.P., Murphy D.J., Gaudeul M., Bruneau A., Luckow M., Lewis G.P., Hughes C.E. 2022. Phylogenomic analysis of 997 nuclear genes reveals the need for extensive generic re-delimitation in Caesalpinioideae (Leguminosae). *PhytoKeys*. 205:3-58.

Ringelberg J.J., Koenen E.J., Sauter B., Aebli A., Rando J.G., Iganci J.R., de Queiroz L.P., Murphy D.J., Gaudeul M., Bruneau A. 2023. Precipitation is the main axis of tropical plant phylogenetic turnover across space and time. *Sci. Adv*. 9:eade4954.

Rivers M., Beech E., Bazos I., Bogunić F., Buira A., Caković D., Carapeto A., Carta A., Cornier B., Fenu G. 2019. European Red List of Trees. International Union for Conservation of Nature and Natural Resources (IUCN).

Sayyari E., Whitfield J.B., Mirarab S. 2018. DiscoVista: interpretable visualizations of gene tree discordance. *Mol. Phylogenet. Evol*. 122:110-115.

Schley R.J., Pennington R.T., Pérez□Escobar O.A., Helmstetter A.J., de la Estrella M., Larridon I., Sabino Kikuchi I. A. B., Barraclough T.G., Forest F., Klitgård B. 2020. Introgression across evolutionary scales suggests reticulation contributes to Amazonian tree diversity. *Mol. Ecol*. 29:4170-4185.

Schley R.J., Twyford A.D., Pennington R.T. 2022. Hybridization: a ‘double-edged sword’ for Neotropical plant diversity. *Bot. J. Linn. Soc*. 199:331-356.

Schumer M., Cui R., Rosenthal G.G., Andolfatto P. 2015. Reproductive isolation of hybrid populations driven by genetic incompatibilities. *PLoS Genet.* 11:e1005041.

Seehausen O. 2004. Hybridization and adaptive radiation. *Trend. Ecol. Evol.* 19:198-207.

Slater G.S.C., Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinform*. 6: 31.

Suarez□Gonzalez A., Hefer C.A., Christe C., Corea O., Lexer C., Cronk Q.C., Douglas C.J. 2016. Genomic and functional approaches reveal a case of adaptive introgression from *Populus balsamifera* (balsam poplar) in *P. atrichocarpa* (black cottonwood). *Mol. Ecol*. 25: 2427-2442.

Suarez-Gonzalez A., Lexer C., Cronk Q.C.B. 2018. Adaptive introgression: a plant perspective. *Biol. Lett*. 14: 10.1098/rsbl.2017.0688.

ter Steege H., Pitman N.C., Sabatier D., Baraloto C., Salomao R.P., Guevara J.E., Phillips O.L., Castilho C.V., Magnusson W.E., Molino J.F., Monteagudo A., Nunez Vargas P., Montero J.C., Feldpausch T.R., Coronado E.N., Killeen T.J., Mostacedo B., Vasquez R., Assis R.L., Terborgh J., Wittmann F., Andrade A., Laurance W.F., Laurance S.G., Marimon B.S., Marimon B.H.Jr., Guimaraes Vieira I.C., Amaral I.L., Brienen R., Castellanos H., Cardenas Lopez D., Duivenvoorden J.F., Mogollon H.F., Matos F.D., Davila N., Garcia-Villacorta R., Stevenson Diaz P.R., Costa F., Emilio T., Levis C., Schietti J., Souza P., Alonso A., Dallmeier F., Montoya A.J., Fernandez Piedade M.T., Araujo-Murakami A., Arroyo L., Gribel R., Fine P.V., Peres C.A., Toledo M., Aymard C.G.A., Baker T.R., Ceron

C., Engel J., Henkel T.W., Maas P., Petronelli P., Stropp J., Zartman C.E., Daly D., Neill D., Silveira M., Paredes M.R., Chave J., Lima Filho D. de A., Jorgensen P.M., Fuentes A., Schongart J., Cornejo Valverde F., Di Fiore A., Jimenez E.M., Penuela Mora M.C., Phillips J.F., Rivas G., van Andel T.R., von Hildebrand P., Hoffman B., Zent E.L., Malhi Y., Prieto A., Rudas A., Ruschell A.R., Silva N., Vos V., Zent S., Oliveira A.A., Schutz A.C., Gonzales T., Trindade Nascimento M., Ramirez-Angulo H., Sierra R., Tirado M., Umana Medina M.N., van der Heijden G., Vela C.I., Vilanova Torre E., Vriesendorp C., Wang O., Young K.R., Baider C., Balslev H., Ferreira C., Mesones I., Torres-Lezama A., Urrego Giraldo L.E., Zagt R., Alexiades M.N., Hernandez L., Huamantupa-Chuquimaco I., Milliken W., Palacios Cuenca W., Pauletto D., Valderrama Sandoval E., Valenzuela Gamarra L., Dexter K.G., Feeley K., Lopez-Gonzalez G., Silman M.R. 2013. Hyperdominance in the Amazonian tree flora. *Science*. 342: 1243092.

The Heliconius Genome Consortium. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*. 487:94-98.

Ulloa Ulloa C., Acevedo-Rodriguez P., Beck S., Belgrano M.J., Bernal R., Berry P.E., Brako L., Celis M., Davidse G., Forzza R.C., Gradstein S.R., Hokche O., Leon B., Leon-Yanez S., Magill R.E., Neill D.A., Nee M., Raven P.H., Stimmel H., Strong M.T., Villasenor J.L., Zarucchi J.L., Zuloaga F.O., Jorgensen P.M. 2017. An integrated assessment of the vascular plant species of the Americas. *Science*. 358: 1614-1617.

Valencia R., Balslev H., Miño G.P.Y. 1994. High tree alpha-diversity in Amazonian Ecuador. *Biodivers. Conserv*. 3: 21-28.

Valencia R., Foster R.B., Villa G., Condit R., Svenning J., Hernandez C., Romoleroux K., Losos E., Magård E., Balslev H. 2004. Tree species distributions and local habitat variation in the Amazon: large forest plot in eastern Ecuador. *J. Ecol*. 92: 214-229.

Wagner N.D., He L., Hörandl E. 2020. Phylogenomic relationships and evolution of polyploid *Salix* species revealed by RAD sequencing data. *Front. Plant Sci*. 11:1077.

WCVP. World Checklist of Vascular Plants, version 2.0. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; http://wcvp.science.kew.org/ Retrieved 31st January 2020. Available from http://wcvp.science.kew.org/.

Wei T., Simko V., Levy M., Xie Y., Jin Y., Zemla J. 2017. Package 'corrplot'. *Statistician*. 56:e24.

Wendel J.F., Doyle J.J. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. In: Soltis DE, Soltis PS, Doyle JJ, editors. Molecular Systematics of plants II: DNA sequencing. New York, NY, USA: Springer; p. 265-296.

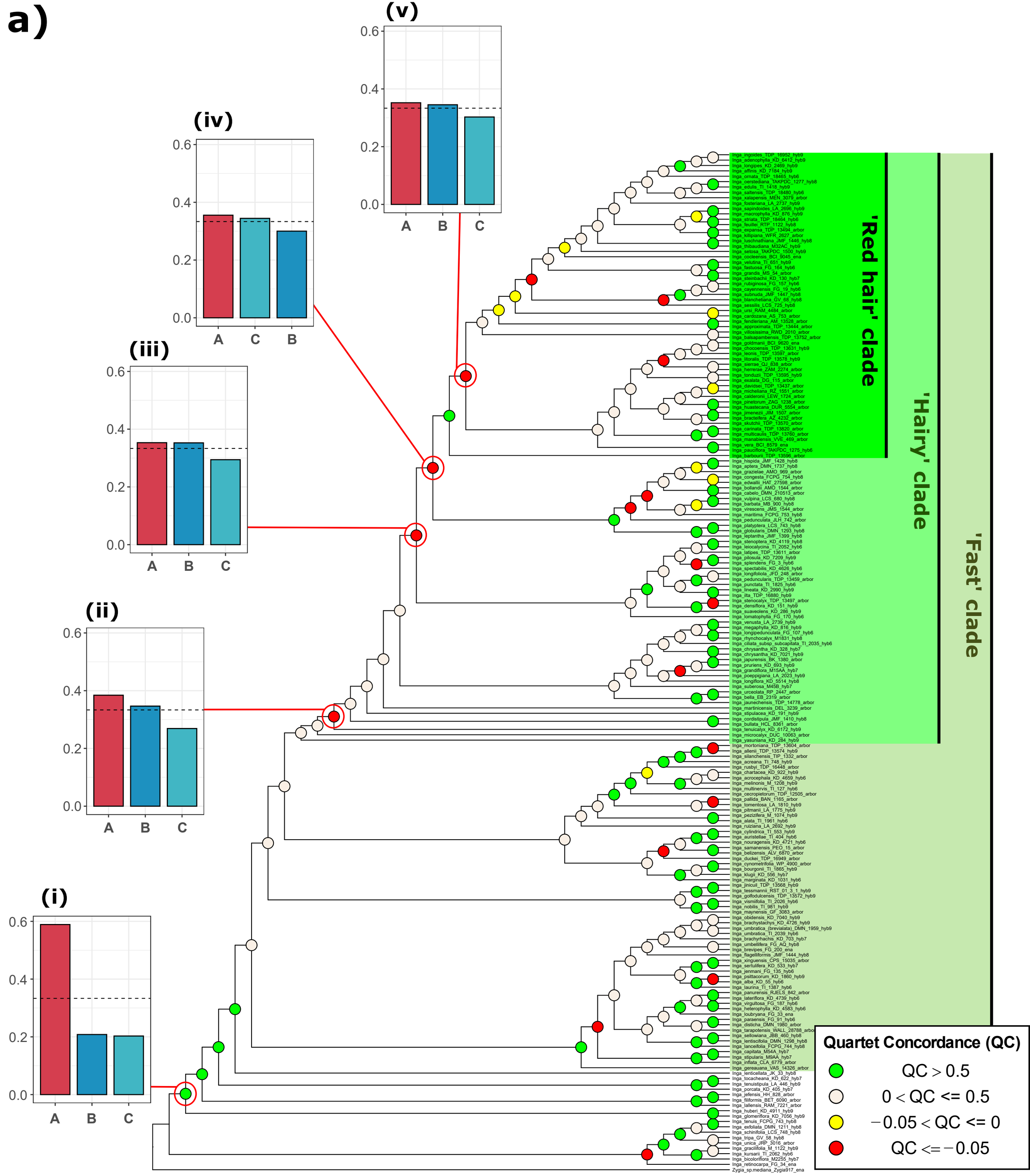
Whitney K.D., Randell R.A., Rieseberg L.H. 2006. Adaptive introgression of herbivore resistance traits in the weedy sunflower *Helianthus annuus*. *Am. Nat*. 167:794-807.

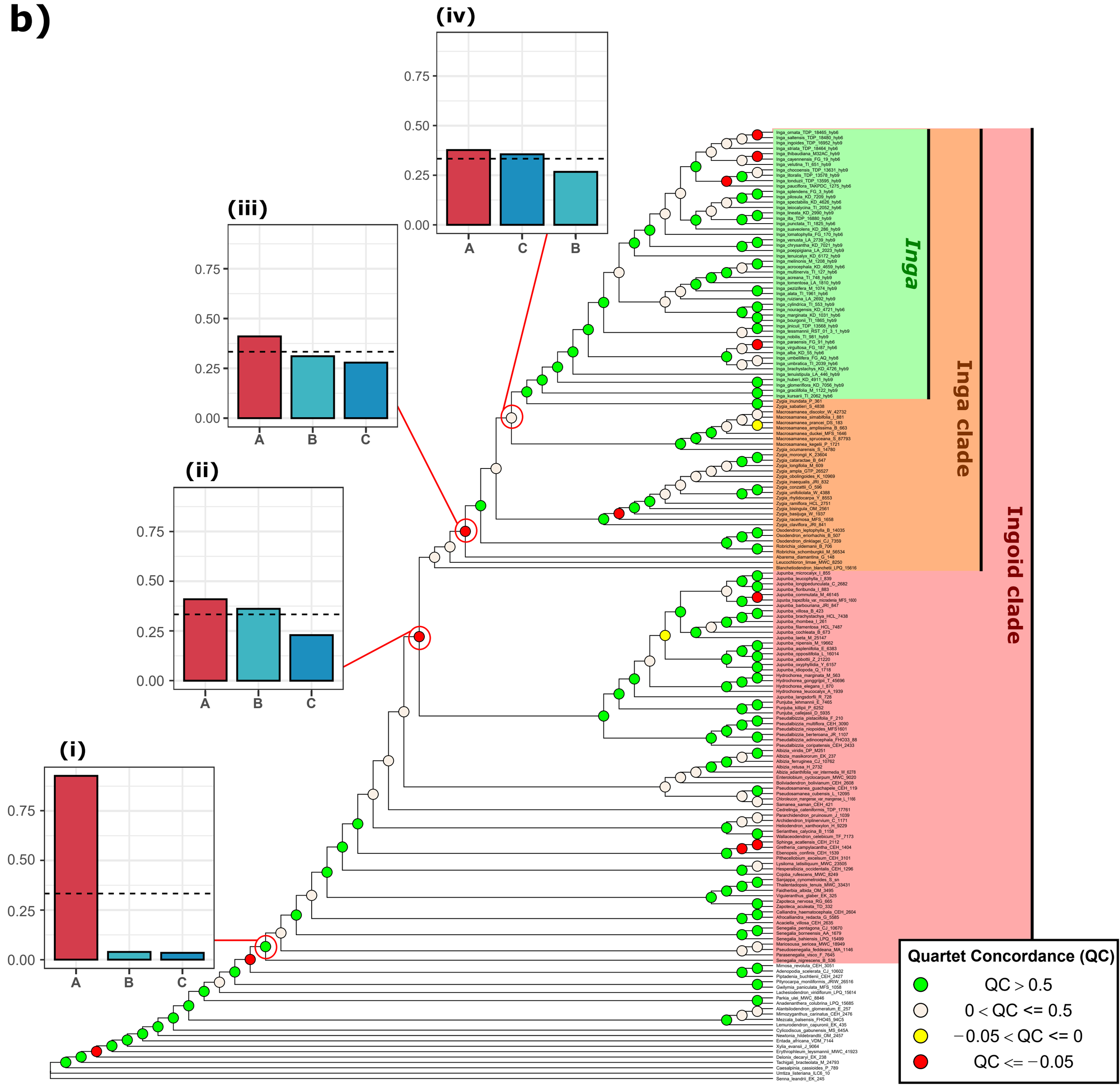
Wickham H. 2016. ggplot2: elegant graphics for data analysis. 3.3.6. New York (NY), USA: Springer-Verlag

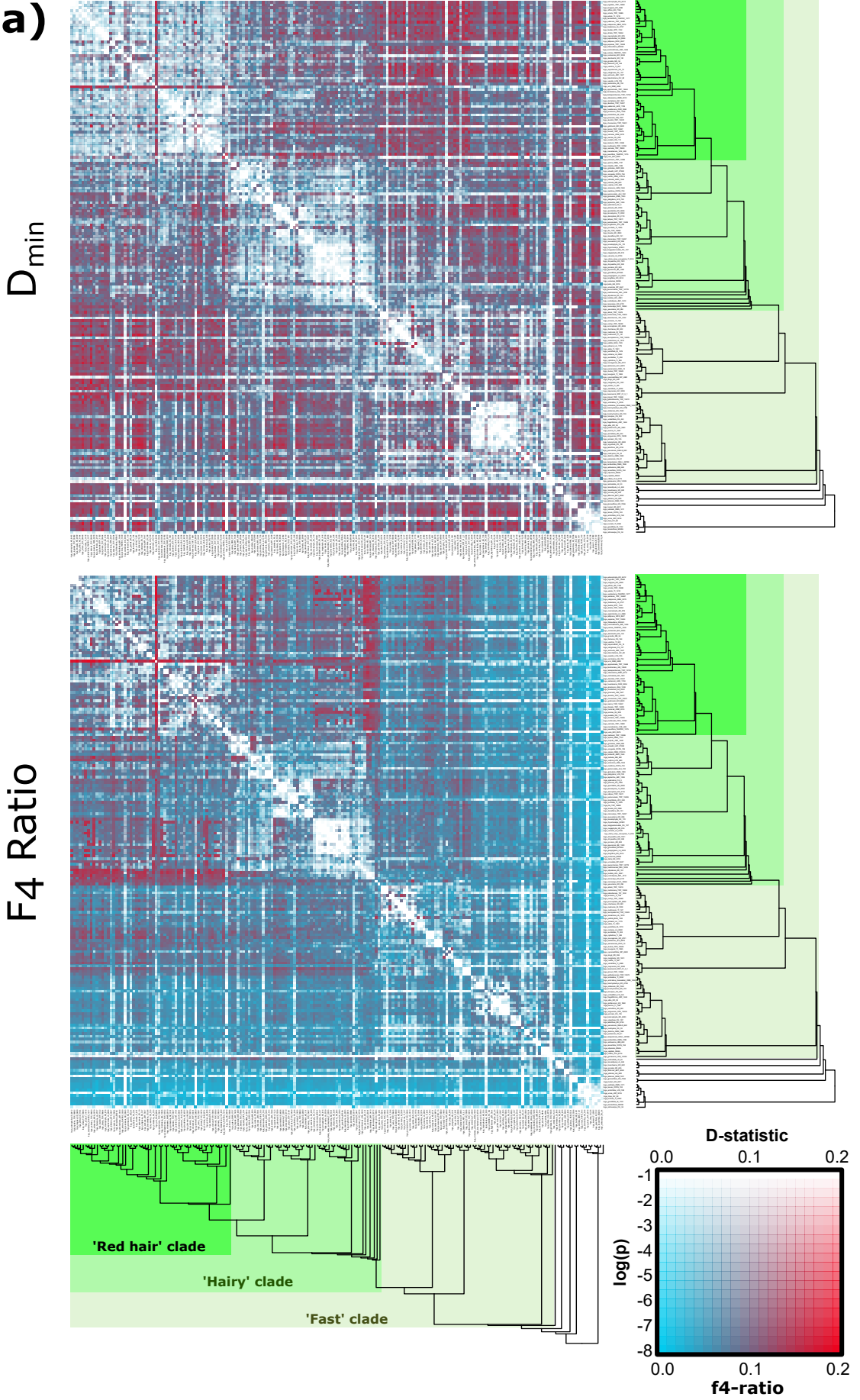
Wright S. 1931. Evolution in Mendelian Populations. *Genetics*. 16: 97-159.

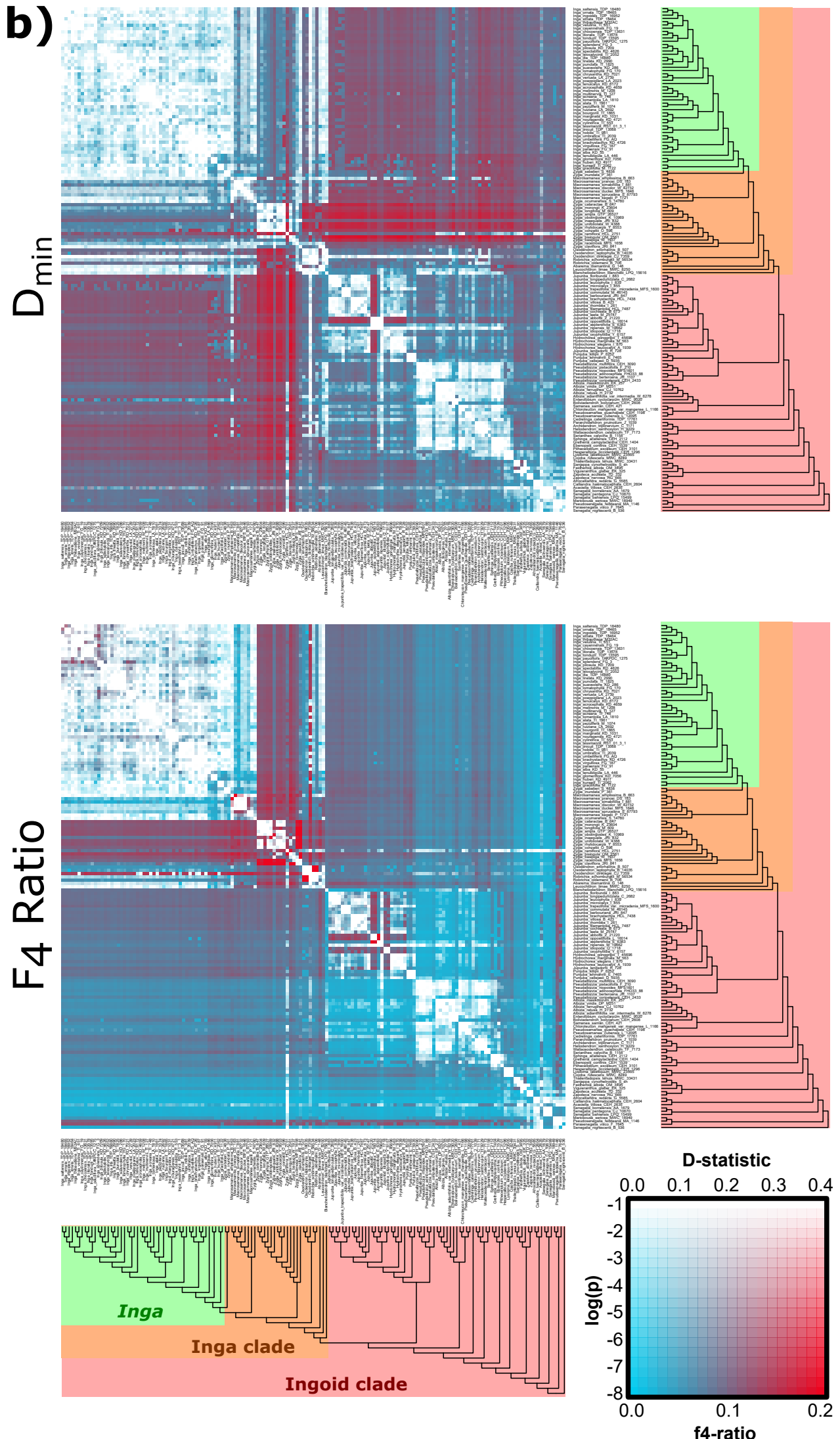
Zhang C., Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinform*. 19:153.

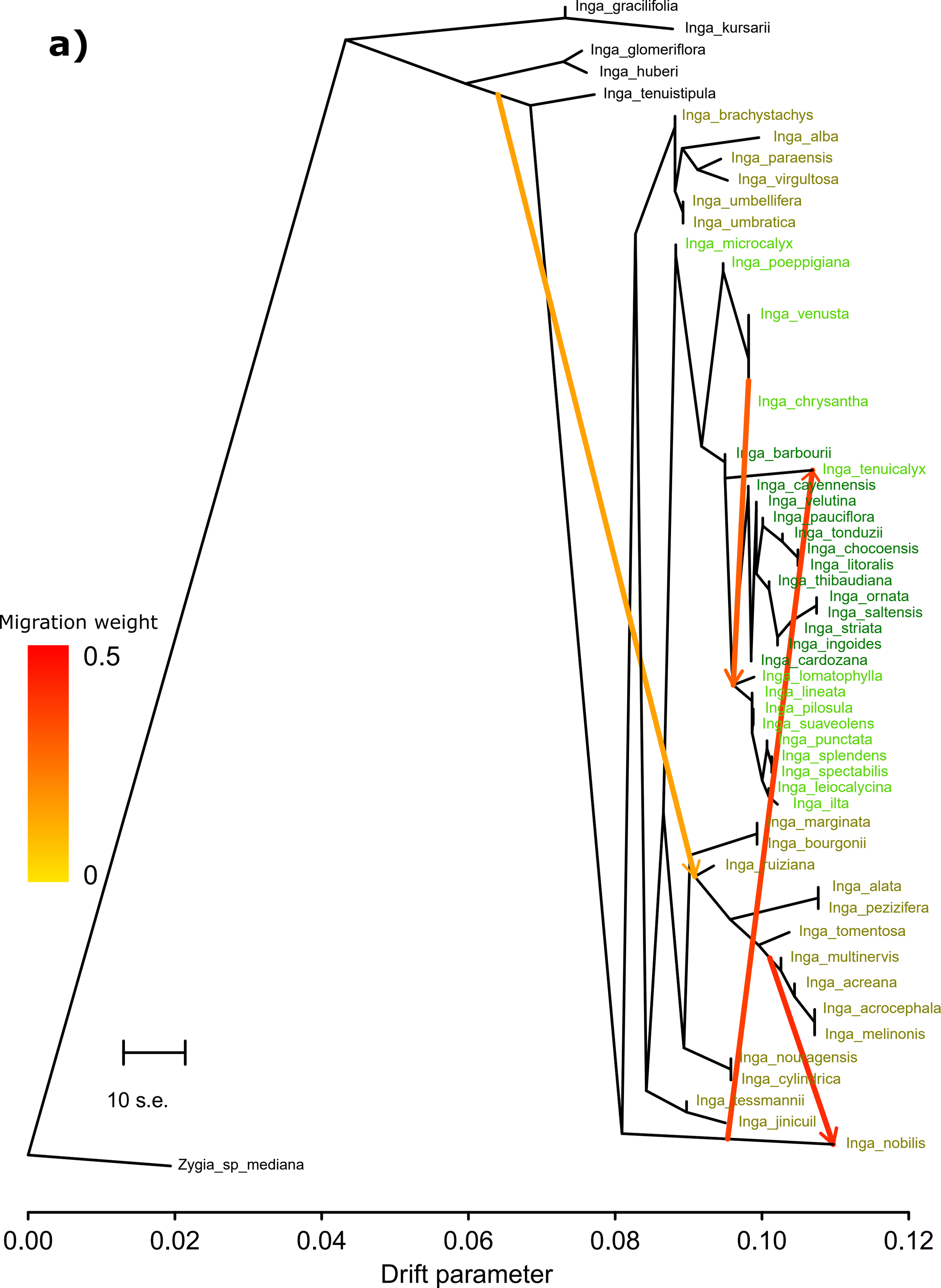
Zhou W., Soghigian J, Xiang Q. 2022. A new pipeline for removing paralogs in target enrichment data. *Syst. Biol*. 71:410-425.

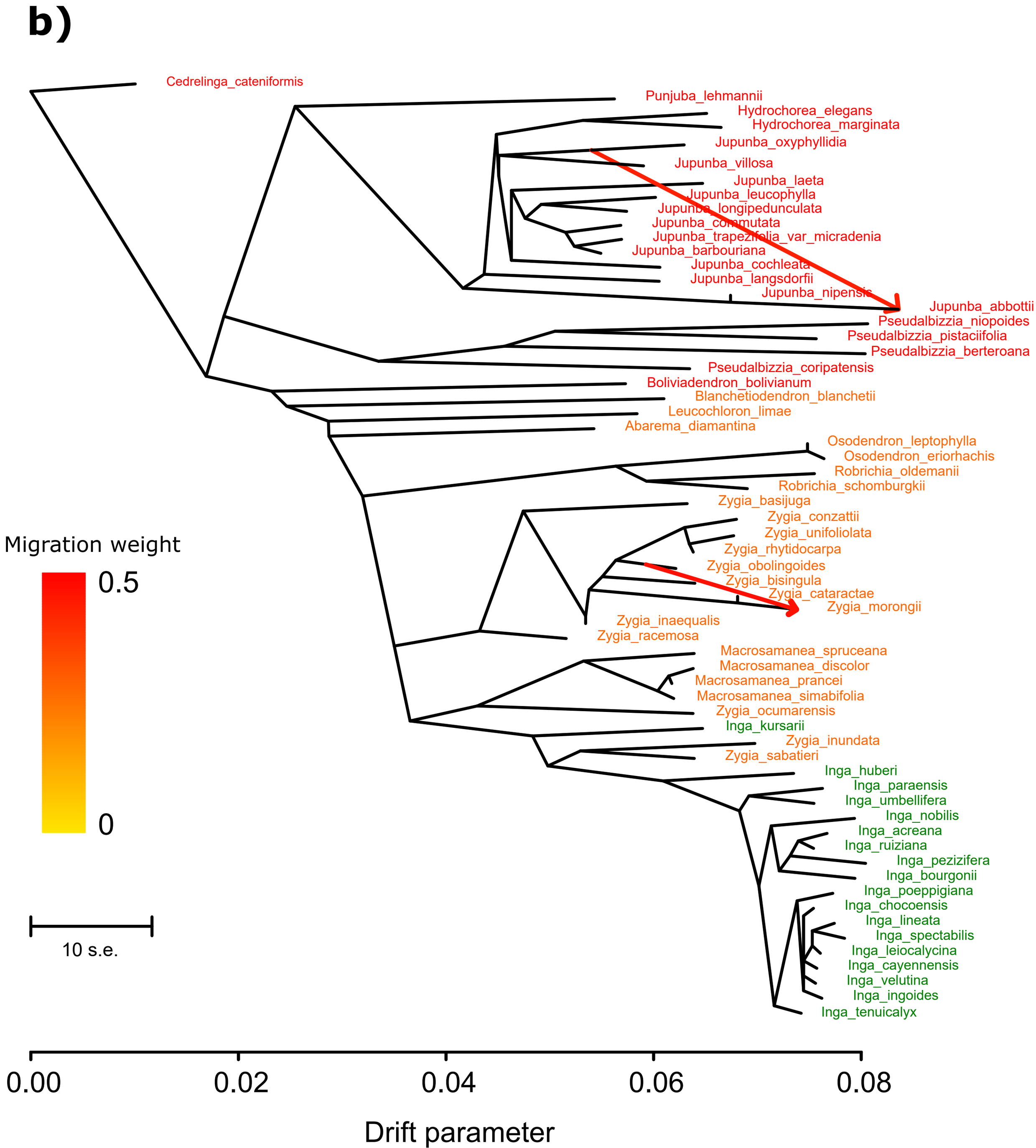

a)
(i)
(ii)
(iii)
(iv)
(v)
0.6
0.4
0.2
0.0
A
B
C
'Red hair' clade
'Hairy' clade
'Fast' clade
Quartet Concordance (QC)
QC > 0.5
0 < QC <= 0.5
−0.05 < QC <= 0
QC <= −0.05

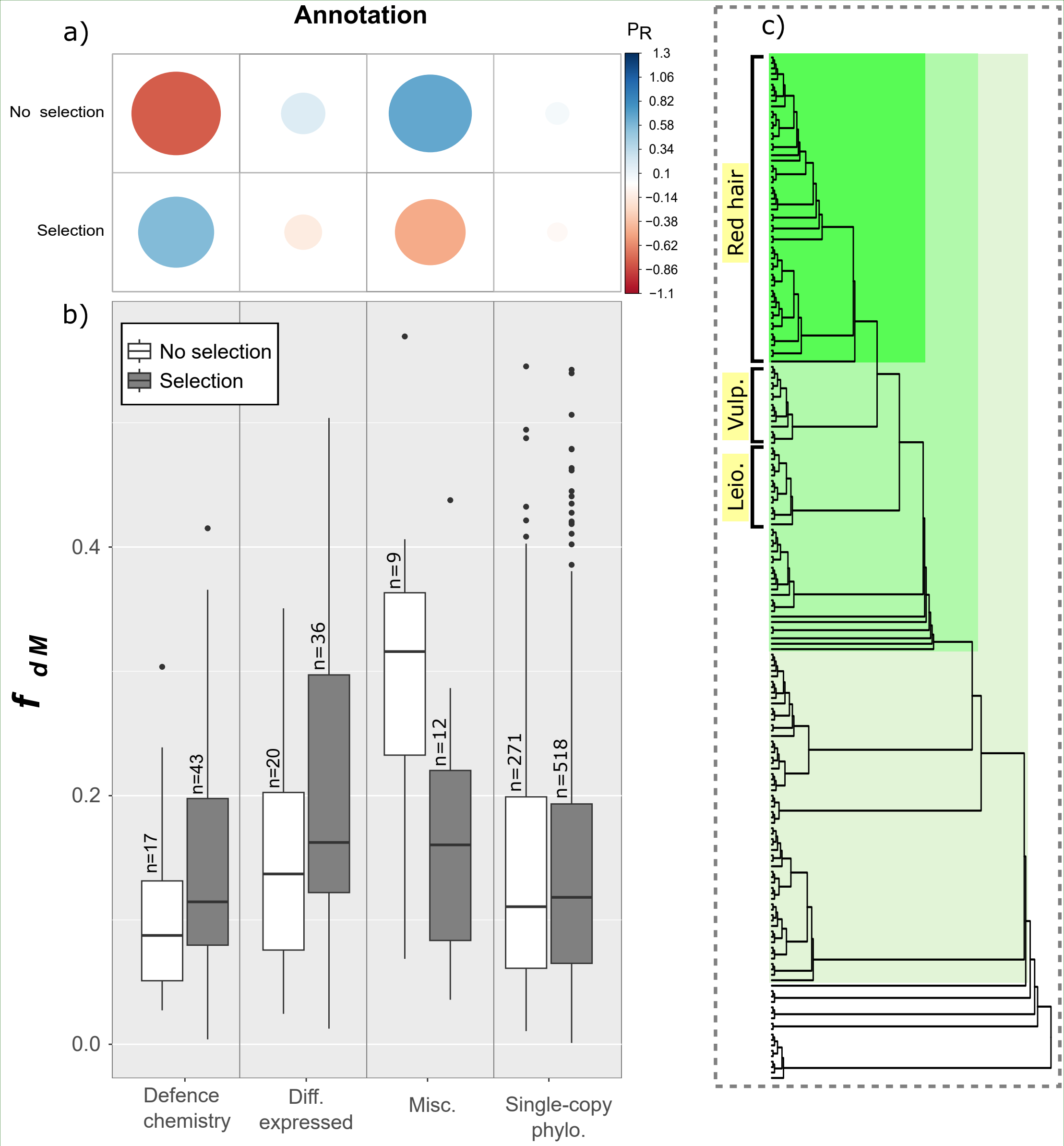

b)
(iv)
(iii)
(ii)
(i)
0.75
0.50
0.25
0.00
A
B
C
Inga
Inga clade
Ingoid clade
Quartet Concordance (QC)
QC > 0.5
0 < QC <= 0.5
−0.05 < QC <= 0
QC <= −0.05


a)
D_{min}
F4 Ratio
'Red hair' clade
'Hairy' clade
'Fast' clade
D-statistic
0.0
0.1
0.2
-1
-2
-3
-4
-5
-6
-7
-8
log(p)
0.0
0.1
0.2
f4-ratio


b)
D_{min}
F4 Ratio
Inga
Inga clade
Ingoid clade
D-statistic
0.0
0.1
0.2
0.3
0.4
log(p)
-1
-2
-3
-4
-5
-6
-7
-8
f4-ratio
0.0
0.1
0.2


a)
Inga_gracilifolia
Inga_kursarii
Inga_glomeriflora
Inga_huberi
Inga_tenuistipula
Inga_brachystachys
Inga_alba
Inga_paraensis
Inga_virgultosa
Inga_umbellifera
Inga_umbratica
Inga_microcalyx
Inga_poeppigiana
Inga_venusta
Inga_chrysantha
Inga_barbourii
Inga_tenuicalyx
Inga_cayennensis
Inga_velutina
Inga_pauciflora
Inga_tonduzii
Inga_chocoensis
Inga_litoralis
Inga_thibaudiana
Inga_ornata
Inga_saltensis
Inga_striata
Inga_ingoides
Inga_cardozana
Inga_lomatophylla
Inga_lineata
Inga_pilosula
Inga_suaveolens
Inga_punctata
Inga_splendens
Inga_spectabilis
Inga_leiocalycina
Inga_ilta
Inga_marginata
Inga_bourgonii
Inga_ruiziana
Inga_alata
Inga_pezizifera
Inga_tomentosa
Inga_multinervis
Inga_acreana
Inga_acrocephala
Inga_melinonis
Inga_nouragensis
Inga_cylindrica
Inga_tessmannii
Inga_jinicuil
Inga_nobilis
Zygia_sp_mediana
Migration weight
0.5
0
10 s.e.
0.00
0.02
0.04
0.06
0.08
0.10
0.12
Drift parameter


b)
Cedrelinga_cateniformis
Punjuba_lehmannii
Hydrochorea_elegans
Hydrochorea_marginata
Jupunba_oxyphyllidia
Jupunba_villosa
Jupunba_laeta
Jupunba_leucophylla
Jupunba_longipedunculata
Jupunba_commutata
Jupunba_trapezifolia_var_micradenia
Jupunba_barbouriana
Jupunba_cochleata
Jupunba_langsdorfii
Jupunba_nipensis
Jupunba_abbottii
Pseudalbizzia_niopoides
Pseudalbizzia_pistaciifolia
Pseudalbizzia_berteroana
Pseudalbizzia_coripatensis
Boliviadendron_bolivianum
Blanchetiodendron_blanchetii
Leucochloron_limae
Abarema_diamantina
Osodendron_leptophylla
Osodendron_eriorhachis
Robrichia_oldemanii
Robrichia_schomburgkii
Zygia_basijuga
Zygia_conzattii
Zygia_unifoliolata
Zygia_rhytidocarpa
Zygia_obolingoides
Zygia_bisingula
Zygia_cataractae
Zygia_morongii
Zygia_inaequalis
Zygia_racemosa
Macrosamanea_spruceana
Macrosamanea_discolor
Macrosamanea_prancei
Macrosamanea_simabifolia
Zygia_ocumarensis
Inga_kursarii
Zygia_inundata
Zygia_sabatieri
Inga_huberi
Inga_paraensis
Inga_umbellifera
Inga_nobilis
Inga_acreana
Inga_ruiziana
Inga_pezizifera
Inga_bourgonii
Inga_poeppigiana
Inga_chocoensis
Inga_lineata
Inga_spectabilis
Inga_leiocalycina
Inga_cayennensis
Inga_velutina
Inga_ingoides
Inga_tenuicalyx
Migration weight
0.5
0
10 s.e.
0.00
0.02
0.04
0.06
0.08
Drift parameter


Annotation
a)
No selection
Selection
P_R
1.3
1.06
0.82
0.58
0.34
0.1
−0.14
−0.38
−0.62
−0.86
−1.1
b)
No selection
Selection
f_{dM}
0.4
0.2
0.0
n=17
n=43
n=20
n=36
n=9
n=12
n=271
n=518
Defence chemistry
Diff. expressed
Misc.
Single-copy phylo.
c)
Red hair
Vulp.
Leio.