



universität
wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

Phylogenetic relationships and the evolutionary history of
Tillandsia, a species-rich genus of Bromeliaceae

verfasst von / submitted by

Gil Yardeni

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Doctor of Philosophy (PhD)

Wien, 2022 / Vienna 2022

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on the student
record sheet:

UA 794 685 437

Dissertationsgebiet lt. Studienblatt /
field of study as it appears on the student record sheet:

Biologie

Betreut von / Supervisor:

Assoz. Prof. Dipl.-Ing. Dr. Ovidiu Paun

Acknowledgments

All scientific work is performed by a community of people, carried through with an investment of time and resources. I wish to thank many people who were involved and made this project possible. I cannot begin to express my gratitude towards my supervisors, Prof. Ovidiu Paun and Dr. Thibault Leroy, for adopting this project and treating it as their own. Their continuous support, attendance and generous sharing of their experience and knowledge allowed me to thrive scientifically. I am indebted to my supervisor Prof. Christian Lexer who initiated the project and whose scientific enthusiasm and excellence were rivaled by his kindness. He provided me with tremendous scientific inspiration and support, advocating for appreciation of the humanity in every person, cooperation over competition and persistent positive approach. It's my privilege to join the line of his past students, through which his mentorship of gentle academics lives on.

I had the privilege to be mentored by excellent researchers working within and with the group. Dr. Jaqueline Hess shared her tremendous knowledge in genomics and bioinformatics, always in an encouraging, patient and professional manner. She helped raise me from a beginner to a confident bioinformatician and considering her among my friends is a pleasure. Dr. Michael H. J. Barfuss shared his expertise in molecular work and is responsible for every successful DNA extraction and for much of my knowledge about bromeliads. I'm thankful for his love for the research system and honest reflections. Dr. Juan Viruel was a wonderful colleague, generously supported the first chapter of this work and was incredibly responsive and present. We had fun sharing cultural understanding for the importance of vinegar, sunlight and bleach. Dr. Margot Paris inspired large parts of this project and contributed experience, ideas and discussions. Her excellent analytical skills assisted in assuring this work was performed on the best side. I'm grateful for the terrific discussions with Prof. Walter Till on anything related to *Tillandsia* and for his expertise at determination. His wealth of knowledge helped connect this work to past and future scientific endeavors.

My gratitude extends towards the scholars and employees in the university of Vienna, besides whom I worked. Dr. Thelma Barbara was extremely kind and supportive, her knowledge and insights were beacon of light in moments of difficulty. Dr. Marcela Van Loo was wonderfully supportive and caring, her humor always lifted the office atmosphere. While

I had only brief opportunities to work with Dr. David Field and Dr. Camille Christer, both contributed so much knowledge in a short amount of time. I'm thankful to Elfriede Grasserbauer and to Juliane Baar for their lab assistance and advice, to Monika Paschinger for the kind administrative assistance and to Jan Zabloudil for responding quickly and efficiently to the many requests while working on the Vienna Scientific Cluster.

I would like to thank Dr. Valeria Guzmán-Jacob for all her help in field collections, identification and teaching me single-rope climbing. It was delightful to work with someone so positive with contagious enthusiasm toward tropical plants. Dr. Thorsten Krömer hosted me so generously in Veracruz, thanks to him my visit was productive, educating and lots of fun. I'm thankful for our collaborators who provided samples: Libertad Silva-Mijangos, Juan P. Pinzón and the SNF Sinergia collaborators.

Prof. Alex Buerkle, Prof. Andrea Betancourt and Dr. Kamil S. Jaron showed unusual kindness in the winter of 2019 - 2020. The words and advice they offered stayed with me throughout the project and beyond. I would also like to acknowledge the scientists who provided advice and productive discussions, especially Dr. Magnus Nordborg, Dr. Kelly Swarts and Prof. Nick Barton. It's a true privilege to learn from them. Dr. Yuval Sapir's ceaseless mentorship, support and friendship have been extremely meaningful throughout this project and in general.

I would like to thank my peers and fellow doctoral students from the bottom of my heart, not only for being fantastic colleagues but mostly for being my friends. Marylaure de la Harpe's experience and work provided the shoulders on which this project stands and I'm grateful for her great help. Huiying Shang was my address for all scientific questions and she passed down knowledge, scripts and advice; Norma Rivera performed excellent lab work and played great Latin music during our fun days in the lab; Aglaia Szukula and I were peers for many good years, her contagious laugh and positive personality were an inspiration to stay true to oneself. Clara Groot Crego was a splendid colleague – thank you especially for many hours of office chatter and much-needed distraction. Niel McNair became a best friend for all the right reasons. I would also like to express my deep thanks to Marie Kristine Brandrud, Jacqueline Heckenhauer, Thomas Wolfe, Mimmi Eriksson, Christina Hedderich, Ahmad Muhammad, Nico Criscuolo, Dana Paun and Matt Thornton. They all provided advice, support, help and terrific company.

I would like to thank my biological family for their continuous support and presence which crossed even growing distances. Finally, I would like to thank my chosen family, the network of friends near and far and my beloved Aram. Their unending care, radical acceptance and unwavering faith are the wind in my sails.

this work is dedicated to my grandfather,
Shlomo Pierre Trefler (1926 – 2020)
for teaching me how to sail a ship against the wind
and the names of plants.

and

to my initial supervisor,
Prof. Christian Lexer (1971 – 2019)
for his invaluable mentorship and scientific inspiration.

Table of Contents

Abstract.....	10
Zusammenfassung.....	12
Introduction.....	14
The genetics of rapid evolutionary radiations	14
Phylogenomics of rapid radiations	17
Study group	22
Research objectives.....	27
References	28
Chapter 1	42
Supporting information	63
Chapter 2.....	146
Supporting Information.....	184
Conclusions and outlook.....	242
References	246
Appendix 1.....	248
Appendix 2.....	250

Abstract

Understanding the drivers and constraints of species formation is at the heart of evolutionary research. The processes that give rise to new species often involve colonization of novel ecological niches and the evolution of novel traits, achieved through an intricate interplay of environment, phenotype and genotype. Modern high-throughput sequencing enables to resolve phylogenomic relationships and to investigate the genomic footprints of evolutionary history. However, research on these topics in recent rapid radiations poses unique challenges: a brief diversification period in adjacent habitats increases the probability of incomplete lineage sorting and introgression among closely related taxa, confounding phylogenetic signal and complicating phylogenomic inference. This thesis focuses on the evolutionary history of the highly diverse and rapidly radiating genus *Tillandsia* (family Bromeliaceae). This genus is the largest and most diverse within the family and presents a variety of ecological adaptations to various habitats, mostly epiphytic and saxicolous. The first main chapter reports the design and efficacy of a target capture bait set specific for Bromeliaceae, devised to address evolutionary questions at both micro- and macro-evolutionary scales. The set targets 1,776 single- and multi-copy genes, including some putatively associated with key innovation traits. Using a wide sampling of members across the family with a particular focus on *Tillandsia*, the taxon-specific bait set was compared to a universal bait set, the Angiosperms353 kit. The comparison considered the kits' performance for both phylogenomics and population genetics, with the expectation that the taxon-specific kit will obtain higher information content and provide increased utility in rapidly radiating clades, compared with the universal kit. Despite retrieving higher information content in the taxon-specific kit, we found that the universal kit performed comparatively well for phylogenomics inference. However, the taxon-specific kit is preferable for detailed investigations, such as estimation of gene tree incongruence and population genomics analyses. The second main chapter explores phylogenomic relationships among 32 species in the clade *Tillandsia*, focusing on rapidly-radiating central-American taxa from subgenus *Tillandsia*. Using whole-genome resequencing data, we inferred relationships among species and explored the evolutionary history of the clade using tree-based and network-based approaches. We found a lack of monophyly and deviations from a tree-like structure, coupled with rampant gene tree discordance. In contrast to classical hypotheses, the uncovered pervasive hybridization along

the phylogeny suggests that the expansion of the subgenus from South into Central America proceeded in several migration waves, followed by episodes of diversification and gene flow. Focusing on hybridization, Patterson's D (ABBA-BABA) and related statistics were used to describe the rates and timing of introgression events. Finally, we discussed the possible contribution of interspecific gene flow to adaptive trait shifts. In summary, the thesis generates novel insights into the evolutionary processes that contribute to a Neotropical rapid radiation and offers future prospects for research in Bromeliaceae at large and specifically in *Tillandsia*.

Zusammenfassung

Das Verständnis der Triebkräfte und Zwänge der Artenbildung steht im Mittelpunkt der Evolutionsforschung. Die Prozesse, die zur Entstehung neuer Arten führen, beinhalten oft die Besiedlung neuartiger ökologischer Nischen und das Auftreten neuartiger Merkmale, die durch ein kompliziertes Zusammenspiel von Umwelt, Phänotyp und Genotyp erreicht werden. Die moderne Hochdurchsatzsequenzierung (HTS) ermöglicht es, phylogenomische Beziehungen zu klären und den genomischen Fußabdruck der Evolutionsgeschichte zu untersuchen. Die Erforschung dieser Themen bei den jüngsten schnellen Ausbreitungen stellt jedoch eine besondere Herausforderung dar: Eine kurze Diversifizierungsphase erhöht die Wahrscheinlichkeit einer unvollständigen Sortierung der Abstammungslinien und der Introgression zwischen eng verwandten Taxa und verwirrt gleichzeitig die phylogenomischen Signale, was die Schlussfolgerungen erschwert. Diese Arbeit konzentriert sich auf die Evolutionsgeschichte der äußerst vielfältigen und sich schnell ausbreitenden Gattung *Tillandsia* (Familie Bromeliaceae). Diese Gattung ist die größte und vielfältigste innerhalb der Familie und weist eine Vielzahl ökologischer Anpassungen an verschiedene Lebensräume auf, die meist epiphytisch und felsbewohnend sind.

Das erste Hauptkapitel beschreibt das Design und die Wirksamkeit eines Target-Capture-Bait-Sets für Bromeliaceae, das speziell entwickelt wurde, um evolutionäre Fragen sowohl auf mikro- als auch auf makro-evolutionärer Ebene zu beantworten. Das Set zielt auf 1.776 Single-Copy- und Multi-Copy-Gene in allen Bromeliaceae ab, darunter auch Gene, die vermutlich mit wichtigen innovativen Merkmalen in Verbindung stehen. Anhand einer breiten Stichprobe von Mitgliedern der Familie mit Schwerpunkt auf *Tillandsia* wurde das taxonspezifische Bait-Set mit einem universellen Bait-Set, dem Angiosperms353-Kit, verglichen. Bei dem Vergleich wurde die Leistung der Kits sowohl für die Phylogenomik als auch für die Populationsgenetik berücksichtigt, wobei die Erwartung bestand, dass das taxonspezifische Kit im Vergleich zum universellen Kit einen höheren Informationsgehalt und einen größeren Nutzen vor allem für die Forschung auf Populationsebene in sich schnell ausbreitenden Kladen bieten würde. Trotz des höheren Informationsgehalts des taxonspezifischen Kits zeigte das universelle Kit eine vergleichsweise gute Leistung bei phylogenomischen Schlussfolgerungen. Für detaillierte Untersuchungen, wie z. B. die Abschätzung der Genbauminkongruenz und populationsgenomische Analysen, ist das taxonspezifische Kit jedoch vorzuziehen.

Im zweiten Hauptkapitel werden die phylogenomischen Beziehungen zwischen 32 Arten der Gattung *Tillandsia* untersucht, wobei der Schwerpunkt auf der mittelamerikanischen Untergattung *Tillandsia* liegt. Unter Verwendung von Ganzgenom-Resequenzierungsdaten haben wir die Beziehungen zwischen den Arten abgeleitet und die Evolutionsgeschichte der Gattung mit Hilfe eines baumbasierten Ansatzes untersucht. Wir stellten fest, dass es keine Monophylie und keine baumähnliche Struktur gibt, und dass der Genbaum sehr uneinheitlich ist. Im Gegensatz zu klassischen Hypothesen deutet die aufgedeckte weit verbreitete Hybridisierung entlang der Phylogenie darauf hin, dass die Ausbreitung der Untergattung nach Mittelamerika in mehreren Migrationswellen erfolgte, gefolgt von Episoden der Diversifizierung und des Genflusses. Mit Blick auf die Hybridisierung wurden Pattersons D (ABBA-BABA), verwandte Statistiken und ein modellbasiertes phylogenetisches Netzwerk verwendet, um die Raten und den Zeitpunkt von Introgressionsereignissen zu beschreiben. Schließlich wurde der mögliche Beitrag des intraspezifischen Genflusses zu adaptiven Merkmalsverschiebungen diskutiert. Zusammenfassend lässt sich sagen, dass diese Arbeit neue Erkenntnisse über die evolutionären Prozesse liefert, die zu einer raschen neotropischen Radiation beitragen, und Zukunftsperspektiven für die Forschung an Bromeliaceae im Allgemeinen und an *Tillandsia* im Besonderen bietet.

Introduction

The genetics of rapid evolutionary radiations

Biodiversity has transformed from an idiosyncratic research topic of highly specialised scholars to an issue at the center of social, economic and political discourses (Escobar, 1998; Gowdy, 1997; Shih et al., 2020; Watermeyer et al., 2021). The aspiration to understand the mechanisms that promote speciation dates back to biologists of the evolutionary synthesis and remains pivotal to evolutionary biology (Brown, 2014; Dobzhansky, 1937; Lewin et al., 2018; Mayr, 1999; Pennisi, 2005; Terborgh, 2015). ‘What drives biological diversity?’ is a question of multifaceted, interdisciplinary nature: diversity differs among taxonomic groups and geographical regions and is determined by a complex interplay of past events, ecological heterogeneity, phenotypic features and genetic makeup (Brown, 2014; Hillebrand, 2004; Linder, 2008; Rabosky, 2009).

Rapid evolutionary radiations, defined as the generation of a large number of species over a short time, give researchers the opportunity to link inherent traits and extrinsic environmental processes. Much attention has been given to rapid radiations, which are so pervasive among the Tree of Life some suggest they generated most of the extant biodiversity (Givnish, 2015; Linder, 2008; Rabosky, 2009; Rundell & Price, 2009; Wiens, 2017). Within the flourishing investigation of rapid radiations several taxa had been thoroughly studied, like the African cichlid fishes (Brawand et al., 2014; Malinsky et al., 2015; Wagner et al., 2012), *Heliconius* butterflies (Merrill et al., 2015; Van Belleghem et al., 2021) and *Anolis* lizards (Losos, 2009; Losos, 1994; Stroud & Losos, 2020). Comparable, albeit more limited research focused on plant groups, for example the New World genus *Lupinus* (Contreras-Ortiz et al., 2018; Drummond et al., 2012), the Neotropical family Bromeliaceae (Givnish et al., 2014; Palma-Silva & Fay, 2020), American live oaks (Cavender-Bares et al., 2015; Eaton et al., 2015) and New Caledonian persimmons (Paun et al., 2016; Samuel et al., 2019). These investigations of plant radiations already provided examples to the mechanisms of diversification and yet much remains unexplored (de La Harpe et al., 2017; Linder, 2008; Suarez-Gonzalez et al., 2018).

When a rapid increase in speciation rates is driven by the adaptation of organisms to new ecological contexts, the process is termed adaptive radiation. Modern definitions of

adaptive radiations developed and gained controversy, but all generally focus on an increase in ecological diversity (Glor, 2010; Schenk, 2021; Schluter, 2000). Schluter (2000) posed four criteria for the diagnosis of adaptive radiations: common ancestry, phenotype-environment correlation, trait utility and rapid speciation. The first and fourth criteria were more deeply studied and fewer studies attempted to present functional evidence of a trait utility and to identify significant correlation between phenotype and environment (Losos & Miles, 2002; Schenk, 2021). Arguably, studies that focus on the tempo of diversification may be overlooking the ecological aspects of adaptive radiation (Givnish, 2015; Schenk, 2021).

In order to describe the evolution of radiations through time, Naciri & Linder (2020) offered a model focused on population genetic structure and trait genetic architecture:

- Initially, a taxonomic lineage disperses into a novel adaptive zone which is usually a new geographic area. This stage may be driven by prior acquisition of traits that allow that lineage to prosper or avoid stress in the novel habitat (Klak et al., 2004). Evolutionary biologists often accompany this with the assumption of relaxed biotic competition – either competitors for the same niche have not yet arrived or antagonists like herbivores or diseases are absent. These hypotheses are termed the ecological release and the enemy release hypothesis, respectively (Hughes et al., 2015; Yoder et al., 2010).
- In the next two steps, the new lineage undergoes (1) geographical or (2) ecological expansion or both. Range expansion happens rapidly and is assumed to be associated with the appearance of novel adaptive traits, especially in the case of adaptive radiations. These adaptations are termed ‘key innovations’, defined as traits that allow a lineage to utilize novel niches (Stroud & Losos, 2016). Populations across the new geographical range experience differing rates and types of selection, population bottleneck, isolation by distance (IBD) or by environment (IBE; Prunier et al., 2017; Sexton et al., 2014).
- As a consequence, populations that adapted sufficiently survive while others perish. The remaining populations may speciate by achieving reproductive isolation or retain genetic connectivity (Lexer et al., 2005; Linck & Battey, 2019; Nei et al., 1983; Nosil, 2008).

The conditions that favour rapid diversification have been rigorously explored in the last few decades with the quick progression of new sequencing technologies. Molecular

biologists studying rapid radiations usually address several challenges: (1) recognizing the evolutionary processes that drive or limit divergence; (2) exploring the balance between selection and neutral processes and (3) addressing the ‘radiation paradox’: How can lineages that expand from small genetic pools develop, in a short amount of time, enough genetic variation to create abundant phenotypic diversity? (Brawand et al., 2014; Gillespie, 2016; Hidalgo et al., 2017; Naciri & Linder, 2020; Pease et al., 2016). Since genomic processes occur across both deep and shallow timescales, addressing questions on the genomic basis of radiations requires connecting two conceptual worlds: macro-evolutionary approaches based on phylogenetics and micro-evolutionary approaches based on population genetics (de La Harpe et al., 2017).

Numerous, mutually-inclusive processes were found to facilitate genetic differentiation in rapid radiations, especially after long-distance dispersal. The roles of standing ancestral variation and hybridization were at the focus of research in the last decades. Both processes are virtually universal in rapid radiations (Novikova et al., 2016; Wu et al., 2018). The persistence of ancestral allelic polymorphism during radiation (incomplete lineage sorting, henceforth ILS) contributes to genetic variability during phases of selection and stochastic allele fixation. Additionally, ample evidence suggests that hybridization and gene flow increase genetic variability by exchange of genetic composition between incipient species. The introduction of alleles upon which selection has already acted is expected to enable diversification through adaptive introgression (Abbott et al., 2013; Meier et al., 2017; Pease et al., 2016; Suarez-Gonzalez et al., 2018). Both whole genome and gene duplications contribute to adaptation by offering genetic variability for selection and the latter may lead to genome structural shifts (Barrier et al., 1999; Couvreur et al., 2010; Soltis & Soltis, 1999). Transposable elements may create genomic changes that generate variability (e.g., Chuong et al., 2015). In cichlid fishes, TE activity was found to alter gene expression during periods of relaxed selection (Brawand et al., 2014; Choudhury & Parisod, 2017). Finally, Epigenetic changes are thought to allow populations to extend the suitable range of habitats, providing more time for other genetic processes to occur. It’s been further suggested that epigenetic changes may generate heritable phenotypic variation (Ashapkin et al., 2016; Flatscher et al., 2012; Paun et al., 2007).

Phylogenomics of rapid radiations

Evolutionary research requires a fundamental understanding of the relationships between studied taxa. Molecular studies joined the field of taxonomy for the task of identifying and delimiting species, applying approaches that investigate the diversity between DNA sequences to infer the level of divergence between taxa. The power to infer the past from contemporary data depends on the type of data obtained, the study system and the properties of the analytical models (Delsuc et al., 2005; Hebert et al., 2003; Philippe et al., 2011). Molecular phylogenies were usually based on sequencing a few highly conserved markers like organellar DNA sequences, microsatellite markers and Sanger-sequenced nuclear genes or genome regions. These remain popular due to the moderate cost efficiencies, the ease of designing the necessary PCR primers and of detecting polymorphism (Barbara et al., 2007; de La Harpe et al., 2017; Schlotteröer et al., 1991; D. X. Zhang & Hewitt, 2003). A single genomic region is then used to infer a gene tree which can be, but often is not an estimate of the ‘true tree’ describing species relationships (Huson & Bryant, 2006). Conserved markers, however, often provide insufficient variation to resolve relationships among closely related taxa. This is especially true for studies of rapidly radiating lineages and incipient species that have recently diverged (Escudero et al., 2020; Straub et al., 2012; Urantowka et al., 2017).

In the last decades, increased accessibility to next-generation sequencing (NGS) transformed phylogenomic approaches and analytical tools advanced accordingly. As molecular biologists now utilize large amounts of multi-locus genomic data, they achieve unprecedented resolution of the relationships between clades across the tree of life (Davey et al., 2011; Escudero et al., 2020; Philippe et al., 2011). Reviewing the current wealth of NGS methods is beyond the scope of this introduction, but generally, the methods can be divided between whole-genome sequencing and reduced representation approaches (e.g., RNA sequencing, target capture and restriction site associated DNA sequencing (RAD-Seq); Baird et al., 2008; de La Harpe et al., 2017; Puritz et al., 2014; Weitemier et al., 2014; Zimmer & Wen, 2013).

A vast volume of data did not, however, immediately clarify the phylogenies of ‘difficult lineages’. First, an increase in the amount of data is coupled with multiplication of technical biases and limitations: random noise, misspecified model parameters, substitutional saturation and errors in data assembly or filtering are a few examples (Kandziora et al., 2022;

Morales-Briones et al., 2021; Philippe et al., 2011). More importantly, extensive genomic data showcases the boundless complexity of evolutionary genomic processes: modern phylogenomic data-sets are sequenced from one or a few individuals per species and alignments of individual loci are used to estimate gene trees. Since recombination decouples genomic regions, the result is a collection of thousands of tree topologies, each representing a different part in the history of a lineage (Hibbins et al., 2020; Kandziora et al., 2022; Wu et al., 2018). Gene tree discordance between the different topologies is unexceptional and indeed is considered quintessential in rapid radiations. An array of successive speciation events result in low sequence variation and increase the probability of ILS, hybridization, and gene duplication and loss, reducing phylogenetic signal and leading to high gene tree discordance (Huson & Bryant, 2006; Maddison, 1997). Some phylogenomic studies approach gene tree discordance as a technical challenge, while others exploit the information it conveys about evolutionary processes (Morales-Briones et al., 2021; Philippe et al., 2011).

The advent of phylogenomics proceeded with novel methods of data-processing and phylogenetic inference from large data-sets. A common approach to the analysis of multiple gene sequences per species is concatenation of sequences and subsequent analysis of the matrix as a single ‘supergene’ (Degnan & Rosenberg, 2009; Gadagkar et al., 2005). Concatenation often results in well-supported phylogenies, however it assumes that all sequences evolved according to one evolutionary tree and may lead to inference of the wrong species tree, especially in the presence of ILS (Degnan & Rosenberg, 2009; Edwards et al., 2007; Kubatko & Degnan, 2007). In contrast, the multi-species coalescence (MSC) addresses phylogenetic reconstruction in the presence of ILS and is considered more robust in several contexts (Edwards et al., 2016; Liu et al., 2015). Furthermore, inference employing the MSC can model gene tree incongruence with methods like BEAST and ASTRAL (Degnan & Rosenberg, 2009; Jiang et al., 2020; Mirarab et al., 2014). The MSC however obtains a non-reticulated species tree which does not account for hybridization and introgression, two relatively common evolutionary processes (but see network extensions: Wen et al., 2016; C. Zhang et al., 2018). In contrast, a few methods can account for ILS and simultaneously infer hybridization explicitly, among them the D-statistic and related statistics (Box 1; Malinsky et al., 2015; Martin & Jiggins, 2017; Patterson et al., 2012), and model-based construction of phylogenetic networks. The latter represents a new frontier in reconstructing evolutionary histories, yet remains computationally challenging and thus only feasible for small data-sets

(Wen et al., 2018; Yu & Nakhleh, 2015). Ultimately, the field of molecular evolution is experiencing more than a mere methodological shift, but rather a philosophical and conceptual one. As evidence accumulates for both data structures and lineages whose history contradicts a simple bifurcating species tree, the predominant use of bifurcating trees to describe diversification is being challenged (Ellstrand, 2014; Mallet et al., 2016; Novikova et al., 2016; Oaks et al., 2022).

Box 1. D-statistics and variations used in the second chapter of this work.

1. **Patterson's D** (ABBA-BABA) – Illustrated in figure 1, a simple test for introgression, robust to the presence of ILS. It exploits asymmetries in the frequencies of nonconcordant gene trees in a three-population tree ordered (((P1,P2),P3),O), where the outgroup O carries the ancestral allele A and the derived allele is denoted as B. Allele frequencies are used to calculate the D-statistic, with values ranging between -1 and 1, indicating an excess of shared derived alleles between P1 and P3 at one end of the range, and between P2 and P3 in the positive range (Patterson et al., 2012; Green et al., 2010).

2. **D_G** – a test for consistency of clade assignment. $D_G(A, G_1; G_2, O)$ is calculated for all possible combinations of individuals, where G_1 and G_2 are assigned to the same clade and A is assigned to a different clade. The results indicate whether individuals assigned to the same clade always share more derived alleles with each-other than with any individual from another clade (Malinsky et al., 2018).

3. **D_{min}** – indicates taxon relationships that deviate from a tree structure. Calculating the minimum value of D across all possible arrangements of a trio (((P1,P2),P3),O), it provides a 'lower bound' estimate of the amount of gene flow within each trio. A positive score signifies the sharing of derived alleles between the species in the trio is inconsistent with a tree structure (Malinsky et al., 2018, 2021).

4. **f_4 -ratio** – known as admixture fraction f , a method for estimating ancestry proportions. The statistics measure correlations in allele frequencies among sets of four populations. Observed values reflect degrees of shared ancestry based on the difference in allele frequencies at biallelic loci, providing a measure for the proportion of the genome involved

in admixture. The f_4 -ratio is a variation informative of admixture between 4 populations P1, P2, P3 and O (Green et al., 2010; Lipson, 2020).

5. $f_b(C)$ – also named f -branch, a statistics that disentangles correlated f_4 -ratio results and assigns gene flow signal to specific, possibly internal, branches on a phylogeny. Using all f -scores on a given tree, $f_b(C)$ captures excess allele sharing involving branches, including internal ones. By utilizing multiple f_4 -ratio calculations, it provides additional information about the timing of introgression events (Malinsky et al., 2018, 2021).

6. f_{dM} - developed to test for introgression in genomic windows, assesses whether the admixture signal is confined to specific loci. As a modification of the f_d statistic, it can quantify shared variation between P3 and P2 (positive values) or between P3 and P1 (negative values). The calculation of f_{dM} further depends on the frequency of the derived allele in P1 and P2 (Malinsky et al., 2015; 2021).

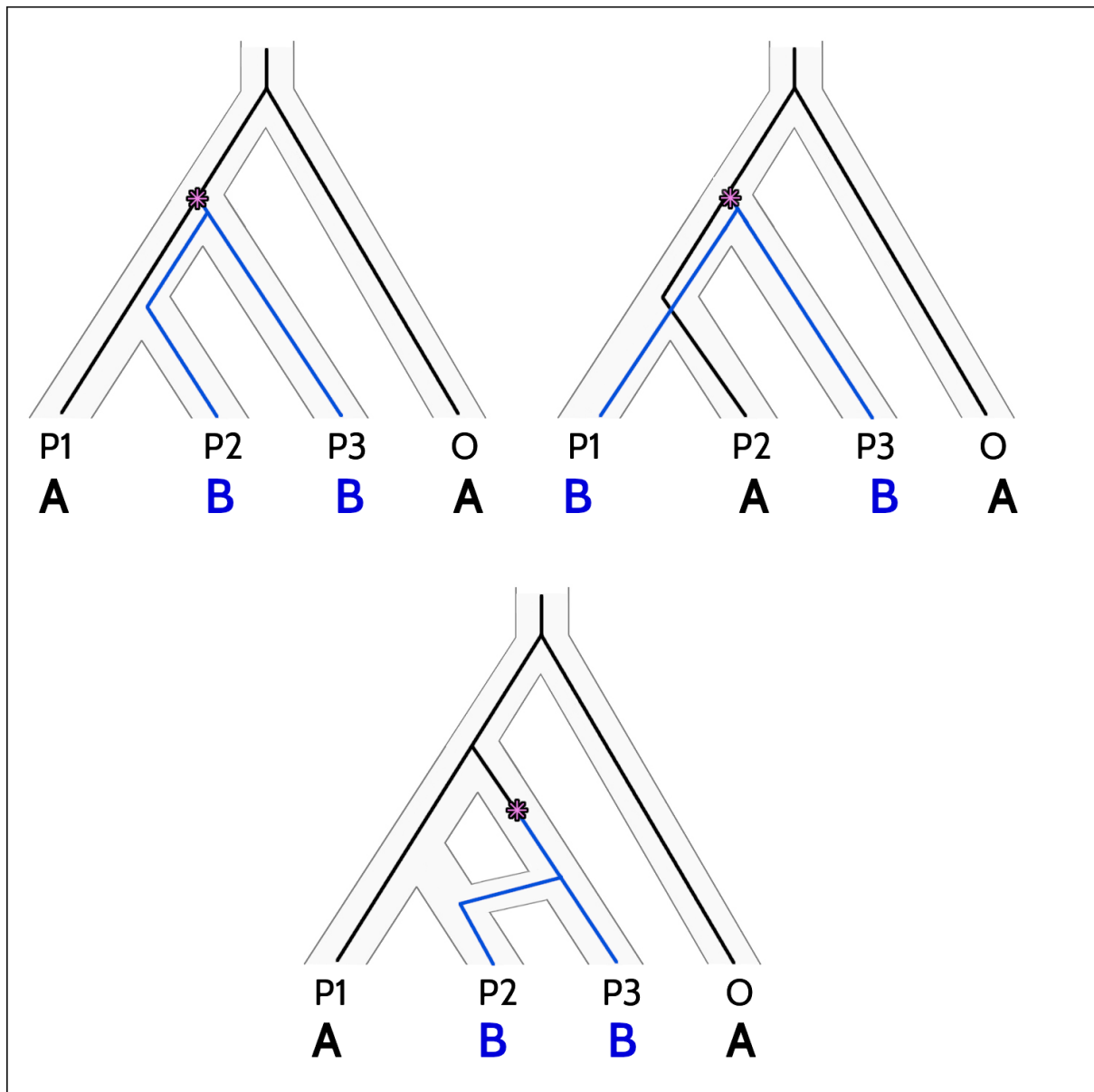


Figure 1. Sharing of derived alleles is informative of gene tree topologies. Above: ABBA sites are those in which P2 and P3 share a derived allele (star), while P1 features the ancestral allele. Under a null hypothesis of only ILS, the frequencies of ABBA and BABA patterns are expected to be equal. Below: Introgression between P2 and P3 will lead to increased frequency of ABBA patterns and positive values of D (see Box 1; adapted from Malinsky et al., 2021).

Study group

The present thesis is dedicated to the plant genus *Tillandsia* (family Bromeliaceae). The taxon was chosen for its impressively high phenotypic diversity, young crown age and unique characteristics. Furthermore, this study complemented larger group efforts to develop molecular resources for studying *Tillandsia*, envisioning it as an emerging model group for studying Neotropical radiations.

Bromeliaceae, a Neotropical radiation of flowering plants, was described by Benzing (2000) as a family that “exemplifies botanical radiation in the extreme” due to its remarkable diversity. This conspicuous monocot angiosperm family of >3,000 species is characterized by unique appearance, with distinctive leaf rosettes that often impound rainwater in central tanks (phytotelmata; figure 2). With high abundance in tropical regions and exquisite beauty, bromeliads fascinated botanists and amateurs for centuries and are still rigorously described, delimited and studied today (Benzing, 2000; Mez, 1904; Picado-Twight, 1913; Schimper, 1888; Tietze, 1906). Bromeliaceae thus constitutes an excellent research system for studying the drivers and constraints of adaptive radiations, especially the evolution of Neotropical biota (Givnish et al., 2011; Palma-Silva & Fay, 2020). According to fossil-calibrated phylogenies, the family originated in the Guayana Shield c. 100 Mya and most of the extant taxa in the family expanded beyond their ancestral distribution, quickly diversifying over a short period of 10-15 Mya. Occurring from northern Patagonia to southeastern USA with a single species in West Africa, bromeliads occupy a range of habitats, from moist mountainous rainforests to dry, rocky sites. Their diversification is associated with orogenic changes, the most notable of which is the uplift of the Andes, and with acquisition of ‘key innovations’, functional traits that allowed them to prosper in diverse adaptive zones (Givnish et al., 2007, 2011; Schulte et al., 2009).

Bromeliad phenotypic adaptations evolved multiple times and in a correlated and contingent manner, creating adaptive syndromes. The association of spatial and temporal patterns in bromeliads allows researchers to hypothesize and test for reciprocity between ecology, habit and rates of diversification (Crayn et al., 2015; Givnish et al., 2014; Schulte et al., 2009). I will further focus on epiphytism which is highly prevalent in *Tillandsia* as in the family altogether: bromeliad epiphytism is associated with humid tropical montane habitats and both are associated with a tank habit, as heavy nutrient-rich rain and detritus derived

from host plants accumulate within the tank. Moreover, the tank hosts a diversity of arthropods and other small animal species and microbes, providing the plants with nutrients through indirect carnivory¹. Epiphytism is further associated with “entangling seeds” that ensure attachment to tall tree branches (Benzing, 2008; Gentry & Dodson, 1987; Givnish et al., 1984, 2014; Romero et al., 2010). Inversely, in dry, terrestrial sites, epiphytism is rather associated with adaptations to prolonged periods of drought. The most commonly observed adaptations are the evolution of CAM photosynthesis, leaf succulence and a dense cover of water-and-nutrient-absorptive trichomes, all which reduce transpiration and provide longer periods of carbon uptake in xeric environments (Benzing, 2008; Givnish et al., 2014; Koch et al., 2020; Mez, 1904). The correlated occurrence of these phenotypes produce the ‘atmospheric’ epiphytic habit, one which is especially common in *Tillandsia* (see below). Notably, CAM evolution is also common in terrestrial bromeliads and at low elevations (Benzing & Renfrow, 1974; Koch et al., 2020). Finally, bromeliads repeatedly developed ornithophilous pollination syndromes which in turn are associated with fertile, humid habitats (Kessler et al., 2020; Krömer et al., 2006). Numerous adaptations favour high rates of diversification: as epiphytism provides a range of micro-niches over a dynamic habitat, entangled seeds offer greater dispersal ability and hummingbird pollination increases diversification, possibly through pollination specialization, limitation of pollen transfer and differential speciation rates mediated through phenology and animal behavior (Gentry & Dodson, 1987; Givnish et al., 2014; Kessler et al., 2020; Krömer et al., 2006).

Bromeliad phylogenetic classification is dynamic and new species and genera are periodically described. Regardless, phylogenetic studies consistently recognize eight subfamilies, out of which Tillandsioideae is the largest with <1,400 species. Diversification in this family is estimated to be associated with the uplift of the Andes (Barfuss et al., 2016; Givnish et al., 2007). Subject to ongoing taxonomic revisions, 6-9 genera are recognized in the subfamily and *Tillandsia* is the most diverse and species-rich with approximately 600 species of predominantly epiphytic plants (Barfuss et al., 2016; Till, 2000).

Archaeological evidence suggests that *Tillandsia* were of great cultural importance before their description by European scientists: *T. usneoides* (‘Spanish moss’) was found in

¹ And the bromeliad research community with stories of the miscellaneous bites and stings they accumulated during fieldwork.

pottery-making societies of the Archaic period in North America and was used for fruit wrapping and pillow filling (Gilmore, 2015; Pierce, 2000). *Tillandsia* was used ritually as decoration and depicted on pottery in religious contexts - probably as early as the Aztec culture - and its ceremonial use continues to this day (Arslanian et al., 1986; Bye Jr, 1979; Gilmore, 2015; Pierce, 2000). The economic and medical uses of *Tillandsia* are historically documented and extend into the 20th century, from the use of leaf decoction to control diabetes and as a contraceptive agent, to modern Brazilian and US farmers using *Tillandsia* as cattle feed (Delaporte et al., 2004; Keller et al., 1981). The family and genus are both of current horticultural interest. Meanwhile, conservation concerns increase due to unsustainable harvesting practices (Estrella-Parra et al., 2019; Hietz, 2005; Wolf & Konings, 2001). Research in this genus therefore holds cultural relevance and importance for conservation.

Several morphological characters were historically employed for generic classification of the taxon (Gardner, 1986; Mez, 1904). Most noteworthy is the work of Gardner (1982) who used flower morphological characters to circumscribe *Tillandsia* into five taxonomic groups (see also: Gardner, 1986). Revisions of *Tillandsia* classification continued as its taxonomy developed amid the age of genomics. Past phylogenetic work on *Tillandsia* was based on a few conserved plastid and nuclear markers and often lacked subgeneric resolution. The taxon's low genetic divergence was early suggested to be representative of an exceptionally young radiation (<5Mya; Barfuss et al., 2005). Barfuss et al. (2016) constructed a sub-family multi-locus phylogeny using four plastid loci, the nuclear DNA gene PHYC and re-evaluated morphology. While they identified subgeneric clades within *Tillandsia*, species relationships were generally poorly supported. Other studies continued to achieve sub-generic resolution and meanwhile encountered taxonomic issues like fuzzy clade and species boundaries (Chew et al., 2010; Pinzón et al., 2016). Recently, Granados Mendoza et al. (2017) expanded the sampling of Barfuss et al. (2016) to cover ca. 30% of the known species' diversity in *Tillandsia*. Their work confirmed two monophyletic *Tillandsia* clades, characterized by a distinct geographic distribution: a clade containing species from subgenus *Tillandsia*, distributed in North and Central America (clade K), and a South and Central American clade containing species from subgenera *Aerobia*, *Anoplophytum*, *Diaphoranthema* and *Phytarrhiza* (clade Q). They estimated that clade K diversified in North- and Central America, whereas within the second most migratory events occurred from the Andes to the Brazilian shield.

Tillandsia generally present the a similar range of phenotypic diversity as that found in the entire family, with great variation in life habits, photosynthetic pathways, pollination syndromes, and presence or absence of absorptive trichomes, among other presumably adaptive traits (Barfuss et al., 2005; Crayn et al., 2015; Givnish et al., 2014; Till, 2000). The combination of traits most commonly result in one of two phenotypes: the “green” *Tillandsia* are adapted to cool, moist habitats, exhibiting predominantly C3 photosynthesis, a pronounced tank habit, low trichome density and no pronounced leaf succulence. The so-called “grey” *Tillandsia* are recognized by adaptations to xeric environments with a strong tendency for CAM photosynthesis, dense leaf trichome cover and pronounced leaf succulence (figure 2; Benzing, 2000; de La Harpe et al., 2020). While instances of intermediate or ‘mixed’ phenotypes are ubiquitous in the genus, the high rates of correlated diversification and the multiple shifts between phenotypic states suggests that *Tillandsia* had undergone adaptive radiation. Little is known about the evolutionary processes and genomic basis of shifts in these adaptive traits: the ancestor of *Tillandsia* is thought to have been mesic and most likely exhibited C3 photosynthesis. Photosynthetic shifts were found to be correlated with changes in rates of gene duplication and loss (de La Harpe et al., 2020). As chromosome numbers are overall stable within the genus, trait shifts most likely did not involve radical genome rearrangements (de La Harpe et al., 2020; Till, 2000; but see Clara Groot Crego et al., in preparation).



Figure 2. Examples of *Tillandsia* species with varying photosynthetic syndromes. **A**, *T. leiboldiana* and **B**, *T. complanata*, are predominantly C3 species with a typical tank habit. **C**, *T. caput-medusae* and **D**, *T. ionantha*, predominantly CAM species (Crayn et al., 2015; de La Harpe et al., 2020). [photographs: A, *Tillandsia leiboldiana* in the Botanischer Garten, Berlin-Dahlem by Krzysztof Ziarnek used under CC BY-NC-SA 4.0. B, Ecuador, Tungurahua, Quebrada Chamanapamba, Rio Ulba by Alexey Yakovlev under CC BY-SA 2.0. C, *Tillandsia caput-medusae* by KENPEI under CC BY-SA 3.0. D, Botanical specimen in the Lyman Plant House, Smith College, Northampton, Massachusetts, USA by Daderot under public domain (CC0 1.0).

Research objectives

This thesis is aimed at exploring the evolutionary history of *Tillandsia* and the genomic processes that accompanied diversification. To that end, we made use of two NGS techniques, target capture sequencing and whole-genome resequencing. In addition, the advantages and limitations of both methods are highlighted in the context of phylogenomics of recent, rapid radiations.

(1) **Chapter 1** examines the potential of target capture sequencing in evolutionary studies in Bromeliaceae with a focus on *Tillandsia*. The chapter describes the design of a taxon-specific bait set for Bromeliaceae and a comparison to a universal bait set for all flowering plants. We sought to bridge the methodological gap between micro- and macro-evolutionary approaches and evaluated the sets' performance in both contexts. Using a family-wide sampling of 72 Bromeliaceae samples, including five species of *Tillandsia* subgenus *Tillandsia*, we performed phylogenomic inference and explored population structure and demographic processes in *Tillandsia*. Our aims were to generate a toolkit for the bromeliad research community, to add knowledge on the utility of target capture in studying rapid radiations, and to present a road-map for the possible use of a single data-set in addressing both macro- and micro- evolutionary questions.

(2) **Chapter 2** employs a tree-based approach to study the evolutionary history of *Tillandsia*. We used whole-genome re-sequencing of 65 individuals representing 32 species, with a focus on clade K of Mexican *Tillandsia* and additional sampling of South American *Tillandsia*. The chapter describes phylogenetic inference and exploration of *Tillandsia* evolutionary history employing a tree-based approach, D-statistics and a network approach. The relationships between the sampled species were inferred using a concatenation-based and a species tree methods, obtaining phylogenies that contradicted previous phylogenetic work on *Tillandsia*. We proceeded to assess the extent of variation in tree topologies along the genomes and deviations from a tree structure. We finally characterized patterns of hybridization between and within different clades, and constructed a species network to understand the evolutionary history of this highly diverse clade. This chapter illuminates the history of *Tillandsia* and provides a case study of exploring genomic processes in a recent and rapid radiation.

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., ... Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, 26(2), 229–246. <https://doi.org/10.1111/j.1420-9101.2012.02599.x>
- Arslanian, R. L., Stermitz, F. R., & Castedo, L. (1986). 3-methoxy-5-hydroxyflavonols from *Tillandsia purpurea*. *Journal of Natural Products*, 49(6), 1177–1178.
- Ashapkin, V. V., Kutueva, L. I., & Vanyushin, B. F. (2016). Epigenetic variability in plants: Heritability, adaptability, evolutionary significance. *Russian Journal of Plant Physiology*, 63(2), 181–192.
- 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(10), e3376.
- Barbara, T., Palma-silva, C., Paggi, G. M., Bered, F., Fay, M. F., & Lexer, C. (2007). Cross-species transfer of nuclear microsatellite markers: Potential and limitations. *Molecular Ecology*, 16(18), 3759–3767.
- Barfuss, M. H. J., Samuel, R., Till, W., & Stuessy, T. F. (2005). Phylogenetic relationships in subfamily Tillandsioideae (Bromeliaceae) based on DNA sequence data from seven plastid regions. *American Journal of Botany*, 92(2), 337–351.
- Barfuss, M. H. J., Till, W., Leme, E. M. C., Pinzón, J. P., Manzanares, J. M., Halbritter, H., Samuel, R., & Brown, G. K. (2016). Taxonomic revision of Bromeliaceae subfam. Tillandsioideae based on a multi-locus DNA sequence phylogeny and morphology. *Phytotaxa*, 279(1), 1–97. <https://doi.org/10.11646/phytotaxa.279.1.1>
- 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. *Molecular Biology and Evolution*, 16(8), 1105–1113. <https://doi.org/10.1093/oxfordjournals.molbev.a026200>
- Benzing, D. H. (2000). *Bromeliaceae: Profile of an Adaptive Radiation*. Cambridge University Press.
- Benzing, D. H. (2008). *Vascular epiphytes: General biology and related biota*. Cambridge University Press.

Benzing, D. H., & Renfrow, A. (1974). The mineral nutrition of Bromeliaceae. *Botanical Gazette*, 135(4), 281–288.

Brawand, D., Wagner, C. E., Li, Y. I., Malinsky, M., Keller, I., Fan, S., Simakov, O., Ng, A. Y., Lim, Z. W., & Bezault, E. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature*, 513(7518), 375–381.

Brown, J. H. (2014). Why are there so many species in the tropics? *Journal of Biogeography*, 41(1), 8–22. <https://doi.org/10.1111/jbi.12228>

Bye Jr, R. A. (1979). Hallucinogenic plants of the Tarahumara. *Journal of Ethnopharmacology*, 1(1), 23–48.

Cavender-Bares, J., González-Rodríguez, A., Eaton, D. A., Hipp, A. A., Beulke, A., & Manos, P. S. (2015). Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): A genomic and population genetics approach. *Molecular Ecology*, 24(14), 3668–3687.

Chew, T., De Luna, E., & González, D. (2010). Phylogenetic relationships of the pseudobulbous *Tillandsia* species (Bromeliaceae) inferred from cladistic analyses of ITS 2, 5.8 S ribosomal RNA gene, and ETS sequences. *Systematic Botany*, 35(1), 86–95.

Choudhury, R. R., & Parisod, C. (2017). *Jumping genes: Genomic ballast or powerhouse of biological diversification*. Wiley Online Library.

Chuong, E. B., Elde, N. C., & Feschotte, C. (2017). Regulatory activities of transposable elements: From conflicts to benefits. *Nature Reviews Genetics*, 18(2), 71–86.

Contreras-Ortiz, N., Atchison, G. W., Hughes, C. E., & Madriñán, S. (2018). Convergent evolution of high elevation plant growth forms and geographically structured variation in Andean *Lupinus* (Fabaceae). *Botanical Journal of the Linnean Society*, 187(1), 118–136. <https://doi.org/10.1093/botlinnean/box095>

Couvreur, T. L. P., Franzke, A., Al-Shehbaz, I. A., Bakker, F. T., Koch, M. A., & Mummenhoff, K. (2010). Molecular Phylogenetics, Temporal Diversification, and Principles of Evolution in the Mustard Family (Brassicaceae). *Molecular Biology and Evolution*, 27(1), 55–71. <https://doi.org/10.1093/molbev/msp202>

- Crayn, D. M., Winter, K., Schulte, K., & Smith, J. A. C. (2015). Photosynthetic pathways in Bromeliaceae: Phylogenetic and ecological significance of CAM and C3 based on carbon isotope ratios for 1893 species. *Botanical Journal of the Linnean Society*, 178(2), 169–221. <https://doi.org/10.1111/boj.12275>
- Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, 12(7), 499–510.
- de La Harpe, M., Paris, M., Hess, J., Barfuss, M. H. J., Serrano-Serrano, M. L., Ghatak, A., Chaturvedi, P., Weckwerth, W., Till, W., Salamin, N., Wai, C. M., Ming, R., & Lexer, C. (2020). Genomic footprints of repeated evolution of CAM photosynthesis in a Neotropical species radiation. *Plant, Cell & Environment*, 43(12), 2987–3001. <https://doi.org/10.1111/pce.13847>
- de La Harpe, M., Paris, M., Karger, D. N., Rolland, J., Kessler, M., Salamin, N., & Lexer, C. (2017). Molecular ecology studies of species radiations: Current research gaps, opportunities and challenges. *Molecular Ecology*, 26(10), 2608–2622. <https://doi.org/10.1111/mec.14110>
- Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24(6), 332–340. <https://doi.org/10.1016/j.tree.2009.01.009>
- Delaporte, R. H., Sarragiotto, M. H., Takemura, O. S., Sanchez, G. M., & Nakamura, C. V. (2004). Evaluation of the antioedematogenic, free radical scavenging and antimicrobial activities of aerial parts of *Tillandsia streptocarpa* Baker–Bromeliaceae. *Journal of Ethnopharmacology*, 95(2–3), 229–233.
- Delsuc, F., Brinkmann, H., & Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. *Nature Reviews Genetics*, 6(5), 361–375. <https://doi.org/10.1038/nrg1603>
- Dobzhansky, T. G. (1937). *Genetics and the origin of species*. New York City. NY Columbia University Press.
- Drummond, C. S., Eastwood, R. J., Miotto, S. T. S., & Hughes, C. E. (2012). Multiple Continental Radiations and Correlates of Diversification in *Lupinus* (Leguminosae): Testing for Key Innovation with Incomplete Taxon Sampling. *Systematic Biology*, 61(3), 443–460. <https://doi.org/10.1093/sysbio/syr126>
- Eaton, D. A., Hipp, A. L., González-Rodríguez, A., & Cavender-Bares, J. (2015). Historical introgression among the American live oaks and the comparative nature of tests for introgression. *Evolution*, 69(10), 2587–2601.

Edwards, S. V., Liu, L., & Pearl, D. K. (2007). High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences*, 104(14), 5936–5941. <https://doi.org/10.1073/pnas.0607004104>

Edwards, S. V., Xi, Z., Janke, A., Faircloth, B. C., McCormack, J. E., Glenn, T. C., Zhong, B., Wu, S., Lemmon, E. M., Lemmon, A. R., Leaché, A. D., Liu, L., & Davis, C. C. (2016). Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. *Molecular Phylogenetics and Evolution*, 94, 447–462. <https://doi.org/10.1016/j.ympev.2015.10.027>

Ellstrand, N. C. (2014). Is gene flow the most important evolutionary force in plants? *American Journal of Botany*, 101(5), 737–753.

Escobar, A. (1998). Whose knowledge, whose nature? Biodiversity, conservation, and the political ecology of social movements. *Journal of Political Ecology*, 5(1), 53–82.

Escudero, M., Nieto-Feliner, G., Pokorny, L., Spalink, D., & Viruel, J. (2020). Editorial: Phylogenomic Approaches to Deal With Particularly Challenging Plant Lineages. *Frontiers in Plant Science*, 11. <https://doi.org/10.3389/fpls.2020.591762>

Estrella-Parra, E., Flores-Cruz, M., Blancas-Flores, G., Koch, S. D., & Alarcón-Aguilar, F. J. (2019). The Tillandsia genus: History, uses, chemistry, and biological activity. *BOLETÍN LATINOAMERICANO Y DEL CARIBE DE PLANTAS MEDICINALES Y AROMÁTICAS*, 18(3), 239–264.

Flatscher, R., Frajman, B., Schönswetter, P., & Paun, O. (2012). Environmental heterogeneity and phenotypic divergence: Can heritable epigenetic variation aid speciation? *Genetics Research International*, 2012.

Gadagkar, S. R., Rosenberg, M. S., & Kumar, S. (2005). Inferring species phylogenies from multiple genes: Concatenated sequence tree versus consensus gene tree. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, 304(1), 64–74. <https://doi.org/10.1002/jez.b.21026>

Gardner, C. S. (1982). *A systematic study of Tillandsia subgenus Tillandsia*. Texas A&M University.

Gardner, C. S. (1986). Preliminary classification of Tillandsia based on floral characters. *Selbyana*, 9(1), 130–146.

Gentry, A. H., & Dodson, C. H. (1987). Diversity and biogeography of neotropical vascular epiphytes. *Annals of the Missouri Botanical Garden*, 74(2), 205–233.

Gillespie, R. G. (2016). Island time and the interplay between ecology and evolution in species diversification. *Evolutionary Applications*, 9(1), 53–73.

Gilmore, Z. I. (2015). Direct radiocarbon dating of Spanish moss (*Tillandsia usneoides*) from early fiber-tempered pottery in the southeastern US. *Journal of Archaeological Science*, 58, 1–8.

Givnish, T. J. (2015). Adaptive radiation versus ‘radiation’ and ‘explosive diversification’: Why conceptual distinctions are fundamental to understanding evolution. *New Phytologist*, 207(2), 297–303. <https://doi.org/10.1111/nph.13482>

Givnish, T. J., Barfuss, M. H. J., Ee, B. V., Riina, R., Schulte, K., Horres, R., Gonsiska, P. A., Jabaily, R. S., Crayn, D. M., Smith, J. A. C., Winter, K., Brown, G. K., Evans, T. M., Holst, B. K., Luther, H., Till, W., Zizka, G., Berry, P. E., & Sytsma, K. J. (2011). Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: Insights from an eight-locus plastid phylogeny. *American Journal of Botany*, 98(5), 872–895. <https://doi.org/10.3732/ajb.1000059>

Givnish, T. J., Barfuss, M. H. J., Ee, B. V., Riina, R., Schulte, K., Horres, R., Gonsiska, P. A., Jabaily, R. S., Crayn, D. M., Smith, J. A. C., Winter, K., Brown, G. K., Evans, T. M., Holst, B. K., Luther, H., Till, W., Zizka, G., Berry, P. E., & Sytsma, K. J. (2014). Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Molecular Phylogenetics and Evolution*, 71, 55–78. <https://doi.org/10.1016/j.ympev.2013.10.010>

Givnish, T. J., Burkhardt, E. L., Happel, R. E., & Weintraub, J. D. (1984). Carnivory in the Bromeliad *Brocchinia reducta*, with a Cost/Benefit Model for the General Restriction of Carnivorous Plants to Sunny, Moist, Nutrient-Poor Habitats. *The American Naturalist*, 124(4), 479–497. <https://doi.org/10.1086/284289>

Givnish, T. J., Millam, K. C., Berry, P. E., & Sytsma, K. J. (2007). Phylogeny, adaptive radiation, and historical biogeography of Bromeliaceae inferred from *ndhF* sequence data. *Aliso: A Journal of Systematic and Floristic Botany*, 23(1), 3–26.

Glor, R. E. (2010). Phylogenetic insights on adaptive radiation. *Annual Review of Ecology, Evolution, and Systematics*, 41, 251–270.

Gowdy, J. M. (1997). The value of biodiversity: Markets, society, and ecosystems. *Land Economics*, 25–41.

- Granados Mendoza, C., Granados-Aguilar, X., Donadío, S., Salazar, G. A., Flores-Cruz, M., Hagsater, E., Starr, J. R., Ibarra-Manríquez, G., Fragoso-Martínez, I., & Magallón, S. (2017). Geographic structure in two highly diverse lineages of *Tillandsia* (Bromeliaceae). *Botany*, 95(7), 641–651. <https://doi.org/10.1139/cjb-2016-0250>
- 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., ... Pääbo, S. (2010). A Draft Sequence of the Neandertal Genome. *Science*, 328(5979), 710–722. <https://doi.org/10.1126/science.1188021>
- Han, T.-S., Zheng, Q.-J., Onstein, R. E., Rojas-Andrés, B. M., Hauenschild, F., Muellner-Riehl, A. N., & Xing, Y.-W. (2020). Polyploidy promotes species diversification of *Allium* through ecological shifts. *New Phytologist*, 225(1), 571–583. <https://doi.org/10.1111/nph.16098>
- Hebert, P. D., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1512), 313–321.
- Hibbins, M. S., Gibson, M. J., & Hahn, M. W. (2020). Determining the probability of hemiplasy in the presence of incomplete lineage sorting and introgression. *eLife*, 9, e63753. <https://doi.org/10.7554/eLife.63753>
- Hidalgo, O., Viales, D., Vallès, J., Garnatje, T., Siljak-Yakovlev, S., Leitch, I. J., & Pellicer, J. (2017). Cytogenetic insights into an oceanic island radiation: The dramatic evolution of pre-existing traits in *Cheirolophus* (Asteraceae: Cardueae: Centaureinae). *Taxon*, 66(1), 146–157.
- Hietz, P. (2005). Conservation of Vascular Epiphyte Diversity in Mexican Coffee Plantations. *Conservation Biology*, 19(2), 391–399. <https://doi.org/10.1111/j.1523-1739.2005.00145.x>
- Hillebrand, H. (2004). On the Generality of the Latitudinal Diversity Gradient. *The American Naturalist*, 163(2), 192–211. <https://doi.org/10.1086/381004>
- Hughes, C. E., Nyffeler, R., & Linder, H. P. (2015). Evolutionary plant radiations: Where, when, why and how? *New Phytologist*, 207(2), 249–253. <https://doi.org/10.1111/nph.13523>
- Huson, D. H., & Bryant, D. (2006). Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution*, 23(2), 254–267. <https://doi.org/10.1093/molbev/msj030>

Jiang, X., Edwards, S. V., & Liu, L. (2020). The multispecies coalescent model outperforms concatenation across diverse phylogenomic data sets. *Systematic Biology*, 69(4), 795–812.

Kandziora, M., Sklenář, P., Kolář, F., & Schmickl, R. (2022). How to Tackle Phylogenetic Discordance in Recent and Rapidly Radiating Groups? Developing a Workflow Using *Loricaria* (Asteraceae) as an Example. *Frontiers in Plant Science*, 12, 765719. <https://doi.org/10.3389/fpls.2021.765719>

Keller, W. J., Bourn, W. M., & Bonfiglio, J. F. (1981). A folk medicine for diabetes mellitus. *Quarterly Journal of Crude Drug Research*, 19(2–3), 49–51.

Kessler, M., Abrahamczyk, S., & Krömer, T. (2020). The role of hummingbirds in the evolution and diversification of Bromeliaceae: Unsupported claims and untested hypotheses. *Botanical Journal of the Linnean Society*, 192(4), 592–608. <https://doi.org/10.1093/botlinnean/boz100>

Klak, C., Reeves, G., & Hedderson, T. (2004). Unmatched tempo of evolution in Southern African semi-desert ice plants. *Nature*, 427(6969), 63–65.

Koch, M. A., Stock, C., Kleinpeter, D., del Río, C., Osses, P., Merklinger, F. F., Quandt, D., & Siegmund, A. (2020). Vegetation growth and landscape genetics of *Tillandsia lomas* at their dry limits in the Atacama Desert show fine-scale response to environmental parameters. *Ecology and Evolution*, 10(23), 13260–13274. <https://doi.org/10.1002/ece3.6924>

Krömer, T., Kessler, M., & Herzog, S. K. (2006). Distribution and Flowering Ecology of Bromeliads along Two Climatically Contrasting Elevational Transects in the Bolivian Andes 1. *Biotropica: The Journal of Biology and Conservation*, 38(2), 183–195.

Kubatko, L. S., & Degnan, J. H. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, 56(1), 17–24.

Lewin, H. A., Robinson, G. E., Kress, W. J., Baker, W. J., Coddington, J., Crandall, K. A., Durbin, R., Edwards, S. V., Forest, F., & Gilbert, M. T. P. (2018). Earth BioGenome Project: Sequencing life for the future of life. *Proceedings of the National Academy of Sciences*, 115(17), 4325–4333.

Lexer, C., Fay, M. F., Joseph, J. A., Nica, M.-S., & Heinze, B. (2005). Barrier to gene flow between two ecologically divergent *Populus* species, *P. alba* (white poplar) and *P. tremula* (European aspen): The role of ecology and life history in gene introgression. *Molecular Ecology*, 14(4), 1045–1057.

- Linck, E., & Battey, C. J. (2019). On the relative ease of speciation with periodic gene flow (p. 758664). *bioRxiv*. <https://doi.org/10.1101/758664>
- Linder, H. P. (2008). Plant species radiations: Where, when, why? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1506), 3097–3105. <https://doi.org/10.1098/rstb.2008.0075>
- Lipson, M. (2020). Applying f4-statistics and admixture graphs: Theory and examples. *Molecular Ecology Resources*, 20(6), 1658–1667.
- Liu, L., Xi, Z., & Davis, C. C. (2015). Coalescent Methods Are Robust to the Simultaneous Effects of Long Branches and Incomplete Lineage Sorting. *Molecular Biology and Evolution*, 32(3), 791–805. <https://doi.org/10.1093/molbev/msu331>
- Losos, J. B. (2009). (2009) *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles*. University of California Press, Berkeley, CA.
- Losos, J. B. (1994). Integrative Approaches to Evolutionary Ecology: Anolis Lizards as Model Systems. *Annual Review of Ecology and Systematics*, 25(1), 467–493. <https://doi.org/10.1146/annurev.es.25.110194.002343>
- Losos, J. B., & Miles, D. B. (2002). Testing the hypothesis that a clade has adaptively radiated: iguanid lizard clades as a case study. *The American Naturalist*, 160(2), 147–157. <https://doi.org/10.1086/341557>
- Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology*, 46(3), 523–536.
- 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(6267), 1493–1498. <https://doi.org/10.1126/science.aac9927>
- Malinsky, M., Matschiner, M., & Svardal, H. (2021). Dsuite—Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, 21(2), 584–595. <https://doi.org/10.1111/1755-0998.13265>
- 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. *Nature Ecology & Evolution*, 2(12), 1940–1955. <https://doi.org/10.1038/s41559-018-0717-x>

- Mallet, J., Besansky, N., & Hahn, M. W. (2016). How reticulated are species? *BioEssays*, 38(2), 140–149. <https://doi.org/10.1002/bies.201500149>
- Martin, S. H., & Jiggins, C. D. (2017). Interpreting the genomic landscape of introgression. *Current Opinion in Genetics & Development*, 47, 69–74. <https://doi.org/10.1016/j.gde.2017.08.007>
- Mayr, E. (1999). *Systematics and the origin of species, from the viewpoint of a zoologist*. Harvard University Press.
- Meier, J. I., Marques, D. A., Mwaiko, S., Wagner, C. E., Excoffier, L., & Seehausen, O. (2017). Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature Communications*, 8(1), 14363. <https://doi.org/10.1038/ncomms14363>
- Merrill, R. M., Dasmahapatra, K. K., Davey, J. W., Dell’Aglia, D. D., Hanly, J. J., Huber, B., Jiggins, C. D., Joron, M., Kozak, K. M., & Llaurens, V. (2015). The diversification of *Heliconius* butterflies: What have we learned in 150 years? *Journal of Evolutionary Biology*, 28(8), 1417–1438.
- Mez, K. (1904). Physiologische Bromeliaceen-studien. *Jahrbücher für Wissenschaftliche Botanik* 40:157–229.
- Mirarab, S., Reaz, R., Bayzid, Md. S., Zimmermann, T., Swenson, M. S., & Warnow, T. (2014). ASTRAL: Genome-scale coalescent-based species tree estimation. *Bioinformatics*, 30(17), i541–i548. <https://doi.org/10.1093/bioinformatics/btu462>
- Morales-Briones, D. F., Kadereit, G., Tefarikis, D. T., Moore, M. J., Smith, S. A., Brockington, S. F., Timoneda, A., Yim, W. C., Cushman, J. C., & Yang, Y. (2021). Disentangling Sources of Gene Tree Discordance in Phylogenomic Data Sets: Testing Ancient Hybridizations in *Amaranthaceae* s.l. *Systematic Biology*, 70(2), 219–235. <https://doi.org/10.1093/sysbio/syaa066>
- Naciri, Y., & Linder, H. P. (2020). The genetics of evolutionary radiations. *Biological Reviews*, 95(4), 1055–1072. <https://doi.org/10.1111/brv.12598>
- Nei, M., Maruyama, T., & Wu, C.-I. (1983). Models of evolution of reproductive isolation. *Genetics*, 103(3), 557–579.
- Nosil, P. (2008). Speciation with gene flow could be common. *Molecular Ecology*, 27(16), 3301–3316.

Novikova, P. Y., Hohmann, N., Nizhynska, V., Tsuchimatsu, T., Ali, J., Muir, G., Guggisberg, A., Paape, T., Schmid, K., Fedorenko, O. M., Holm, S., Säll, T., Schlötterer, C., Marhold, K., Widmer, A., Sese, J., Shimizu, K. K., Weigel, D., Krämer, U., ... Nordborg, M. (2016). Sequencing of the genus *Arabidopsis* identifies a complex history of nonbifurcating speciation and abundant trans-specific polymorphism. *Nature Genetics*, 48(9), 1077–1082. <https://doi.org/10.1038/ng.3617>

Oaks, J. R., Wood, P. L., Siler, C. D., & Brown, R. M. (2022). Generalizing Bayesian phylogenetics to infer shared evolutionary events (p. 2021.07.23.453597). *bioRxiv*. <https://doi.org/10.1101/2021.07.23.453597>

Palma-Silva, C., & Fay, M. F. (2020). Bromeliaceae as a model group in understanding the evolution of Neotropical biota. *Botanical Journal of the Linnean Society*, 192(4), 569–586. <https://doi.org/10.1093/botlinnean/boaa003>

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(3), 1065–1093.

Paun, O., Fay, M. F., Soltis, D. E., & Chase, M. W. (2007). Genetic and epigenetic alterations after hybridization and genome doubling. *Taxon*, 56(3), 649–656. <https://doi.org/10.2307/25065849>

Paun, O., Turner, B., Trucchi, E., Munzinger, J., Chase, M. W., & Samuel, R. (2016). Processes Driving the Adaptive Radiation of a Tropical Tree (Diospyros, Ebenaceae) in New Caledonia, a Biodiversity Hotspot. *Systematic Biology*, 65(2), 212–227. <https://doi.org/10.1093/sysbio/syv076>

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 Biology*, 14(2), e1002379. <https://doi.org/10.1371/journal.pbio.1002379>

Pennisi, E. (2005). What determines species diversity? *Science*, 309(5731), 90–90.

Philippe, H., Brinkmann, H., Lavrov, D. V., Littlewood, D. T. J., Manuel, M., Wörheide, G., & Baurain, D. (2011). Resolving difficult phylogenetic questions: Why more sequences are not enough. *PLoS Biology*, 9(3), e1000602.

Picado-Twight, C. (1913). Les broméliacées épiphytes considérées comme milieu biologique. Las bromeliáceas epífitas consideradas como medio biológico. *Bulletin Scientifique de La France et de La Belgique*, 47(3), 215–360.

- Pierce, S. (2000). The use of *Tillandsia* species in ritual adornment in Qosqo, Peru. *Journal of Bromeliad Society*, 50, 195–201.
- Pinzón, J. P., Ramírez-Morillo, I. M., Carnevali, G., Barfuss, M. H., Till, W., Tun, J., & Ortiz-Díaz, J. J. (2016). Phylogenetics and evolution of the *Tillandsia utriculata* complex (Bromeliaceae, Tillandsioideae) inferred from three plastid DNA markers and the ETS of the nuclear ribosomal DNA. *Botanical Journal of the Linnean Society*, 181(3), 362–390.
- Prunier, R., Akman, M., Kremer, C. T., Aitken, N., Chuah, A., Borevitz, J., & Holsinger, K. E. (2017). Isolation by distance and isolation by environment contribute to population differentiation in *Protea repens* (Proteaceae L.), a widespread South African species. *American Journal of Botany*, 104(5), 674–684.
- Puritz, J. B., Matz, M. V., Toonen, R. J., Weber, J. N., Bolnick, D. I., & Bird, C. E. (2014). Demystifying the RAD fad. *Molecular Ecology*, 24(23), 5937–5942.
- Rabosky, D. L. (2009). Ecological limits and diversification rate: Alternative paradigms to explain the variation in species richness among clades and regions. *Ecology Letters*, 12(8), 735–743.
- Romero, G., Q., Nomura, F., Gonçalves, A. Z., Dias, N. Y., Mercier, H., Conforto, E. de C., & Rossa-Feres, D. de C. (2010). Nitrogen fluxes from treefrogs to tank epiphytic bromeliads: An isotopic and physiological approach. *Oecologia*, 162(4), 941–949.
- Rundell, R. J., & Price, T. D. (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology & Evolution*, 24(7), 394–399. <https://doi.org/10.1016/j.tree.2009.02.007>
- Samuel, R., Turner, B., Duangjai, S., Munzinger, J., Paun, O., Barfuss, M. H. J., & Chase, M. W. (2019). Systematics and evolution of the Old World Ebenaceae, a review with emphasis on the large genus *Diospyros* and its radiation in New Caledonia. *Botanical Journal of the Linnean Society*, 189(2), 99–114. <https://doi.org/10.1093/botlinnean/boy081>
- Schenk, J. J. (2021). The Next Generation of Adaptive Radiation Studies in Plants. *International Journal of Plant Sciences*, 182(4), 245–262. <https://doi.org/10.1086/713445>
- Schimper, A. F. W. (1888). *Die epiphytische vegetation Amerikas* (Issue 2). G. Fischer.
- Schlotteröer, C., Amos, B., & Tautz, D. (1991). Conservation of polymorphic simple sequence loci in cetacean species. *Nature*, 354(6348), 63–65.

Schluter, D. (2000). *The ecology of adaptive radiation*. OUP Oxford.

Schulte, K., Barfuss, M. H. J., & Zizka, G. (2009). Phylogeny of Bromelioideae (Bromeliaceae) inferred from nuclear and plastid DNA loci reveals the evolution of the tank habit within the subfamily. *Molecular Phylogenetics and Evolution*, 51(2), 327–339. <https://doi.org/10.1016/j.ympev.2009.02.003>

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by environment or distance: Which pattern of gene flow is most common? *Evolution*, 68(1), 1–15.

Shih, W. Y., Mabon, L., & De Oliveira, J. A. P. (2020). Assessing governance challenges of local biodiversity and ecosystem services: Barriers identified by the expert community. *Land Use Policy*, 91, 104291.

Soltis, D. E., & Soltis, P. S. (1999). Polyploidy: Recurrent formation and genome evolution. *Trends in Ecology & Evolution*, 14(9), 348–352.

Straub, S. C. K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R. C., & Liston, A. (2012). Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *American Journal of Botany*, 99(2), 349–364. <https://doi.org/10.3732/ajb.1100335>

Stroud, J. T., & Losos, J. B. (2016). Ecological Opportunity and Adaptive Radiation. *Annual Review of Ecology, Evolution, and Systematics*, 47(1), 507–532. <https://doi.org/10.1146/annurev-ecolsys-121415-032254>

Stroud, J. T., & Losos, J. B. (2020). Bridging the Process-Pattern Divide to Understand the Origins and Early Stages of Adaptive Radiation: A Review of Approaches With Insights From Studies of Anolis Lizards. *Journal of Heredity*, 111(1), 33–42. <https://doi.org/10.1093/jhered/esz055>

Suarez-Gonzalez, A., Lexer, C., & Cronk, Q. C. B. (2018). Adaptive introgression: A plant perspective. *Biology Letters*, 14(3), 20170688. <https://doi.org/10.1098/rsbl.2017.0688>

Terborgh, J. W. (2015). Toward a trophic theory of species diversity. *Proceedings of the National Academy of Sciences*, 112(37), 11415–11422.

Tietze, M. (1906). Physiologische Bromeliaceen Studien II. *Die Entwicklung Der Wasseraufnehmenden Bromeliaceen-Trichome*. *Z. Naturwissen*, 78, 1–50.

Till, W. (2000). *Tillandsioideae*. In D.H. Benzing, ed. *Bromeliaceae: Profile of an Adaptive Radiation* (p. Pp. 555-571). Cambridge University Press.

- Urantowka, A. D., Krocak, A., & Mackiewicz, P. (2017). The influence of molecular markers and methods on inferring the phylogenetic relationships between the representatives of the Arini (parrots, Psittaciformes), determined on the basis of their complete mitochondrial genomes. *BMC Evolutionary Biology*, 17(1), 166. <https://doi.org/10.1186/s12862-017-1012-1>
- Van Belleghem, S. M., Lewis, J. J., Rivera, E. S., & Papa, R. (2021). Heliconius butterflies: A window into the evolution and development of diversity. *Current Opinion in Genetics & Development*, 69, 72–81.
- Wagner, C. E., Harmon, L. J., & Seehausen, O. (2012). Ecological opportunity and sexual selection together predict adaptive radiation. *Nature*, 487(7407), 366–369. <https://doi.org/10.1038/nature11144>
- Watermeyer, K. E., Guillera-Aroita, G., Bal, P., Burgass, M. J., Bland, L. M., Collen, B., Hallam, C., Kelly, L. T., McCarthy, M. A., & Regan, T. J. (2021). Using decision science to evaluate global biodiversity indices. *Conservation Biology*, 35(2), 492–501.
- Weitemier, K., Straub, S. C. K., Cronn, R. C., Fishbein, M., Schmickl, R., McDonnell, A., & Liston, A. (2014). Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences*, 2(9), 1400042. <https://doi.org/10.3732/apps.1400042>
- Wen, D., Yu, Y., & Nakhleh, L. (2016). Bayesian Inference of Reticulate Phylogenies under the Multispecies Network Coalescent. *PLOS Genetics*, 12(5), e1006006. <https://doi.org/10.1371/journal.pgen.1006006>
- Wen, D., Yu, Y., Zhu, J., & Nakhleh, L. (2018). Inferring Phylogenetic Networks Using PhyloNet. *Systematic Biology*, 67(4), 735–740. <https://doi.org/10.1093/sysbio/syy015>
- Wiens, J. J. (2017). What explains patterns of biodiversity across the Tree of Life? *BioEssays*, 39(3), 1600128. <https://doi.org/10.1002/bies.201600128>
- Wolf, J. H., & Konings, C. J. (2001). Toward the sustainable harvesting of epiphytic bromeliads: A pilot study from the highlands of Chiapas, Mexico. *Biological Conservation*, 101(1), 23–31.
- 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. *Molecular Ecology*, 27(16), 3301–3316.

Yoder, J. B., Clancey, E., Des Roches, S., Eastman, J. M., Gentry, L., Godsoe, W., Hagey, T. J., Jochimsen, D., Oswald, B. P., & Robertson, J. (2010). Ecological opportunity and the origin of adaptive radiations. *Journal of Evolutionary Biology*, 23(8), 1581–1596.

Yu, Y., & Nakhleh, L. (2015). A maximum pseudo-likelihood approach for phylogenetic networks. *BMC Genomics*, 16(10), S10. <https://doi.org/10.1186/1471-2164-16-S10-S10>

Zhang, C., Ogilvie, H. A., Drummond, A. J., & Stadler, T. (2018). Bayesian inference of species networks from multilocus sequence data. *Molecular Biology and Evolution*, 35(2), 504–517.

Zhang, D.-X., & Hewitt, G. M. (2003). Nuclear DNA analyses in genetic studies of populations: Practice, problems and prospects. *Molecular Ecology*, 12(3), 563–584.

Zimmer, E. A., & Wen, J. (2013). Reprint of: Using nuclear gene data for plant phylogenetics: Progress and prospects. *Molecular Phylogenetics and Evolution*, 66(2), 539–550. <https://doi.org/10.1016/j.ympev.2013.01.005>

Chapter 1

Taxon-specific or universal? Using target capture to study the evolutionary history of rapid radiations

Gil Yardeni¹, Juan Viruel², Margot Paris³, Jaqueline Hess^{1,4}, Clara Groot Crego^{1,5}, Marylaure de La Harpe¹, Norma Rivera¹, Michael H. J. Barfuss¹, Walter Till¹, Valeria Guzmán-Jacob⁶, Thorsten Krömer⁷, Christian Lexer¹, Ovidiu Paun¹, Thibault Leroy¹

¹Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria

²Royal Botanic Gardens, Kew, Richmond, UK

³Unit of Ecology & Evolution, Department of Biology, University of Fribourg, Fribourg, Switzerland

⁴Department of Soil Ecology, Helmholtz Centre for Environmental Research, UFZ, Halle (Saale), Germany

⁵Vienna Graduate School of Population Genetics, Vienna, Austria

⁶Biodiversity, Macroecology and Biogeography, University of Goettingen, Göttingen, Germany

⁷Centro de Investigaciones Tropicales, Universidad Veracruzana, Xalapa, Mexico

Christian Lexer, Ovidiu Paun, and Thibault Leroy shared last authorship.

Status: published, *Molecular Ecology Resources*, September 2021. DOI: 10.1111/1755-0998.13523

Contribution: Project planning, data collection, bait kit design, molecular laboratory work, data analysis and interpretation, original draft preparation.

Taxon-specific or universal? Using target capture to study the evolutionary history of rapid radiations

Gil Yardeni¹  | Juan Viruel² | Margot Paris³  | Jaqueline Hess^{1,4} | Clara Groot Crego^{1,5} | Marylaure de La Harpe¹ | Norma Rivera¹ | Michael H. J. Barfuss¹  | Walter Till¹ | Valeria Guzmán-Jacob⁶ | Thorsten Krömer⁷ | Christian Lexer¹  | Ovidiu Paun¹  | Thibault Leroy¹

¹Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria

²Royal Botanic Gardens, Kew, Richmond, UK

³Unit of Ecology & Evolution, Department of Biology, University of Fribourg, Fribourg, Switzerland

⁴Department of Soil Ecology, Helmholtz Centre for Environmental Research, UFZ, Halle (Saale), Germany

⁵Vienna Graduate School of Population Genetics, Vienna, Austria

⁶Biodiversity, Macroecology and Biogeography, University of Göttingen, Göttingen, Germany

⁷Centro de Investigaciones Tropicales, Universidad Veracruzana, Xalapa, Mexico

Correspondence

Gil Yardeni, Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, A-1030 Vienna, Austria.

Email: gil.c.yardeni@gmail.com

Funding information

Christian Lexer professorship, Grant/Award Number: BE772002

Abstract

Target capture has emerged as an important tool for phylogenetics and population genetics in nonmodel taxa. Whereas developing taxon-specific capture probes requires sustained efforts, available universal kits may have a lower power to reconstruct relationships at shallow phylogenetic scales and within rapidly radiating clades. We present here a newly developed target capture set for Bromeliaceae, a large and ecologically diverse plant family with highly variable diversification rates. The set targets 1776 coding regions, including genes putatively involved in key innovations, with the aim to empower testing of a wide range of evolutionary hypotheses. We compare the relative power of this taxon-specific set, Bromeliad1776, to the universal Angiosperms353 kit. The taxon-specific set results in higher enrichment success across the entire family; however, the overall performance of both kits to reconstruct phylogenetic trees is relatively comparable, highlighting the vast potential of universal kits for resolving evolutionary relationships. For more detailed phylogenetic or population genetic analyses, for example the exploration of gene tree concordance, nucleotide diversity or population structure, the taxon-specific capture set presents clear benefits. We discuss the potential lessons that this comparative study provides for future phylogenetic and population genetic investigations, in particular for the study of evolutionary radiations.

KEYWORDS

Bromeliaceae, phylogenomics, plant radiation, population structure, target capture, *Tillandsia*
La captura selectiva de secuencias de ADN ha surgido como una herramienta importante para la filogenética y la genética de poblaciones en taxones no-modelo. Mientras que el desarrollo de sondas de captura específicas para cada taxón requiere un esfuerzo sostenido, las colecciones de sondas universales disponibles pueden tener una potencia disminuida para la reconstrucción de relaciones filogenéticas poco

Christian Lexer, Ovidiu Paun, and Thibault Leroy shared last authorship.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *Molecular Ecology Resources* published by John Wiley & Sons Ltd.

profundas o de radiaciones rápidas. Presentamos aquí un conjunto de sondas para la captura selectiva desarrollado recientemente para Bromeliaceae, una familia de plantas extensa, ecológicamente diversa y con tasas de diversificación muy variables. El conjunto de sondas se centra en 1776 regiones de codificación, incluyendo genes supuestamente implicados en rasgos de innovación clave, con el objetivo de potenciar la comprobación de una amplia gama de hipótesis evolutivas. Comparamos la potencia relativa de este conjunto de sondas diseñado para un taxón específico, Bromeliad1776, con la colección universal Angiosperms353. El conjunto específico da lugar a un mayor éxito de captura en toda la familia. Sin embargo, el rendimiento global de ambos kits para reconstruir árboles filogenéticos es relativamente comparable, lo que pone de manifiesto el gran potencial de los kits universales para resolver las relaciones evolutivas. Para análisis filogenéticos o de genética de poblaciones más detallados, como por ejemplo la exploración de la congruencia de los árboles de genes, la diversidad de nucleótidos o la estructura de la población, el conjunto de captura específico para Bromeliaceae presenta claras ventajas. Discutimos las lecciones potenciales que este estudio comparativo proporciona para futuras investigaciones filogenéticas y de genética de poblaciones, en particular para el estudio de las radiaciones evolutivas.

1 | INTRODUCTION

Targeted sequencing approaches have emerged as a promising tool for studying evolutionary relationships in nonmodel taxa, enabling researchers to retrieve large data sets while requiring few genomic resources (Bossert & Danforth, 2018; Escudero et al., 2020; McDonnell et al., 2021; Soto-Gomez et al., 2019). Using custom baits, the method largely retrieves the same loci across a wide taxonomic scale, obtains comparable and mergeable data sets and may be combined with genome-skimming (Lemmon & Lemmon, 2013; Weitemier et al., 2014). Pre-existing knowledge of the targeted loci further provides opportunities to address specific questions on both deep and shallow timescales (Hale et al., 2020; Lemmon et al., 2012). Finally, the method does not necessarily require a reference genome, is highly cost-effective and, with the ability to sequence herbarium samples, reduces the need for extensive sampling campaigns (Blaimer et al., 2016; Hale et al., 2020; Weitemier et al., 2014). Target capture has been successfully applied to resolve phylogenies in diverse groups, from arthropods such as bees (*Xylocopa*, Blaimer et al., 2016; Apidae, Bossert et al., 2019) and Araneae (Hexathelidae, Hedin et al., 2018) to mammals (Cetacea, McGowen et al., 2020), and in numerous plant groups (*Heuchera*, Folk et al., 2015; Gesneriaceae, Ogutcen et al., 2021; Zingiberales, Sass et al., 2016 to name a few). The method's utility for studies at microevolutionary scales has been to date marginally explored, but several studies have pointed to the ability to analyse genomic diversity and estimate population genomic parameters (Choquet et al., 2019; Christmas et al., 2017; Derrien & Ramos-Onsins, 2020; de La Harpe et al., 2019; Sanderson et al., 2020). Nonetheless, the development of probes for target enrichment may pose several challenges: first, the need to identify regions conserved enough to ensure recovery, yet polymorphic

enough to provide ample information (Soto-Gomez et al., 2019; Villaverde et al., 2018). Second, probe design requires detecting regions without pervasive copy number polymorphism (Kadlec et al., 2017; Lemmon et al., 2012), a particular challenge for angiosperms and other groups, where duplication events are ubiquitous (Van de Peer et al., 2017).

In contrast, universal kits offer an attractive alternative that require reduced efforts to establish, and provide comparable data sets across wider ranges of taxa (Johnson et al., 2019; Kadlec et al., 2017). Such kits were designed to retrieve single-copy markers, for example, in the broad scope of amphibians (Hime et al., 2021), anthozoans (Quattrini et al., 2018), vertebrates (Lemmon et al., 2012) or angiosperms (Johnson et al., 2019). In the latter example, the Angiosperms353 kit is designed to target 353 single-copy genes across angiosperms. So far the kit has been employed successfully in resolving phylogenies, including but not limited to *Nepenthes* (Murphy et al., 2020), *Schefflera* (Shee et al., 2020) and the rapid radiations of *Burmeistera* (Bagley et al., 2020) and *Veronica* (Thomas et al., 2021), establishing the kit as an eminent tool in macroevolutionary research. Its utility at microevolutionary levels is yet to be fully realized, although several works have established its suitability to deliver informative signals at a lower taxonomic level (Beck et al., 2021) and in acquiring population genomics parameters (Slimp et al., 2021). The use of highly conserved markers in a universal kit may, however, limit resolution power. Generally, taxon-specific baits are expected to deliver a higher information content and hence more accurate results (Kadlec et al., 2017), as enrichment success is known to drop with the level of divergence between sequences used for probe design and the targeted taxa (Liu et al., 2019). However, one study comparing the power of the universal Angiosperms353 kit and a taxon-specific kit to resolve phylogenomic relationship in

Cyperaceae reported surprisingly similar performance (Larridon et al., 2020) and similar findings were reported in Malinae (Ufimov et al., 2021) and in Ochnaceae (Shah et al., 2021). It remains to be established whether these findings apply to other taxa and other evolutionary scales, including at population level, where ample genomic variability is required to resolve intraspecific relationships and investigate patterns of genetic differentiation.

Until recently, the technology available to investigate evolutionary questions in rapidly evolving groups featuring high net diversification rates has presented major obstacles, in particular for nonmodel groups. Decreasing costs of sequencing coupled with an ever-growing plethora of bioinformatic tools for data processing and downstream analysis has led to an increase in the use of methods like whole-genome sequencing, RNA sequencing and restriction-site associated DNA sequencing (RAD-Seq) in lieu of traditional methods employing few conserved markers (de La Harpe et al., 2017; McKain et al., 2018; Weitemier et al., 2014; Zimmer & Wen, 2013). Whole-genome sequencing however remains costly, posing barriers for research targeting large numbers of samples, organisms with large genomes and nonmodel organisms for which the availability of high-quality genomic resources is often limited (Hollingsworth et al., 2016; Supple & Shapiro, 2018). While RAD-seq is an affordable alternative and widely used in population genetics, the resulting data sets may fall short when screened for homologous sequences across distantly related lineages (but see, e.g., Heckenhauer et al., 2018). Additionally, RAD-seq is less feasible when using degraded DNA from herbarium samples, and the use of short and inconsistently represented loci across phylogenetic sampling may result in low information content and difficulties in assessing paralogy (Jones & Good, 2016; Lemmon & Lemmon, 2013; McKain et al., 2018).

Rapid evolutionary radiations are key stages in the evolutionary history across the Tree of Life and highly recurrent, hence an essential part of biodiversity research (Gavrillets & Losos, 2009; Givnish et al., 2014; Hughes et al., 2015; Soltis et al., 2019; Soltis & Soltis, 2004). Fast evolving groups provide potent opportunities to investigate important questions in evolutionary biology, such as the interplay between ecological and evolutionary processes in shaping biodiversity. A few notable study systems are the cichlid fish (McGee et al., 2020; Salzburger, 2018), *Heliconius* butterflies (Dasmahapatra et al., 2012; Moest et al., 2020), *Anolis* lizards (McGlothlin et al., 2018; Stroud & Losos, 2020), Darwin's finches (Lamichhaney et al., 2015; Zink & Vázquez-Miranda, 2019), white-eyes birds (Moyle et al., 2009) and New World lupins (Nevado et al., 2016). Nevertheless, much remains unknown about the genomic basis underlying species diversification outside these intensively studied systems.

Research of rapidly diversifying lineages presents several challenges. First, a brief diversification period typically leads to imperfect reproductive barriers and incomplete lineage sorting, reflected in significant gene tree discordance and ambiguous relationships (Degnan & Rosenberg, 2009; Lamichhaney et al., 2015; Pease et al., 2016; Straub et al., 2014). In addition, understanding 'speciation through time' poses a methodological challenge and requires connecting two conceptual worlds: macroevolutionary investigations,

concerned with spatial and ecological patterns over deeper timescales, and microevolutionary approaches, providing insight into the processes acting during population divergence and speciation (Bragg et al., 2016; de La Harpe et al., 2017). Resolving phylogenomic relationships and disentangling the contribution of different genomic processes through time typically require large-scale genomic data sets and thorough taxon sampling efforts (Lemmon & Lemmon, 2013; Linder, 2008; Straub et al., 2012).

Here, we present Bromeliad1776, a new bait set for targeted sequencing, designed to address a wide range of evolutionary hypotheses in Bromeliaceae: from producing robust phylogenies to studying the interplay of genomic processes during speciation and the genetic basis of trait shifts, such as photosynthetic and pollination syndrome. This highly diverse Neotropical radiation provides an excellent research system for studying the drivers and constraints of rapid adaptive radiation (Benzing, 2000; Givnish et al., 2011; Loiseau et al., 2021; Mota et al., 2020; Palma-Silva & Fay, 2020; Wöhrmann et al., 2020). Bromeliaceae as a whole is considered an adaptive radiation (Benzing, 2000; Givnish et al., 2011) and contains several rapidly radiating lineages, most notably within Bromelioideae (Aguirre-Santoro et al., 2020) and Tillandsioideae (Loiseau et al., 2021). It is a species-rich and charismatic monocot family, consisting of over 3000 species, including crops in the genus *Ananas* and other economically important species (Luther, 2008). Members of the family are characterized by a distinctive leaf rosette that often impounds rainwater in central tanks (phytotelmata). A diversity of arthropods and other animal species and microbes reside in bromeliad tanks, in some cases even leading to protocarnivory and other forms of nutrient acquisition (Givnish et al., 1984; Leroy et al., 2016). Bromeliads present a diversity of repeatedly evolving adaptive traits, which allowed them to occupy versatile habitats and ecological niches (Benzing, 2000). CAM photosynthesis, water-absorbing trichomes, formation of tank habit, extensive rates of epiphytism and a diversity of pollination syndromes are some of the adaptations correlated with high rates of diversification within the family (Benzing, 2000; Crayn et al., 2004; Givnish et al., 2014; Kessler et al., 2020; Quezada & Gianoli, 2011).

To assess the utility of the Bromeliad1776 kit, we performed a comparison between our taxon-specific kit and the universal Angiosperms353 kit using several methods across different evolutionary timescales. We present Bromeliad1776 in the light of methodological considerations on bait design, data handling, analyses and other practical considerations.

2 | MATERIALS AND METHODS

2.1 | Custom bait design

Whole-genome sequences and gene models from *Ananas comosus* v.3 (Ming et al., 2015) were used to design a bait set aiming to target (i) single-copy protein coding genes distributed across the whole genome, (ii) genes previously described as associated with key

innovation traits in Bromeliaceae (see below), (iii) markers previously used for phylogenomic inference in Bromeliaceae and (iv) genes orthologous to those in the Angiosperms353 bait set. The 1776 selected genes are detailed in Table S1.

Genes in subset *i* were selected based on genetic diversity parameters calculated using whole-genome sequence and RNAseq data previously published by de La Harpe et al. (2020); data publicly available online at SRA Bioproject (PRJNA649109) with the POPGENOME R package v.2.1.6 (Pfeifer et al., 2014). Genomic regions were retained in this category if they shared at least 70% identity between *A. comosus* and *T. sphaerocephala*, and if they had nucleotide diversity (π) values not exceeding the 90% quantile of the (π) distribution across genes for four *Tillandsia* species (*Tillandsia australis*, *Tillandsia fasciculata*, *Tillandsia floribunda* and *T. sphaerocephala*; data and analysis performed by de La Harpe et al. (2020)). We further excluded genes with a total exonic size smaller than 1100 bp, or individual exons smaller than 120 bp. Next, copy number variation was calculated based on clustering of *A. comosus* and *Tillandsia* transcriptome assemblies to generate three copy number categories—'single copy', 'low copy' (i.e. less than five copies) and 'high copy' (i.e. five or more copies). We included only single-copy genes in the design for bait subset *i*. Finally, we excluded genes that were located in genomic regions outside those assigned to linkage groups in the *A. comosus* reference (Ming et al., 2015). A total of 1243 genes were identified for this part.

The bait subset of genes associated with key innovative traits in Bromeliaceae (subset *ii* above) included (1) genes putatively under positive selection along branches relevant to C3/CAM shifts (de La Harpe et al., 2020), (2) genes that exhibit differential gene expression between CAM and C3 *Tillandsia* species (de La Harpe et al., 2020) and (3) genes putatively associated with photosynthetic and developmental functions, or with flavonoid and anthocyanin biosynthesis, according to the literature (e.g. Goolsby et al., 2018; Ming et al., 2015; Palma-Silva et al., 2016; Wai et al., 2017). *Ananas comosus* genes with the highest match scores (calculated as lowest E-score in BLASTP, Madden (2003)) against the sequences of genes from the literature were added to the bait set (see Table S2 for details). A total of 1612 genes underpinning innovative traits were included in the bait design, regardless of criteria used for subset *i* for size, similarity and duplication rate.

Markers previously used for phylogenomic inference in Bromeliaceae (subset *iii*) were obtained from the literature, spanning 13 genes (e.g. Barfuss et al., 2016; Machado et al., 2020; Schulte et al., 2009, see TS2 for full list). Genes orthologous to those in the Angiosperms353 bait set (Johnson et al., 2019) were identified using the orthologous gene models from *A. comosus* based on gene annotations (Ming et al., 2015) or using BLASTP (Madden, 2003), totalling 281 genes.

Finally, we used a draft genome of *T. fasciculata* (Jaqueline Hess, personal communication) to exclude from all candidates genes that exhibited multiple BLASTN hits, if they have not been previously described as duplicated within the genus (de La Harpe et al., 2020). Specifically, we excluded genes that matched another genomic

sequence of at least 100 bp with high similarity score ($>80\%$) and low *E*-value ($<10^{-5}$). In an additional round of filtering performed by the manufacturer of the final bait set, Arbor Biosciences (Ann Arbor, MI, USA), multicopy genes with sequences that are more than 95% identical were collapsed into a single sequence and baits with more than 70% GC content or containing at least 25% repeated sequences were excluded. In addition, targets including exons smaller than 80 bp were completed with regions flanking the exons according to the *A. comosus* reference genome. The final kit included 1776 genes: 801 genes in subset *i*, 681 genes associated with key innovative traits, 13 genes representing phylogenetic markers and 281 genes orthologous to the Angiosperms353 set. Probes were designed with 57,445 80-mer baits tiling across targets in 2x coverage, targeting approximately 2.3 Mbp. The kit is subsequently referred to as the Bromeliad1776 bait set. Further specifications can be found in Tables S1 and S2 and in the github repository: https://github.com/giyany/Bromeliad1776/tree/main/MS_2021_scripts.

2.2 | Plant material collection

We sampled a total of 70/72 Bromeliaceae samples (for Angiosperms353 and for Bromeliad1776, accordingly; Table S3), including 56 accessions from the Tillandsioideae subfamily and 16 representing the other subfamilies, except Navioideae. The divergence time between Tillandsioideae and subfamily Bromelioideae to which *A. comosus* belongs is estimated at 15 Mya (according to Givnish et al., 2014). Within Tillandsioideae, we sampled 38/40 individuals from five species of the *Tillandsia* subgenus *Tillandsia* ('clade K' in Barfuss et al., 2016; Sampling in Mexican populations illustrated in Figure S1).

2.3 | Library preparation & enrichment

DNA extractions were performed using a modified CTAB protocol (Doyle & Doyle, 1987), purified using Nucleospin® gDNA cleanup kit from Macherey-Nagel (Hudlow et al., 2011) following the supplier's instructions with a twofold elution step and finally quantified with Qubit® 3.0 Fluorometer (Life Technologies).

For each sample, 200 ng DNA was sheared using Bioruptor® Pico sonication device (Diagenode) aiming for an average insert size of 350 bp, dried in a speed vacuum Eppendorf concentrator 5301 (Eppendorf) and eluted in 30 L ddH₂O. Genomic libraries were prepared using the NEBNext® Ultra TM II DNA Library Prep Kit for Illumina® (New England Biolabs) using reagents at half volumes following Hale et al. (2020) and using 11 PCR cycles, increased up to 13 cycled for libraries with low genomic output. Samples were double-indexed with NEBNext® Multiplex Oligos for Illumina® (New England Biolabs). Fragment sizes were inspected with Agilent Bioanalyzer (Agilent Technologies), and concentrations were measured with Qubit® 3.0 Fluorometer. Subpools of 11–14 equimolar genomic libraries were prepared using phylogenetic proximity and

DNA concentrations of the genomic libraries, which ranged from 2.62 to 118.0 ng/L, following Soto-Gomez et al. (2019).

We used the Angiosperms353 and the Bromeliad1776 bait sets from Arbor Biosciences to enrich each subpool of genomic libraries independently with a single hybridization reaction of myBaits® target capture kits from Arbor Biosciences, following Hale et al. (2020). Average fragment size and DNA yield were estimated for each subpool using Agilent Bioanalyzer and Qubit® 3.0 Fluorometer. Subpools were then pooled in equimolar conditions and sequenced at Vienna BioCenter Core Facilities on Illumina® NextSeq™ 550 (2 × 150 bp, Illumina). Sequencing was conducted independently for either bait kit.

2.4 | Data processing

The raw sequence data in BAM format were demultiplexed using `DEML v.1.1.3` (Renaud et al., 2015) and `SAMTOOLS VIEW v.1.7` (Li et al., 2009), converted to fastq using `BAMTOOLS v.2.4.0` (Barnett et al., 2011) and quality checked using `FASTQC v.0.11.7` (Andrews, 2010). Reads were then trimmed for adapter content and quality using `TRIMGALORE v.0.6.5` (Krueger, 2019), a wrapper tool around `FASTQC` and `CUTADAPT`, using settings `--fastqc --retain unpaired`. Sequence quality and adapter removal were confirmed with FastQC reports.

Quality and adapter-trimmed reads were aligned to *A. comosus* reference genome v.3 (Ming et al., 2015) using `BOWTIE2` (Langmead & Salzberg, 2012) with the `--very-sensitive-local` option to increase sensitivity and accuracy. `Samtools` (Li et al., 2009) was then used to remove low-quality mapping and sort alignments by position, and PCR duplicates were marked using `MarkDuplicates` from `PICARDTOOLS v.2.25` (Picard Toolkit, 2019). Summary statistics of the mapping step were generated using `samtools stats`. Variants were called using `FREEBAYES v.1.3.2-dirty` (Garrison & Marth, 2012), and sites marked as MNP/complex were decomposed and normalized using the script 'vcfallelicprimitives' from `VCFLIB` (Garrison, 2012). Next, AN/AC field was calculated using `BCFTOOLS v.1.7` (Li, 2011) and variant calls were filtered using `VCFLIB` (Garrison & Marth, 2012) and `bcftools`. Given that freebayes does not perform automatic variant filtering steps, we identified sets of parameters that generate reliable final SNP sets, based on two independent criteria: the highest transition/transversion ratios as reported by `SNPSIFT` (SnEff Cingolani et al., 2012) and the lowest π_N/π_S (see Section 2.7). After a detailed evaluation, we used the following criteria to generate two high-quality SNP sets, one for each bait set: we considered genotype calls with per-sample coverage below 10×—as missing (NA) and excluded variants (i) marked as indels or neighbouring indels within a distance of 3 bp, (ii) with depth of coverage at the SNP level lower than 500×, (iii) with less than 10 reads supporting the alternate allele at the SNP level or (iv) with more than 40% missing data. All genes in the Bromeliad1776 that passed the filtering criteria were included in the SNP set, regardless of their function. Summary statistics of the final SNP sets were generated using the script `vcf2genocountsmatrix.py`, namely the total number of SNPs, the proportion of on-target SNPs and the proportion of SNPs in some specific genomic contexts, with

A. comosus genome v.3 as a reference. The full data processing script `align_and_trim.sh` and the `vcf2genocountsmatrix.py` script are both available at <https://github.com/giyan/Bromeliad1776>.

2.5 | Bait specificity and efficiency

To explore bait specificity, we calculated the percentage of high-quality trimmed reads on-target using `samtools stats` and `BEDTOOLS INTERSECT v.2.25.0` (Quinlan & Hall, 2010) using the script `calculat_bait_target_specifity.sh` (available from <https://github.com/giyan/Bromeliad1776>). Targets for Bromeliad1776 were defined as the bait sequences plus their 500-bp flanking regions. Targets for Angiosperms353 were defined using orthogroups to *A. comosus*: gene annotations from the bait set were used to assign genes to orthogroups using `ORTHOFINDER` (Emms & Kelly, 2019). When several orthogroups were found for a single Angiosperms353 gene, we included all, resulting in 559 *A. comosus* genes assigned to orthogroups. Within the orthogroups, targets were again defined as exonic regions plus their 500 bp flanking regions.

To provide insights into determinants of bait capture success, we calculated bait efficiency for all baits of Bromeliad1776. For each bait, efficiency was calculated as the number of high-quality reads uniquely mapping to each bait target region, averaged over samples. We then tested for the correlation of capture efficiency to several bait characteristics (copy number, GC content, number and size of exons in targeted gene, size of baits and phylogenetic distance to *A. comosus*) with a generalized linear model or Kruskal–Wallis test in `R v.4.0.3` (R Core Team, 2020) using a negative binomial family.

2.6 | Phylogenomic analyses

We inferred phylogenomic relationships for all samples using two methods: a concatenation method, and a coalescent-based species tree estimation. The latter method was included as concatenation methods do not account for gene tree incongruence, which may result in high support for an incorrect topology (Kubatko & Degnan, 2007), especially in the presence of notable incomplete lineage sorting. In addition, gene tree incongruence analysis provides insight into molecular genome evolution, including the extent of incomplete lineage sorting and other genomic processes such as hybridization and introgression (Galtier & Daubin, 2008; Wendel & Doyle, 1998).

We used the variant and nonvariant genotypes to create a phylip matrix with `vcf2PHYLIP v.2.0` (Ortiz, 2019) and constructed a maximum-likelihood species tree for each bait set with `RAXML-NG v.0.9.0` (Kozlov et al., 2019), using 250 bootstrap replicates and a GTR model with an automatic MRE-based bootstrap convergence test. Next, we constructed a species tree using `ASTRAL-III v.5.7.7` (hereafter: ASTRAL, Zhang et al., 2018). For both the Angiosperms353 and the Bromeliad1776 sets, we separated the matrix into independent genomic windows, defining each window as a gene according to the known exons and a 500-bp flanking region. For Angiosperms353,

we extracted the 559 genes (assigned to orthogroups as explained above) as genomic windows using bedtools intersect. For Bromeliad1776, genomic windows were extracted using the *A. comosus* gene sequences included in bait design. All loci and all accessions were included in species tree inference regardless of the percentage of missing data, since taxon completeness of individual gene trees is important for statistical consistency of this approach, and we expected only low levels of fragmentary sequences (Mirarab, 2019; Nute et al., 2018). After excluding genes with zero coverage, 269 genes and 1600 genes were included in species tree inference for Angiosperms353 and Bromeliad1776, respectively.

For each gene, a maximum-likelihood gene tree was inferred using PARGENES (Morel et al., 2019) with RAXML-NG (Kozlov et al., 2019), using a GTR model with an automatic MRE-based bootstrap convergence test. Loci with insufficient signal may reduce the accuracy of species tree estimation (Mirarab, 2019), hence, in all gene trees, nodes with a bootstrap support smaller than 10 were collapsed using Newick utilities (Junier & Zdobnov, 2010). A species tree was then generated in ASTRAL with quartet support and posterior probability for each tree topology. The number of conflicting gene trees was calculated using phyparts and visualized using the script phypartspiecharts.py (available from <https://github.com/mossmatters/MJPythonNotebooks>).

2.7 | Population structure and nucleotide diversity estimates

To explore the genetic structure within the *Tillandsia* species complex, we focused on five species from 15 localities (Table S3 and Figure S1). We first used PLINK v.1.9 (Chang et al., 2015) to filter out SNPs in linkage disequilibrium. Population structure was further explored through individual ancestry analysis, with identity-by-descent matrix calculated by plink and inference of population structure using ADMIXTURE v.1.3. with *K* values ranging from one to ten, and 30 replicates for each *K*, using a block optimization method (Alexander & Lange, 2011). A summary of the ADMIXTURE results was obtained and presented using pong (Behr et al., 2016). The set of LD-pruned biallelic SNPs was further filtered to allow a maximum of 10% missing data and used to perform a principal components analysis (PCA) with SNPRELATE v.1.20.1 (Zheng et al., 2012). Finally, for each *Tillandsia* species, we used the strategy of Leroy et al. (2021) to compute synonymous (π_s) and non-synonymous (π_n) nucleotide diversities and Tajima's *D*, from fasta sequences using seq_stat_coding (Leroy et al., 2021).

3 | RESULTS

3.1 | Higher mapping rates and capture efficiency for taxon-specific set

On average, 4,401,958 (803,464–12,693,516) paired-end reads per accession were generated per Angiosperms353 library and 2,962,023 (1,282,762–6,298,880) per Bromeliad1776 library.

Overall, the mapping rates to the *A. comosus* reference genome were higher for libraries enriched with Bromeliad1776, with an average mapping rate of 82.3% (61.8%–95.9%) and 42.8% (22.1%–77.9%), for Bromeliad1776 and Angiosperms353, respectively (Figure S2, Table S4). Higher mapping rates were recorded for subfamilies Bromelioideae and Puyoideae, as compared to Tillandsioideae, for both the Angiosperms353 and Bromeliad1776 sets (see Figures S3 and S4, respectively). This may reflect the effect of reference bias, and in the case of Bromeliad1776, it may be further amplified by our kit design based on *A. comosus* (subfamily Bromelioideae). Bait specificity was high for Bromeliad1776 with on average 90.4% reads on-target (76.5%–94.2%), while for Angiosperms353 bait specificity was 14.0% (4.6%–30.1%; see Figure S2). Mapping rates and bait specificity were positively correlated for both bait sets (GLM, $p < .01$).

3.2 | Bait efficiency depends on the genomic context

We investigated factors that may influence bait efficiency, starting with the contribution of gene copy number variation. We assumed three categories regarding the number of paralogs per orthogroup: single-copy, low-copy (i.e. less than five copies) and high-copy (i.e. five or more copies). The number of gene copies had a significant effect on bait efficiency and post hoc Dunn's test supported significant differences in efficiency for comparisons between low-copy and high-copy, and between single-copy and low-copy ($P = 2.8^{-44}$). Low-copy genes exhibit the lowest enrichment success, suggesting that the bait efficiency is not simply correlated with the number of gene copies (Figure 1). We also recovered a significant effect of the intragenic GC content and GC content of the baits on bait efficiency (GLM, $P = 1.5^{-68}$). Finally, we investigated the possible link between efficiency and gene structure. Average exon sizes ($P < 2.0^{-16}$) and total number of exons per gene ($P = 1.1^{-89}$) were also positively correlated with enrichment success. The size of the smallest exon for all targeted genes was however not correlated with bait efficiency. Sequence similarity, measured as per cent of identity between *Tillandsia* sequences and those of *A. comosus*, was positively correlated with capture success ($P = 4.8^{-13}$; Figure 1).

3.3 | Both kits provided a large number of SNPs

After variant calling and filtering, we identified 47,390 and 209,186 high-quality SNPs for the Angiosperms353 and the Bromeliad1776 bait sets, respectively. On average, missing data represented 23.7% of genotype calls per individual in Angiosperms353, but only 6.3% for the Bromeliad1776 kit. The differences in amount of missing data are likely associated with the higher mean depth per site across the Bromeliad1776 kit (6602), as compared to Angiosperms353 (3437). Focusing on the subgenus *Tillandsia*, we identified 15,622 SNPs for Angiosperms353 (including a total of

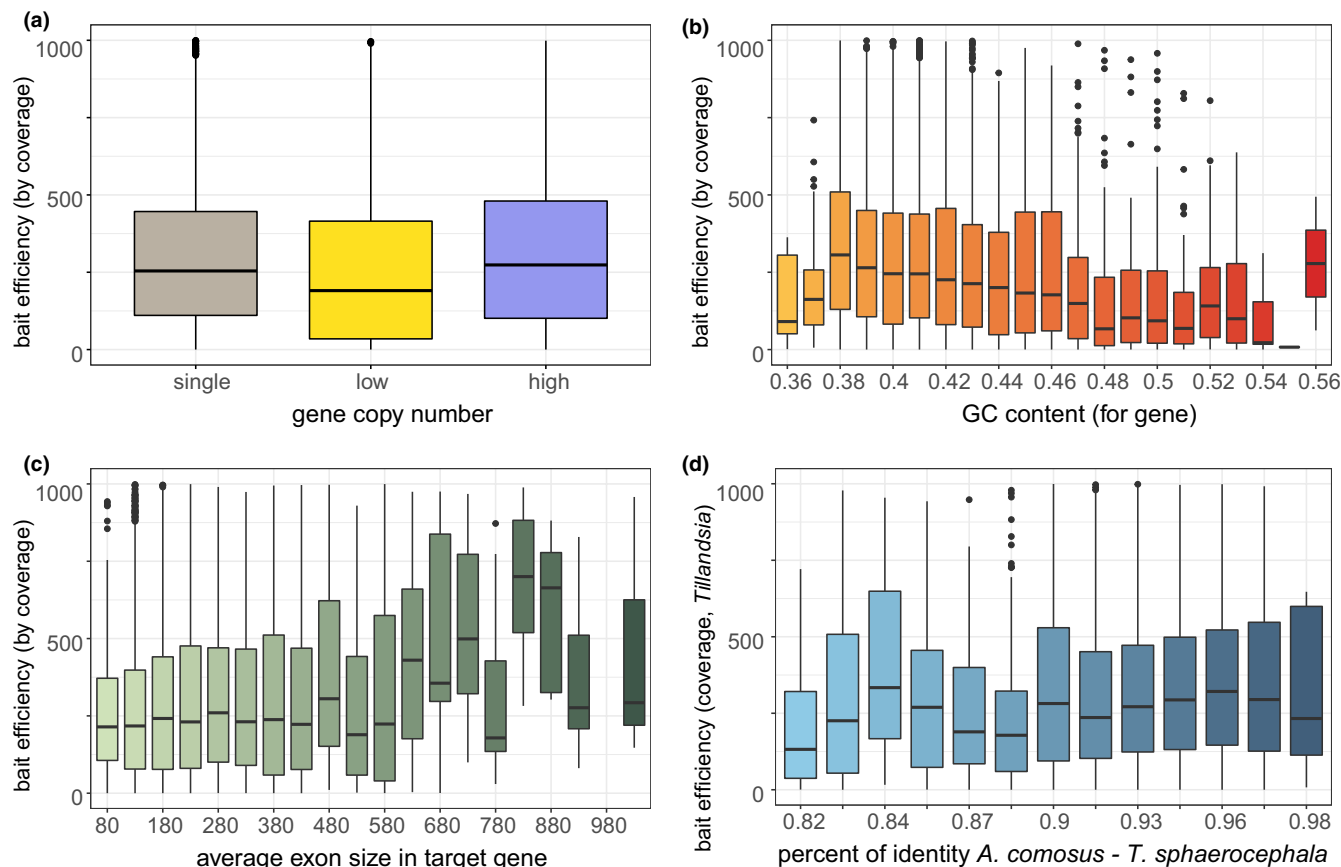


FIGURE 1 Effects of (a) putative gene copy number, (b) gene GC content, (c) average exon size, and (d) per cent of identity on bait efficiency in Bromeliad1776 bait set, measured as the number of high-quality reads uniquely mapping to bait target region across samples. Continuous variable was binned and y-values higher than 1000 excluded for visualization in b–d

18.9% missing data) compared to 65,473 polymorphic sites (2.9% missing data) for Bromeliad1776. In both full data sets and the subset including only *Tillandsia* samples, Bromeliad1776 recovered more variants in intronic regions compared with Angiosperms353. Angiosperms353 recovered a large proportion of off-target SNPs, whereas in Bromeliad1776 approximately 15% of the SNPs were recovered from flanking regions (Table 1). We discuss ascertainment bias that may rise due to the nonrandom selection of markers in the supporting information.

3.4 | Similar phylogenomic resolution in concatenation method, Bromeliad1776 outperforms Angiosperms353 for species tree reconstruction

The Angiosperms353 and Bromeliad1776-based maximum-likelihood phylogenetic trees recovered the same backbone phylogeny of Bromeliaceae, clustering subfamily Tillandsioideae and the subgenus *Tillandsia* with high bootstrap values (Figure S5). Neither set obtained high support for interpopulation structure for *Tillandsia gymnotrya*, but highly supported nodes separated *T. fasciculata* accessions from Mexico and from other locations, and the populations of *T. punctulata* for the Bromeliad1776 data set were

similarly separated. The tree topologies were identical, with the notable exception of the placements of *Tillandsia biflora* and *Racinaea ropalocarpa* and the genus *Deuterocohnia* (Figure S5, purple arrow). Overall, internal nodes are strongly supported for both sets, except for *Hechtia carlsoniae* as sister to Tillandsioideae, which is poorly supported for both sets. While several internal nodes are slightly less supported for the Angiosperms353 set, overall these results demonstrate the efficacy of both kits in phylogenomic reconstruction using concatenation approaches, indicating that as few as 47 k SNPs within variable regions provide reliable information to resolve phylogenetic relationships within the recent evolutionary radiation of *Tillandsia*.

Species trees as inferred with ASTRAL for both data sets likewise provided an overall strong local posterior support (Figure 2, see also Figures S8 and S9). Several nodes however exhibit lower local posterior support values for the Angiosperms353 tree than for the Bromeliad1776 tree. The topology for the Bromeliad1776 ASTRAL tree was similar to the ML tree, but differed again by placing *Deuterocohnia* as sister taxa to *Puyoideae* only. In the Angiosperms353 tree, the topology differed from both ML trees and the ASTRAL Bromeliad1776 tree in several nodes. *H. carlsoniae* was placed as a sister taxa to all other subfamilies in the Angiosperm353 phylogeny. Notably, the placement of *Catopsis* and *Glomeropitcrania* differed,

TABLE 1 Number and characteristics of the variants obtained for Angiosperms353 and Bromeliad1776

	indv Nr.	SNP Nr.	Site mean depth	SNPs in exonic regions	SNPs in intronic regions	SNPs in intergenic regions	On-target SNPs	Flanking SNPs	Off-target SNPs
intra-genic vcf									
Angiosperms353	70	47,390	3447	40,628 (85.7%)	4376 (9.2%)	2386 (5.1%)	8424 (17.8%)	3488 (7.4%)	35,478 (74.8%)
Bromeliad1776	72	209,186	6601.7	170,893 (81.7%)	35,790 (17.1%)	2503 (1.2%)	162,924 (77.9%)	37,661 (18.0%)	8601 (4.11%)
pop-level vcf									
Angiosperms353	38	15,622	1837.8	13,345 (85.5%)	1442 (9.2%)	835 (5.3%)	3032 (19.4%)	1129 (7.22%)	11,461 (73.4%)
Bromeliad1776	40	65,473	3914.9	54,636 (83.5%)	9967 (15.2%)	870 (1.3%)	51,405 (78.5%)	10,588 (16.2%)	3480 (5.3%)

as well as the placement of *Cipurosis subandinai*, *T. biflora* and *R. ropalocarpa*. Several internal nodes were poorly supported, such as the node separating the tribe Catopsideae and core Tillandsioideae, and the nodes separating Tillandsioideae from all other subfamilies. The differences in topology between the Angiosperms353 ASTRAL tree to all other trees (ML trees and Bromeliad1776 ASTRAL tree) together with the low posterior support suggest lower resolution power and a poor fit of this data set for resolving a species tree.

The length and average size of the input gene trees different among sets, with average window length of 304.6 bp and 819.9 bp and average gene tree support of 16.9 and 38.9 for Angiosperms353 and Bromeliad1776 bait sets, respectively (Figure 2). An examination of gene tree concordance constructed with Bromeliad1776 data set allowed us to identify variable levels of gene tree conflict among nodes (Figure 2). Gene tree discordance was especially high for the split between Tillandsioideae and other subfamilies, as well as for the split between Puyoideae and taxa assigned to Bromelioideae. Furthermore, gene tree discordance and the proportion of uninformative gene trees were especially high for splits among clades within the K.1 and K.2 clades of subgenus *Tillandsia*. A similar analysis with Angiosperms353 yielded evidence for gene tree discordance, but a considerable number of gene trees were reported to be noninformative (grey part of the pie charts), especially within subgenus *Tillandsia* (Figure 2).

3.5 | Strong interspecific structure, but little evidence for within-species population structure

After LD-pruning and retaining maximum 10% missing data, 1025 and 32,941 biallelic SNPs were included for the *Tillandsia* PCA analysis of the Angiosperms353 and Bromeliad1776 data sets, respectively. Overall, both data sets provided evidence for interspecific structure, but not for population structure, with Bromeliad1776 resulting in border-line higher resolution (slightly better separating *T. foliosa* from *T. fasciculata*). The percentage of explained variance was higher in the Bromeliad1776 set (19.3% and 16.5% for PC1 and PC2) as compared to the Angiosperms353 data set (14.5% and 11.8%, see Figures 3 and S6). Based on these two PCAs, we found no evidence for spatial genetic structure within each species, since accessions did not cluster by geographic origin on the two PCs presented, or any other PCs we investigated (see Figure S6).

In addition to PCA, we performed ADMIXTURE analyses based on 9804 and 42,613 variants for the Angiosperms353 and Bromeliad1776 sets, respectively (Figure 4). We used a cross-validation strategy to identify the best *K* and found clear support for *K* = 5 for the Bromeliad1776 set (Figure S7). In contrast, the CV pattern for the Angiosperms353 set varied widely, providing limited information about the best *K*. Lowest CV values were however observed for *K* = 9 with locally low values for *K* = 5 and *K* = 3 (Figure S7). We further investigated the ADMIXTURE bar plots at different values of *K*. For *K* = 5, very similar patterns can be observed for both sets, with the recovered clusters reflecting the

Bromeliad1776

Angiosperms353

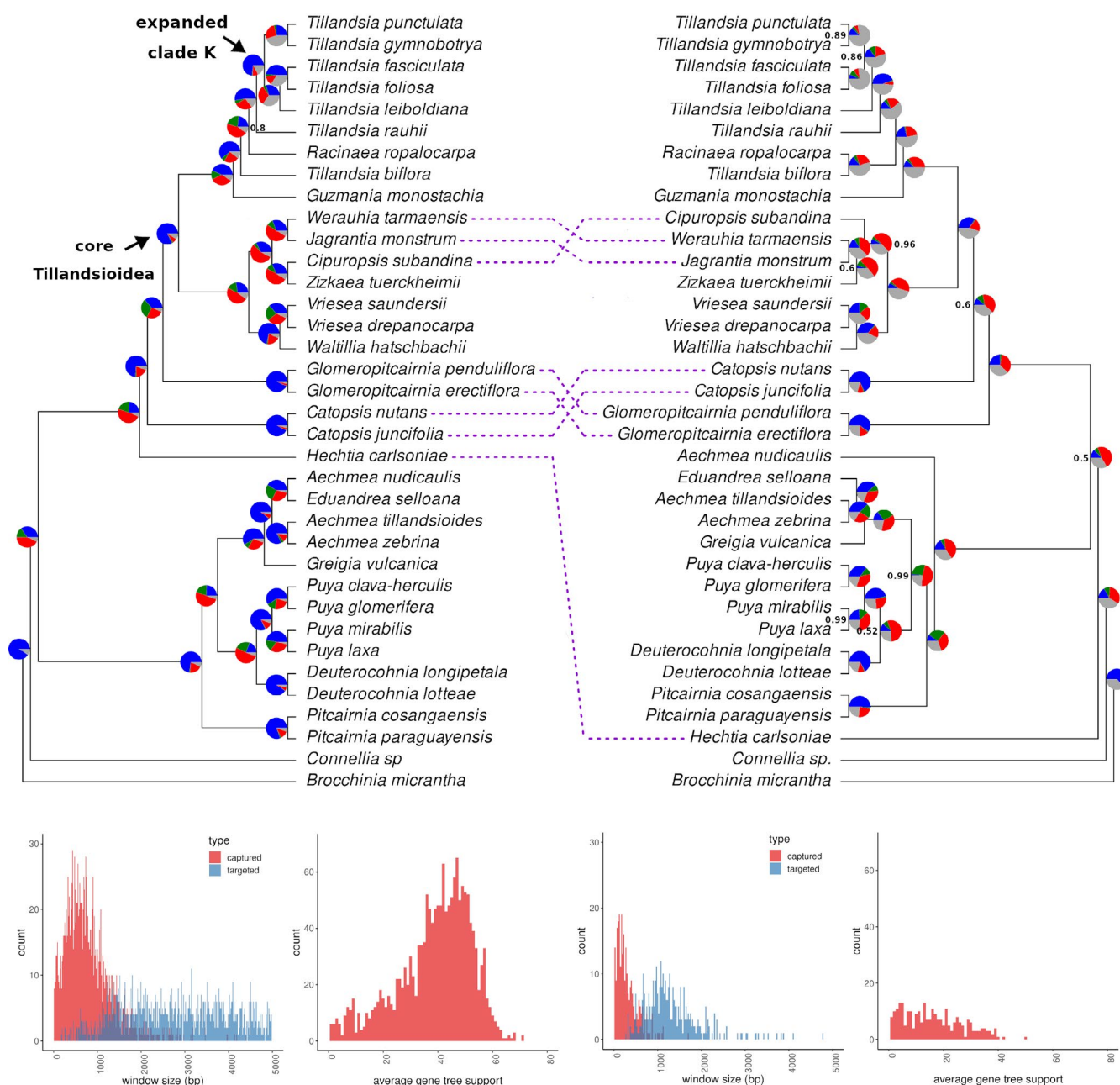


FIGURE 2 Coalescent-based species trees generated ASTRAL-III for samples enriched with Bromeliad1776 (left) and Angiosperms353 (right, flipped for mirroring), on 269 and 1600 genes for each set, respectively. Node values represent local posterior probabilities (pp) for the main topology and are equal to 1 unless noted otherwise. Pie charts at the nodes show levels of gene tree discordance: the percentages of concordant gene trees (blue), the top alternative bipartition (green), other conflicting topologies (red) and uninformative gene trees (grey). At bottom, length and average bootstrap support for gene trees from either data set, according to the design of the bait set used for enrichment: Angiosperms353 (right) and Bromeliad1776 (left). Each gene was considered a single genomic window

expected species boundaries. The main difference between the two data sets was the ability of the Bromeliad1776 set to reach a more consistent solution ('consensus') among 30 runs, especially at large K , as compared to the runs based on the Angiosperms353 bait set. The Bromeliad1776 was also able to distinguish between different sampling localities of *T. punctulata* and of *T. fasciculata* at $K = 7-8$ (Figure 4).

3.6 | Distinct diversities hint at different demographic processes

Nucleotide diversity estimates were calculated for the Bromeliad1776 data set only, due to difficulties obtaining a reliable SNP set with Angiosperms353 (see Section 2.4). Averaged levels of nucleotide diversity at synonymous sites π_s greatly varied

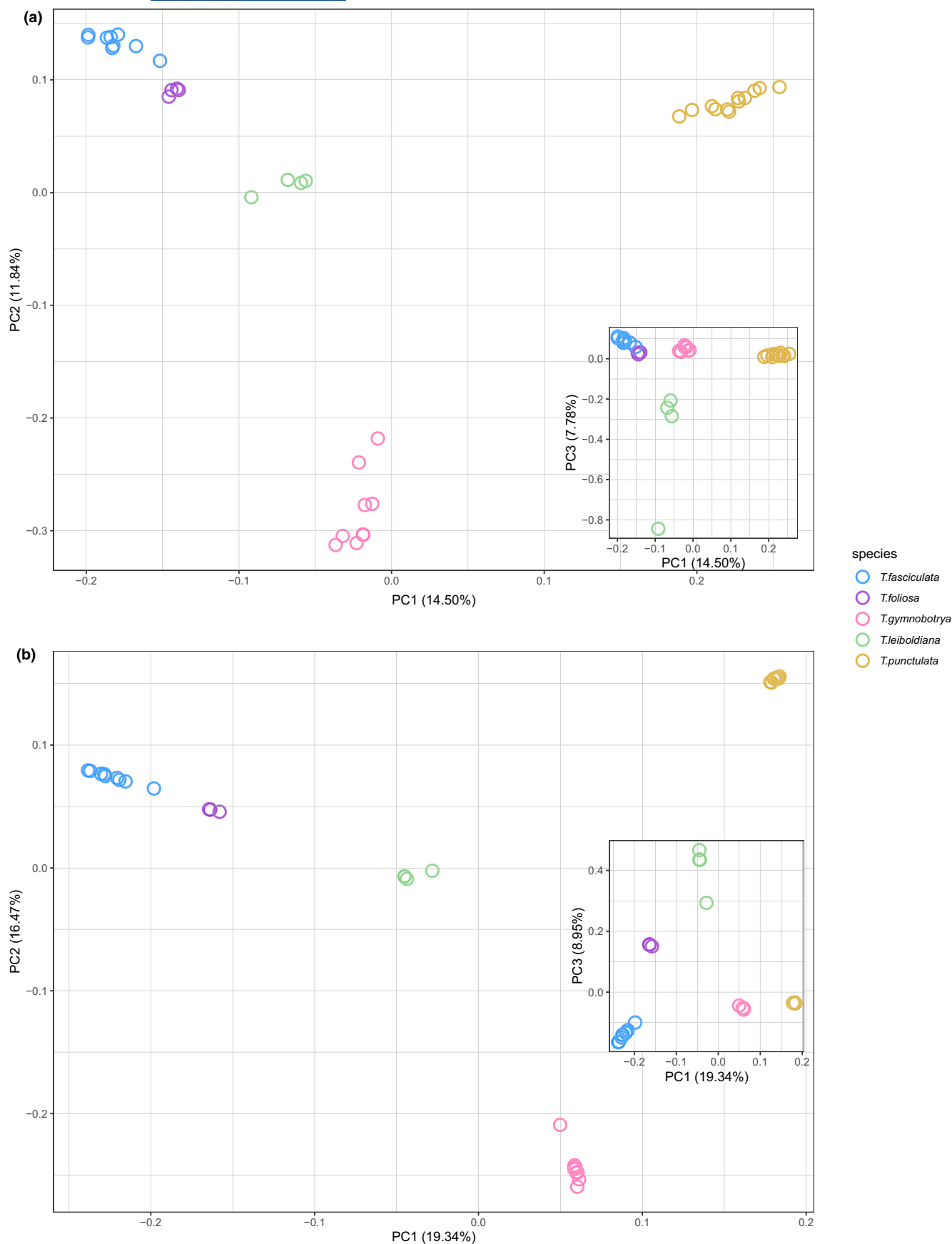


FIGURE 3 Principal component analysis (PCA) plot for samples of *Tillandsia* subgenus *Tillandsia* enriched with two bait sets: (a) Angiosperms353 (1025 variants); (b) Bromeliad1776 (32,941 variants). Colours indicate different species according to legend

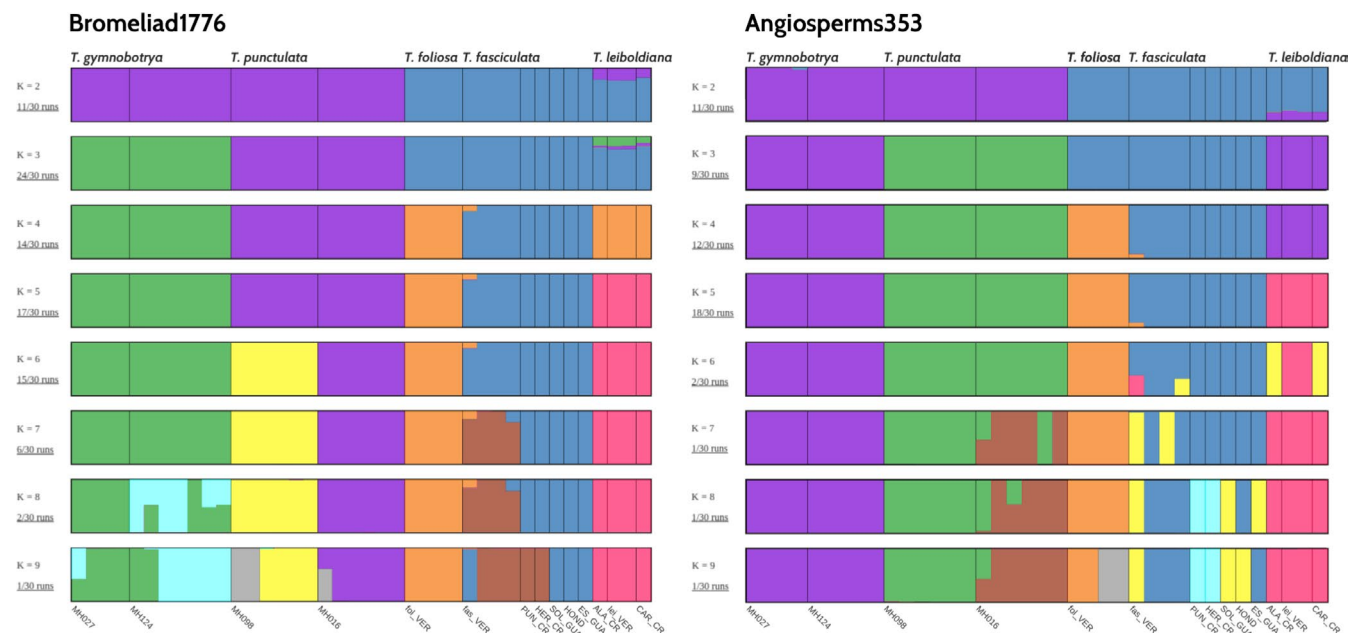


FIGURE 4 Population structure of 5 *Tillandsia* subgenus *Tillandsia* species from 14 sampling locations inferred with the ADMIXTURE software. Samples were enriched with either of two bait sets: Angiosperms353 (9804 variants after LD-pruning) and Bromeliad1776 (42,613 variants after LD-pruning), showing values of $K = 2$ to $K = 9$. Colours represent genetically differentiated groups, while each accession is represented by a vertical bar

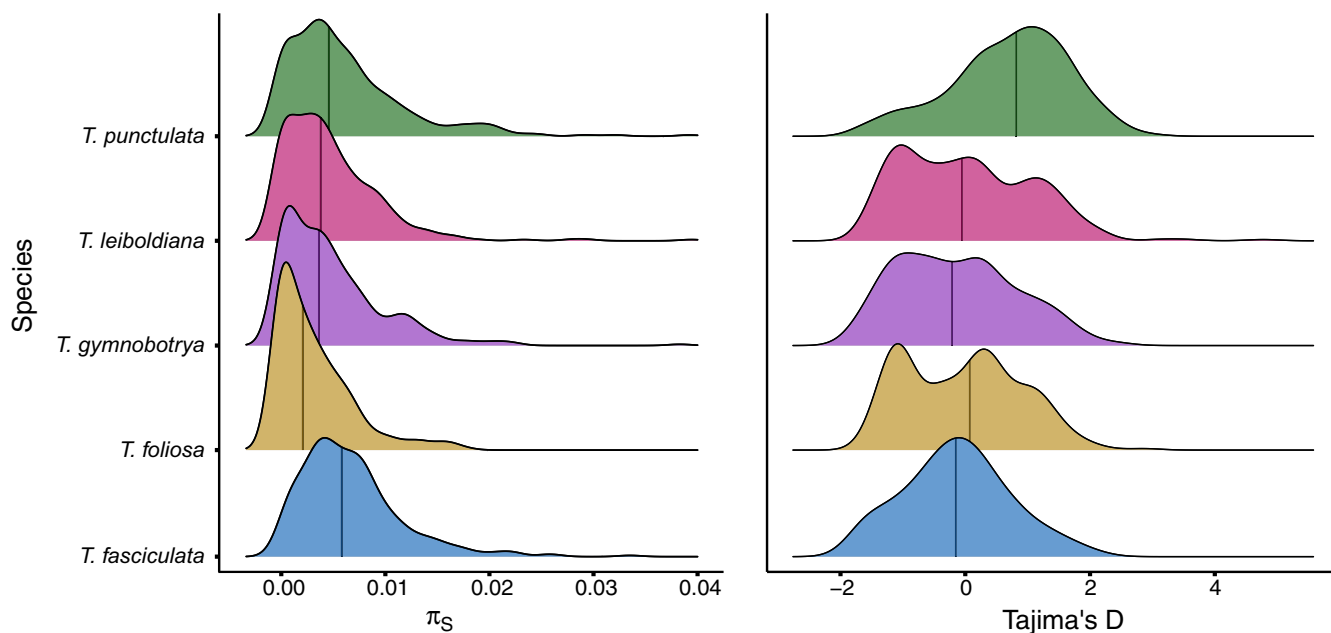


FIGURE 5 Distribution of Tajima's D and synonymous (π_s) nucleotide diversity within each species for the Bromeliad1776 kit

among species, from 4.1×10^{-3} to 8.1×10^{-3} for *T. foliosa* and *T. fasciculata*, respectively (Table S5; Figure 5). Given the recent divergence of these different species and their roughly similar life history traits, they are expected to share relatively similar mutation rates; hence, the observed differences in π_s are expected to translate into differences of long-term N_e . Looking at the distribution of π_s across genes, we found broader or narrower distributions depending on the species, which explains the observed

differences in averaged π_s , as typically represented by the median of the distribution (vertical bars, Figure 5). Most species exhibit distributions of Tajima's D (Figure 5) that are centred around zero, with the notable exception of *T. punctulata*. The distribution of this species is shifted towards positive Tajima's D values, therefore indicating a recent population contraction, suggesting that this species experienced a unique demographic trajectory as compared to the other species.

4 | DISCUSSION

4.1 | A taxon-specific bait set performs marginally better for phylogenomics

In this study, we compared the information content and performance of a taxon-specific bait set and a universal bait set for addressing questions on evolutionary processes at different scales in a highly diverse Neotropical plant group, including recently radiated clades. We found that the taxon-specific kit provided a greater number of segregating sites, yet contrary to our expectations, the abundance of information content did not directly translate to a greater resolution power.

The universal and taxon-specific sets performed comparably when investigating macroevolutionary patterns: the inferred species trees are remarkably consistent between the two bait sets (Figures 2 and S5). Notably, both sets were sufficiently informative to reconstruct the relationships among the fastest radiating clades. These results resonate with previous comparative works (e.g. in *Burmeistera*, Bagley et al., 2020; in *Buddleja*, Chau et al., 2018; and in *Cyperus*, Larridon et al., 2020), where taxon-specific markers provided higher gene assembly success, but a comparable number of segregating sites for phylogenetic inference, indicating that universal bait sets are nearly as effective as taxon-specific bait sets, even in fast evolving taxa. The main advantage of the bromeliad taxon-specific set is its ability to provide additional resolution for deeper examination of gene tree incongruence (Figure 2), currently a fundamental tool in phylogenomic research (Edwards, 2009; Morales-Briones et al., 2021; Pease et al., 2016).

The taxon-specific bait set performed marginally better to address hypotheses at more recent evolutionary scales and provided arguably clearer evidence for inference of species genomic structure using clustering methods. In fact, genetic markers obtained from both data sets provided sufficient information to infer species but no geographic structure, suggesting that *Tillandsia* could be characterized by high gene dispersal among populations. Considering that the Angiosperms353 kit has shown potential to provide within-species signal, as recently demonstrated by Beck et al. (2021) on *Solidago ulmifolia*, and to estimate demographic parameters from herbarium specimen (Simp et al., 2021), we would expect the taxon-specific set to accurately reveal a geographical genetic structure. However, the present study is generally based on small sample sizes per species ($n = 4-8$), mostly sampled within a limited geographic range, limiting our ability to draw robust conclusions on the levels of intraspecific population structure.

The Bromeliad1776 kit provided a substantially larger number of segregating sites (more than 200 k vs. 47 k in Angiosperms353; Table 1, Figure S2) due to higher enrichment success, following the expectation for higher sequence variation in custom-made loci (Figure 1, see also Bragg et al., 2016; de La Harpe et al., 2019; Kadlec et al., 2017). We accordingly found that rates of molecular divergence are distinctly correlated with enrichment success in our sampling (Figure 1), following the expectation that a universal kit will provide fewer segregating sites.

However, the difference in resolution power between the kits cannot be ascribed solely to the different numbers of SNPs, but rather to the length and variability of the obtained regions. The topology obtained with the Angiosperms353 data set under the multispecies coalescent model was substantially different from all other inferred trees and the input gene trees provided a low power to detect patterns of gene tree discordance (Figure 2). We additionally observed that the highly conserved regions targeted by Angiosperms353 are shorter in comparison to Bromeliad1776 targets and thus result in shorter input windows for species tree inference (Figure 2). Hence, the patterns of gene tree discordance in the Angiosperms353 data set likely indicate incorrect gene tree estimation or other model misspecifications, rather than a biological signal. Specifically, coalescence-based methods are sensitive to gene tree estimation error (Zhang et al., 2018) and perform better with gene trees estimated from unlinked loci long enough and variable enough to render sufficient signal per gene tree—this is especially true for data sets with many taxa. The high rates of uninformative genes trees, found in almost half of the intergenic nodes in the Angiosperms353 data set, are expected with increased levels of gene tree error, which in turn reduce the accuracy of ASTRAL (Mirarab, 2019; Sayyari & Mirarab, 2016). In contrast, the Bromeliad1776 ASTRAL tree (Figure 2, left and Figure S9) resolved phylogenetic relationships among taxa with high posterior probability and a topology similar to the ML tree. Gene tree discordance analysis revealed high incongruence around certain nodes, possibly reflecting rapid speciation events.

Since inference of phylogenetic relationships under the multispecies coalescent and exploration of gene tree discordance are both pivotal to phylogenomic research (Degnan & Rosenberg, 2009; Edwards et al., 2016; Pease et al., 2016), a taxon-specific kit provides a clear advantage especially in recent rapid radiations, where gene tree conflict and incomplete lineage sorting are expected to be prevalent (Dornburg et al., 2019; Kubatko & Degnan, 2007; Roch & Warnow, 2015). In that regard, inference of the species tree with the Bromeliad1776 is a tool to drive further hypotheses concerning evolutionary and demographic processes in the evolution of *Tillandsia*. Moreover, the features of the loci targeted provide an important opportunity to study selection (see Section 4.3).

4.2 | Insights on Bromeliaceae phylogeny and demographic processes in *Tillandsia*

Both bait sets resolved the phylogeny of Bromeliaceae, including the fastest evolving lineages of the subfamily Tillandsioideae. The results generally agreed with previous findings of the relationships among taxa (Givnish et al., 2011, 2014). Several findings that contrast with the expected known phylogeny may point at a complexity of genomic processes in the evolutionary history of Bromeliaceae subfamilies. Both the ML tree and species tree did not support a monophyly of the subfamily Pitcairnioideae, which was represented by four samples and two genera in our phylogeny: *Deuterochonia* and *Pitcarnia*. Rather, the genus *Deuterochonia* was sister to

subfamily Puyoideae or sister to both Puyoideae and Bromelioideae subfamilies, inconsistent with the results of Barfuss et al. (2016) and Granados Mendoza et al. (2017). Interestingly, in a visualization of gene tree discordance we found high levels of incongruence and a high percentage of trees supporting an alternative topology in the node splitting the genera, indicating that several genomic processes such as hybridization and incomplete lineage sorting may have accompanied divergence in this group, contributing to the phylogenetic conflict and extending the challenges in resolving these evolutionary relationships. Within the core Tillandsioideae, the tribes Tillandsieae and Vrieseae were found to be monophyletic, in accordance with previous work on the subfamily (Barfuss et al., 2016). Finally, within our focal group *Tillandsia* subgenus *Tillandsia*, clade K as suggested by Barfuss et al. (2016) and clades K.1 and K.2 as proposed by Granados Mendoza et al. (2017) were all well supported, further in agreement with their interpretation of Mexico and Central America as a centre of diversity for subgenus *Tillandsia*. Within *Tillandsia*, incongruence was prominent at the recent splits within clade K.1. and clade K.2 as expected in a recent rapid radiation, a result of high levels of incomplete lineage sorting, hybridization and introgression (Berner & Salzburger, 2015).

When applied to methods in population genetics, we obtained some evidence for a difference in demographic processes and in the level of genetic variation among species. This was especially true for the taxon-specific bait set: for example, the bait set differentiated between populations of *T. punctulata* and *T. fasciculata*, but not *T. gymnotrya* in a maximum-likelihood tree and ancestry analysis (Figures 4 and S5), indicating differences in interpopulation genetic structure among species. The evidence for different demographic processes in these species extended to estimates of Tajima's *D*, where lower values may indicate a recent bottleneck. In addition, we found a unique distribution of nucleotide diversity for *T. foliosa*, possibly reflecting a low effective population size for this endemic species in contrast to the closely related, but widespread *T. fasciculata*. In all cases, our limited sampling given the large size of the family constrains our ability to draw conclusions of a 'true' phylogeny and to account for population structure. Our finding however suggests that nuclear markers obtained with a target capture technique can highlight genomic processes and be further applied to address questions in population genomics with a wider sampling scheme.

4.3 | Future prospects and implications for research in Bromeliaceae and rapid radiations

Beyond the scope of this study, the availability of a bait set kit for Bromeliaceae provides a prime genetic resource for investigating several topical research questions on the origin and maintenance of Bromeliaceae diversity. Manifold studies of bromeliad phylogenomics set forth the challenges of resolving species-level phylogenies with a small number of markers, particularly in young and speciose groups (Goetze et al., 2017; Granados Mendoza et al., 2017; Loiseau

et al., 2021; Versieux et al., 2012). This particularly curated bait set allows highly efficient sequencing across taxa: within our study, we found high mapping success with 82.3% average read mapping. As expected, we documented a difference in enrichment success among taxa, explained by divergence time to the reference used for bait design (see Figure S4), suggesting possible deviations from the assumptions of nonrandomly distributed missing data that may mislead phylogenetic inference (Lemmon et al., 2009; Streicher et al., 2016; Xi et al., 2016). However, given the large enrichment success, downstream analysis with deliberate methodology can account for possible biases and provide robust inference with strict data filtering (Molloy & Warnow, 2018; Streicher et al., 2016). Hence, target enrichment with Bromeliad1776 can produce large data sets with consistent representation between taxa, allowing repeatability between studies and retaining the possibility for global synthesis by including sequence baits orthologous to the universal Angiosperms353 bait set. Moreover, with specific knowledge of the loci targeted in this set, the ability to obtain the same sequences across taxa and experiments and to differentiate genic regions with the use of *A. comosus* models, this bait set offers a broad utility for research in population genomics.

Another important feature in the Bromeliad1776 set is the inclusion of genes putatively associated with key innovative traits in Bromeliaceae with a focus on C3/CAM shifts. Little is known about the molecular basis of the CAM pathway, an adaptation to arid environments which evolved independently and repeatedly in over 36 plant families (Chen et al., 2020; Heyduk et al., 2019; Silvera et al., 2010). CAM phenotypes are considered key adaptations in Bromeliaceae, associated with expansion into novel ecological niches. In *Tillandsia*, C3/CAM shifts were found to be particularly associated with increased rates of diversification (Crayn et al., 2004; Givnish et al., 2014; de La Harpe et al., 2020). The Bromeliad1776 bait set offers opportunities to address specific questions on the relationship between rapid diversification and photosynthetic syndromes in this clade, including testing for gene sequence evolution. Additionally, the inclusion of multicopy genes, combined with newly developed pipelines for studying gene duplication and ploidy (Morales-Briones et al., 2021; Viruel et al., 2019), is beneficial for studying the role of gene duplication and loss in driving diversification. With the increasing ubiquity of target baits as a genomic tool, we expect to see additional pipelines and applications emerging, further expanding the utility of target capture for both macro- and microevolutionary research.

5 | CONCLUSIONS

Even as whole-genome sequencing becomes increasingly economically feasible, target capture is expected to remain popular due to its extensive applications in research. We found that evaluating the differences in resolution power between universal and taxon-specific bait sets is far from a trivial task, and we attempted to lay out a methodological roadmap for researchers wishing to reconstruct the

complex evolutionary history of rapidly diversifying lineages. While a taxon-specific set offers exciting opportunities beyond phylogenomic and into research of molecular evolution, its development is highly time-consuming, requires community-based knowledge and may cost months of work when compared with out-of-the-box universal kits. Our results suggest that universal kits can continue to be employed when aiming to reconstruct phylogenies, in particular as this may offer the possibility to use previously published data to generate larger data sets. However, for those wishing to deeply investigate evolutionary questions in certain lineages, a taxon-specific kit offers certain benefits during data processing stages, where knowledge of the design scheme and gene models is extremely useful, and the possible return of costs is especially high for taxa emerging as model groups. We furthermore encourage groups designing taxon-specific kits to include also universal probes, furthering the mission to complete the tree of life.

ACKNOWLEDGEMENTS

This paper is dedicated to Christian Lexer, a wonderful mentor, friend and colleague. The research was supported with funding from the Christian Lexer professorship start-up BE772002. The analyses benefited from the Vienna Scientific Cluster (VSC) and the Montpellier Bioinformatics Biodiversity (MBB) platform services. We thank Huiying Shang and Aram Drevekenin for help with code, Kelly Swarts, Claus Vogl and Matt Johnson for insightful discussions and advice. We thank the members of Swiss SNSF Sinergia project CRSII3_147630 for accession sampling and three anonymous reviewers for their comments on an earlier version of this manuscript.

AUTHOR CONTRIBUTIONS

CL, MP and GY conceived the study. CL provided funding. TK coordinated sample collection, MdLH, VGJ and GY collected data. MHJB and WT identified species. GY designed bait kit, with guidance from JH and MP. CGC, JV, NR, MHJB and GY performed molecular work. GY and TL analysed the data using feedback from JV and OP. GY wrote the manuscript with significant input from all co-authors.

OPEN RESEARCH BADGES



This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results.

DATA AVAILABILITY STATEMENT

Targeted sequencing reads generated for this project are available at NCBI-SRA under BioProject PRJNA759878; for accession numbers, see Table S4. The probe set and the relevant supporting information are available in Dryad (<https://doi.org/10.5061/dryad.mpg4f4r11>). The bioinformatics scripts are available at https://github.com/giyan/Bromeliad1776/tree/main/MS_2021_scripts.

ORCID

Gil Yardeni <https://orcid.org/0000-0002-6848-7156>

Margot Paris <https://orcid.org/0000-0001-7328-3820>

Michael H. J. Barfuss <https://orcid.org/0000-0001-7172-9454>

Christian Lexer <https://orcid.org/0000-0002-7221-7482>

Ovidiu Paun <https://orcid.org/0000-0002-8295-4937>

REFERENCES

- Aguirre-Santoro, J., Salinas, N. R., & Michelangeli, F. A. (2020). The influence of floral variation and geographic disjunction on the evolutionary dynamics of *Ronnbergia* and *Wittmackia* (Bromeliaceae: Bromelioideae). *Botanical Journal of the Linnean Society*, 192(4), 609–624. <https://doi.org/10.1093/botlinnean/boz087>
- Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, 12(1), 246. <https://doi.org/10.1186/1471-2105-12-246>
- Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Babraham Bioinformatics, Babraham Institute.
- Bagley, J. C., Uribe-Convers, S., Carlsen, M. M., & Muchhala, N. (2020). Utility of targeted sequence capture for phylogenomics in rapid, recent angiosperm radiations: Neotropical *Burmeistera* bellflowers as a case study. *Molecular Phylogenetics and Evolution*, 152, 106769. <https://doi.org/10.1016/j.ympev.2020.106769>
- Barfuss, M. H., Till, W., Leme, E. M., Pinzón, J. P., Manzanares, J. M., Halbritter, H., Samuel, R., & Brown, G. K. (2016). Taxonomic revision of Bromeliaceae subfam. Tillandsioideae based on a multi-locus DNA sequence phylogeny and morphology. *Phytotaxa*, 279(1), 1–97. <https://doi.org/10.11646/phytotaxa.279.1.1>
- Barnett, D. W., Garrison, E. K., Quinlan, A. R., Strömberg, M. P., & Marth, G. T. (2011). BamTools: A C++ API and toolkit for analyzing and managing BAMfiles. *Bioinformatics*, 27(12), 1691–1692. <https://doi.org/10.1093/bioinformatics/btr174>
- Beck, J. B., Markley, M. L., Zielke, M. G., Thomas, J. R., Hale, H. J., Williams, L. D., & Johnson, M. G. (2021). Is Palmer's elm leaf goldenrod real? The Angiosperms353 kit provides within-species signal in *Solidago ulmifolia* s.l. *bioRxiv*. <https://doi.org/10.1101/2021.01.07.425781>
- Behr, A. A., Liu, K. Z., Liu-Fang, G., Nakka, P., & Ramachandran, S. (2016). Pong: Fast analysis and visualization of latent clusters in population genetic data. *Bioinformatics*, 32(18), 2817–2823. <https://doi.org/10.1093/bioinformatics/btw327>
- Benzing, D. H. (2000). *Bromeliaceae: Profile of an adaptive radiation*. Cambridge University Press.
- Berner, D., & Salzburger, W. (2015). The genomics of organismal diversification illuminated by adaptive radiations. *Trends in Genetics*, 31(9), 491–499. <https://doi.org/10.1016/j.tig.2015.07.002>
- Blaimer, B. B., Lloyd, M. W., Guillory, W. X., & Brady, S. G. (2016). Sequence capture and phylogenetic utility of genomic ultraconserved elements obtained from pinned insect specimens. *PLoS One*, 11(8), e0161531. <https://doi.org/10.1371/journal.pone.0161531>
- Bossert, S., & Danforth, B. N. (2018). On the universality of target-enrichment baits for phylogenomic research. *Methods in Ecology and Evolution*, 9(6), 1453–1460. <https://doi.org/10.1111/2041-210X.12988>
- Bossert, S., Murray, E. A., Almeida, E. A. B., Brady, S. G., Blaimer, B. B., & Danforth, B. N. (2019). Combining transcriptomes and ultraconserved elements to illuminate the phylogeny of Apidae. *Molecular Phylogenetics and Evolution*, 130, 121–131. <https://doi.org/10.1016/j.ympev.2018.10.012>
- Bragg, J. G., Potter, S., Bi, K., & Moritz, C. (2016). Exon capture phylogenomics: Efficacy across scales of divergence. *Molecular Ecology Resources*, 16(5), 1059–1068. <https://doi.org/10.1111/1755-0998.12449>

- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Giga-Science*, 4(1), 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Chau, J. H., Rahfeldt, W. A., & Olmstead, R. G. (2018). Comparison of taxon-specific versus general locus sets for targeted sequence capture in plant phylogenomics. *Applications in Plant Sciences*, 6(3), e1032. <https://doi.org/10.1002/aps3.1032>
- Chen, L.-Y., Xin, Y., Wai, C. M., Liu, J., & Ming, R. (2020). The role of cis-elements in the evolution of crassulacean acid metabolism photosynthesis. *Horticulture Research*, 7(1), 1–8. <https://doi.org/10.1038/s41438-019-0229-0>
- Choquet, M., Smolina, I., Dhanasiri, A. K. S., Blanco-Bercial, L., Kopp, M., Jueterbock, A., Sundaram, A. Y. M., & Hoarau, G. (2019). Towards population genomics in non-model species with large genomes: A case study of the marine zooplankton *Calanus finmarchicus*. *Royal Society Open Science*, 6(2), 180608. <https://doi.org/10.1098/rsos.180608>
- Christmas, M. J., Biffin, E., Breed, M. F., & Lowe, A. J. (2017). Targeted capture to assess neutral genomic variation in the narrow-leaf hopbush across a continental biodiversity refugium. *Scientific Reports*, 7(1), 41367. <https://doi.org/10.1038/srep41367>
- Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*, 6(2), 80–92. <https://doi.org/10.4161/fly.19695>
- Crayn, D. M., Winter, K., & Smith, J. A. C. (2004). Multiple origins of crassulacean acid metabolism and the epiphytic habit in the Neotropical family Bromeliaceae. *Proceedings of the National Academy of Sciences of the United States of America*, 101(10), 3703–3708. <https://doi.org/10.1073/pnas.0400366101>
- Dasmahapatra, K. K., Walters, J. R., Briscoe, A. D., Davey, J. W., Whibley, A., Nadeau, N. J., ... The Heliconius Genome Consortium (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, 487(7405), 94–98. <https://doi.org/10.1038/nature11041>
- de La Harpe, M., Hess, J., Loiseau, O., Salamin, N., Lexer, C., & Paris, M. (2019). A dedicated target capture approach reveals variable genetic markers across micro- and macroevolutionary time scales in palms. *Molecular Ecology Resources*, 19(1), 221–234. <https://doi.org/10.1111/1755-0998.12945>
- de La Harpe, M., Paris, M., Hess, J., Barfuss, M. H. J., Serrano-Serrano, M. L., Ghatak, A., Chaturvedi, P., Weckwerth, W., Till, W., Salamin, N., Wai, C. M., Ming, R., & Lexer, C. (2020). Genomic footprints of repeated evolution of CAM photosynthesis in a Neotropical species radiation. *Plant, Cell & Environment*, 43(12), 2987–3001. <https://doi.org/10.1111/pce.13847>
- de La Harpe, M., Paris, M., Karger, D. N., Rolland, J., Kessler, M., Salamin, N., & Lexer, C. (2017). Molecular ecology studies of species radiations: Current research gaps, opportunities and challenges. *Molecular Ecology*, 26(10), 2608–2622. <https://doi.org/10.1111/mec.14110>
- Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, 24(6), 332–340. <https://doi.org/10.1016/j.tree.2009.01.009>
- Derrien, T., & Ramos-Onsins, S. (2020). Assessing a novel sequencing-based approach for population genomics in non-model species. *Peer Community in Genomics*, 1, 100002. <https://doi.org/10.24072/pci.genomics.100002>
- Dornburg, A., Su, Z., & Townsend, J. P. (2019). Optimal rates for phylogenetic inference and experimental design in the era of genome-scale data sets. *Systematic Biology*, 68(1), 145–156. <https://doi.org/10.1093/sysbio/syy047>
- Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19(1), 11–15.
- Edwards, S. V. (2009). Is a new and general theory of molecular systematics emerging? *Evolution*, 63(1), 1–19. <https://doi.org/10.1111/j.1558-5646.2008.00549.x>
- Edwards, S. V., Xi, Z., Janke, A., Faircloth, B. C., McCormack, J. E., Glenn, T. C., Zhong, B., Wu, S., Lemmon, E. M., Lemmon, A. R., Leaché, A. D., Liu, L., & Davis, C. C. (2016). Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. *Molecular Phylogenetics and Evolution*, 94, 447–462. <https://doi.org/10.1016/j.ympev.2015.10.027>
- Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20(1), 238. <https://doi.org/10.1186/s13059-019-1832-y>
- Escudero, M., Nieto-Feliner, G., Pokorny, L., Spalink, D., & Viruel, J. (2020). Editorial: Phylogenomic approaches to deal with particularly challenging plant lineages. *Frontiers in Plant Science*, 11, 591762. <https://doi.org/10.3389/fpls.2020.591762>
- Folk, R. A., Mandel, J. R., & Freudenstein, J. V. (2015). A protocol for targeted enrichment of intron-containing sequence markers for recent radiations: A phylogenomic example from *Heuchera* (Saxifragaceae). *Applications in Plant Sciences*, 3(8), 1500039. <https://doi.org/10.3732/apps.1500039>
- Galtier, N., & Daubin, V. (2008). Dealing with incongruence in phylogenomic analyses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1512), 4023–4029. <https://doi.org/10.1098/rstb.2008.0144>
- Garrison, E. (2012). Vcflib: A C++ library for parsing and manipulating VCF files. Github. <https://github.com/ekg/vcflib>
- Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv:1207.3907 [q-bio].
- Gavrilets, S., & Losos, J. B. (2009). Adaptive radiation: Contrasting theory with data. *Science*, 323(5915), 732–737. <https://doi.org/10.1126/science.1157966>
- Givnish, T. J., Barfuss, M. H. J., Ee, B. V., Riina, R., Schulte, K., Horres, R., Gonsiska, P. A., Jabaily, R. S., Crayn, D. M., Smith, J. A. C., Winter, K., Brown, G. K., Evans, T. M., Holst, B. K., Luther, H., Till, W., Zizka, G., Berry, P. E., & Sytsma, K. J. (2014). Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Molecular Phylogenetics and Evolution*, 71, 55–78. <https://doi.org/10.1016/j.ympev.2013.10.010>
- Givnish, T. J., Barfuss, M. H. J., Van Ee, B., Riina, R., Schulte, K., Horres, R., Gonsiska, P. A., Jabaily, R. S., Crayn, D. M., Smith, J. A. C., Winter, K., Brown, G. K., Evans, T. M., Holst, B. K., Luther, H., Till, W., Zizka, G., Berry, P. E., & Sytsma, K. J. (2011). Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: Insights from an eight-locus plastid phylogeny. *American Journal of Botany*, 98(5), 872–895. <https://doi.org/10.3732/ajb.1000059>
- Givnish, T. J., Burkhardt, E. L., Happel, R. E., & Weintraub, J. D. (1984). Carnivory in the Bromeliad *Brocchinia reducta*, with a cost/benefit model for the general restriction of carnivorous plants to sunny, moist, nutrient-poor habitats. *The American Naturalist*, 124(4), 479–497. <https://doi.org/10.1086/284289>
- Goetze, M., Zanella, C. M., Palma-Silva, C., Büttow, M. V., & Bered, F. (2017). Incomplete lineage sorting and hybridization in the evolutionary history of closely related, endemic yellow-flowered *Aechmea* species of subgenus *Ortgiesia* (Bromeliaceae). *American Journal of Botany*, 104(7), 1073–1087. <https://doi.org/10.3732/ajb.1700103>
- Goolsby, E. W., Moore, A. J., Hancock, L. P., Vos, J. M. D., & Edwards, E. J. (2018). Molecular evolution of key metabolic genes during transitions to C4 and CAM photosynthesis. *American Journal of Botany*, 105(3), 602–613. <https://doi.org/10.1002/ajb2.1051>
- Granados Mendoza, C., Granados-Aguilar, X., Donadio, S., Salazar, G. A., Flores-Cruz, M., Hagsater, E., Starr, J. R., Ibarra-Manríquez, G., Fragoso-Martínez, I., & Magallón, S. (2017). Geographic structure in two highly diverse lineages of *Tillandsia* (Bromeliaceae). *Botany-Botanique*, 95(7), 641–651. <https://doi.org/10.1139/cjb-2016-0250>

- Hale, H., Gardner, E. M., Viruel, J., Pokorny, L., & Johnson, M. G. (2020). Strategies for reducing per-sample costs in target capture sequencing for phylogenomics and population genomics in plants. *Applications in Plant Sciences*, 8(4), e11337. <https://doi.org/10.1002/aps3.11337>
- Heckenhauer, J., Samuel, R., Ashton, P. S., Abu Salim, K., & Paun, O. (2018). Phylogenomics resolves evolutionary relationships and provides insights into floral evolution in the tribe Shoreeae (Dipterocarpaceae). *Molecular Phylogenetics and Evolution*, 127, 1–13. <https://doi.org/10.1016/j.ympev.2018.05.010>
- Hedin, M., Derkarabetian, S., Ramirez, M. J., Vink, C., & Bond, J. E. (2018). Phylogenomic reclassification of the world's most venomous spiders (Mygalomorphae, Aracinae), with implications for venom evolution. *Scientific Reports*, 8(1), 1636. <https://doi.org/10.1038/s41598-018-19946-2>
- Heyduk, K., Moreno-Villena, J. J., Gilman, I. S., Christin, P.-A., & Edwards, E. J. (2019). The genetics of convergent evolution: Insights from plant photosynthesis. *Nature Reviews Genetics*, 20(8), 485–493. <https://doi.org/10.1038/s41576-019-0107-5>
- Hime, P. M., Lemmon, A. R., Lemmon, E. C. M., Prendini, E., Brown, J. M., Thomson, R. C., Kratovil, J. D., Noonan, B. P., Pyron, R. A., Peloso, P. L. V., Kortyna, M. L., Keogh, J. S., Donnellan, S. C., Mueller, R. L., Raxworthy, C. J., Kunte, K., Ron, S. R., Das, S., Gaitonde, N., ... Weisrock, D. W. (2021). Phylogenomics reveals ancient gene tree discordance in the amphibian tree of life. *Systematic Biology*, 70(1), 49–66. <https://doi.org/10.1093/sysbio/syaa034>
- Hollingsworth, P. M., Li, D.-Z., van der Bank, M., & Twyford, A. D. (2016). Telling plant species apart with DNA: From barcodes to genomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1702), 20150338. <https://doi.org/10.1098/rstb.2015.0338>
- Hudlow, W. R., Krieger, R., Meusel, M., Sehhat, J. C., Timken, M. D., & Buoncristiani, M. R. (2011). The NucleoSpin® DNA Clean-up XS kit for the concentration and purification of genomic DNA extracts: An alternative to microdialysis filtration. *Forensic Science International. Genetics*, 5(3), 226–230. <https://doi.org/10.1016/j.fsigen.2010.03.005>
- Hughes, C. E., Nyffeler, R., & Linder, H. P. (2015). Evolutionary plant radiations: Where, when, why and how? *New Phytologist*, 207(2), 249–253. <https://doi.org/10.1111/nph.13523>
- 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.-S., 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. *Systematic Biology*, 68(4), 594–606. <https://doi.org/10.1093/sysbio/syy086>
- Jones, M. R., & Good, J. M. (2016). Targeted capture in evolutionary and ecological genomics. *Molecular Ecology*, 25(1), 185–202. <https://doi.org/10.1111/mec.13304>
- Junier, T., & Zdobnov, E. M. (2010). The Newick utilities: High-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics (Oxford, England)*, 26(13), 1669–1670. <https://doi.org/10.1093/bioinformatics/btq243>
- Kadlec, M., Bellstedt, D. U., Maitre, N. C. L., & Pirie, M. D. (2017). Targeted NGS for species level phylogenomics: “made to measure” or “one size fits all”? *PeerJ*, 5, e3569. <https://doi.org/10.7717/peerj.3569>
- Kessler, M., Abrahamczyk, S., & Krömer, T. (2020). The role of hummingbirds in the evolution and diversification of Bromeliaceae: Unsupported claims and untested hypotheses. *Botanical Journal of the Linnean Society*, 192(4), 592–608. <https://doi.org/10.1093/botlinnean/boz100>
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAXML NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21), 4453–4455. <https://doi.org/10.1093/bioinformatics/btz305>
- Krueger, F. (2019). *Trim Galore: A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (Reduced Representation Bisulfite-Seq) libraries*. Babraham Institute, Babraham Bioinformatics.
- Kubatko, L. S., & Degnan, J. H. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, 56(1), 17–24. <https://doi.org/10.1080/10635150601146041>
- Lamichhane, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerová, M., Rubin, C.-J., Wang, C., Zamani, N., Grant, B. R., Grant, P. R., Webster, M. T., & Andersson, L. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature*, 518(7539), 371–375. <https://doi.org/10.1038/nature14181>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Larridon, I., Villaverde, T., Zuntini, A. R., Pokorny, L., Brewer, G. E., Epitawalage, N., Fairlie, I., Hahn, M., Kim, J., Maguilla, E., Maurin, O., Xanthos, M., Hipp, A. L., Forest, F., & Baker, W. J. (2020). Tackling rapid radiations with targeted sequencing. *Frontiers in Plant Science*, 10, 1655. <https://doi.org/10.3389/fpls.2019.01655>
- Lemmon, A. R., Brown, J. M., Stanger-Hall, K., & Lemmon, E. M. (2009). The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. *Systematic Biology*, 58(1), 130–145. <https://doi.org/10.1093/sysbio/syp017>
- Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61(5), 727–744. <https://doi.org/10.1093/sysbio/sys049>
- Lemmon, E. M., & Lemmon, A. R. (2013). High-throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, 44(1), 99–121. <https://doi.org/10.1146/annurev-ecolsys-110512-135822>
- Leroy, C., Carrias, J.-F., Céréghino, R., & Corbara, B. (2016). The contribution of microorganisms and metazoans to mineral nutrition in bromeliads. *Journal of Plant Ecology*, 9(3), 241–255. <https://doi.org/10.1093/jpe/rtv052>
- Leroy, T., Rousselle, M., Tilak, M.-K., Caizergues, A. E., Scornavacca, C., Recuerda, M., Fuchs, J., Illera, J. C., De Swardt, D. H., Blanco, G., Thébaud, C., Milá, B., & Nabholz, B. (2021). Island songbirds as windows into evolution in small populations. *Current Biology*, 31(6), 1303–1310.e4. <https://doi.org/10.1016/j.cub.2020.12.040>
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21), 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Linder, H. P. (2008). Plant species radiations: Where, when, why? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1506), 3097–3105. <https://doi.org/10.1098/rstb.2008.0075>
- Liu, Y., Johnson, M. G., Cox, C. J., Medina, R., Devos, N., Vanderpoorten, A., Hedenäs, L., Bell, N. E., Shevock, J. R., Aguero, B., Quandt, D., Wickett, N. J., Shaw, A. J., & Goffinet, B. (2019). Resolution of the ordinal phylogeny of mosses using targeted exons from organellar and nuclear genomes. *Nature Communications*, 10(1), 1485. <https://doi.org/10.1038/s41467-019-09454-w>
- Loiseau, O., Mota Machado, T., Paris, M., Koubinová, D., Dexter, K. G., Versieux, L. M., Lexer, C., & Salamin, N. (2021). Genome skimming reveals widespread hybridization in a Neotropical flowering plant radiation. *Frontiers in Ecology and Evolution*, 9, 322. <https://doi.org/10.3389/fevo.2021.668281>

- Luther, H. E. (2008). *An alphabetical list of bromeliad binomials* (11th ed.). The Marie Selby Botanical Gardens Sarasota, The Bromeliad Society International.
- Machado, T. M., Loiseau, O., Paris, M., Weigand, A., Versieux, L. M., Stehmann, J. R., Lexer, C., & Salamin, N. (2020). Systematics of *Vriesea* (Bromeliaceae): Phylogenetic relationships based on nuclear gene and partial plastome sequences. *Botanical Journal of the Linnean Society*, 192(4), 656–674. <https://doi.org/10.1093/botlinnean/boz102>
- Madden, T. (2003). *The BLAST Sequence Analysis Tool*. National Center for Biotechnology Information (US).
- McDonnell, A. J., Baker, W. J., Dodsworth, S., Forest, F., Graham, S. W., Johnson, M. G., Pokorny, L., Tate, J., Wicke, S., & Wickett, N. J. (2021). Exploring Angiosperms353: Developing and applying a universal toolkit for flowering plant phylogenomics. *Applications in Plant Sciences*, 9(7), 11443. <https://doi.org/10.1002/aps3.11443>
- McGee, M. D., Borstein, S. R., Meier, J. I., Marques, D. A., Mwaiko, S., Taabu, A., Kishu, M. A., O'Meara, B., Bruggmann, R., Excoffier, L., & Seehausen, O. (2020). The ecological and genomic basis of explosive adaptive radiation. *Nature*, 586(7827), 75–79. <https://doi.org/10.1038/s41586-020-2652-7>
- McGlothlin, J. W., Kobiela, M. E., Wright, H. V., Mahler, D. L., Kolbe, J. J., Losos, J. B., & Brodie, E. D. (2018). Adaptive radiation along a deeply conserved genetic line of least resistance in *Anolis* lizards. *Evolution Letters*, 2(4), 310–322. <https://doi.org/10.1002/evl3.72>
- McGowen, M. R., Tsagkogeorga, G., Álvarez-Carretero, S., dos Reis, M., Struebig, M., Deaville, R., Jepson, P. D., Jarman, S., Polanowski, A., Morin, P. A., & Rossiter, S. J. (2020). Phylogenomic resolution of the Cetacean tree of life using target sequence capture. *Systematic Biology*, 69(3), 479–501. <https://doi.org/10.1093/sysbio/syz068>
- McKain, M. R., Johnson, M. G., Uribe-Convers, S., Eaton, D., & Yang, Y. (2018). Practical considerations for plant phylogenomics. *Applications in Plant Sciences*, 6(3), e1038. <https://doi.org/10.1002/aps3.1038>
- Ming, R., VanBuren, R., Wai, C. M., Tang, H., Schatz, M. C., Bowers, J. E., Lyons, E., Wang, M.-L., Chen, J., Biggers, E., Zhang, J., Huang, L., Zhang, L., Miao, W., Zhang, J., Ye, Z., Miao, C., Lin, Z., Wang, H., ... Yu, Q. (2015). The pineapple genome and the evolution of CAM photosynthesis. *Nature Genetics*, 47(12), 1435–1442. <https://doi.org/10.1038/ng.3435>
- Mirarab, S. (2019). *Species tree estimation using ASTRAL: Practical considerations*. arXiv:1904.03826 [q-bio].
- Moest, M., Van Belleghem, S. M., James, J. E., Salazar, C., Martin, S. H., Barker, S. L., Moreira, G. R. P., Mérot, C., Joron, M., Nadeau, N. J., Steiner, F. M., & Jiggins, C. D. (2020). Selective sweeps on novel and introgressed variation shape mimicry loci in a butterfly adaptive radiation. *PLOS Biology*, 18(2), e3000597. <https://doi.org/10.1371/journal.pbio.3000597>
- Molloy, E. K., & Warnow, T. (2018). To include or not to include: The impact of gene filtering on species tree estimation methods. *Systematic Biology*, 67(2), 285–303. <https://doi.org/10.1093/sysbio/syx077>
- Morales-Briones, D. F., Gehrke, B., Huang, C.-H., Liston, A., Ma, H., Marx, H. E., & Tank, D. C., & Yang, Y. (2021). Analysis of paralogs in target enrichment data pinpoints multiple ancient polyploidy events in *Alchemilla* s.l. (Rosaceae). *Systematic Biology*, 1063–1157. <https://doi.org/10.1093/sysbio/syab032>
- Morel, B., Kozlov, A. M., & Stamatakis, A. (2019). ParGenes: A tool for massively parallel model selection and phylogenetic tree inference on thousands of genes. *Bioinformatics*, 35(10), 1771–1773. <https://doi.org/10.1093/bioinformatics/bty839>
- Mota, M. R., Pinheiro, F., Leal, B. S. D. S., Sardelli, C. H., Wendt, T., & Palma-Silva, C. (2020). From micro- to macroevolution: Insights from a Neotropical bromeliad with high population genetic structure adapted to rock outcrops. *Heredity*, 125(5), 353–370. <https://doi.org/10.1038/s41437-020-0342-8>
- Moyle, R. G., Filardi, C. E., Smith, C. E., & Diamond, J. (2009). Explosive Pleistocene diversification and hemispheric expansion of a “great speciator”. *Proceedings of the National Academy of Sciences of the United States of America*, 106(6), 1863–1868. <https://doi.org/10.1073/pnas.0809861105>
- Murphy, B., Forest, F., Barraclough, T., Rosindell, J., Bellot, S., Cowan, R., Golos, M., Jebb, M., & Cheek, M. (2020). A phylogenomic analysis of *Nepenthes* (Nepenthaceae). *Molecular Phylogenetics and Evolution*, 144, 106668. <https://doi.org/10.1016/j.ympev.2019.106668>
- Nevado, B., Atchison, G. W., Hughes, C. E., & Filatov, D. A. (2016). Widespread adaptive evolution during repeated evolutionary radiations in New World lupins. *Nature Communications*, 7(1), 12384. <https://doi.org/10.1038/ncomms12384>
- Nute, M., Chou, J., Molloy, E. K., & Warnow, T. (2018). The performance of coalescent based species tree estimation methods under models of missing data. *BMC Genomics*, 19(5), 286. <https://doi.org/10.1186/s12864-018-4619-8>
- Ogutten, E., Christe, C., Nishii, K., Salamin, N., Möller, M., & Perret, M. (2021). Phylogenomics of Gesneriaceae using targeted capture of nuclear genes. *Molecular Phylogenetics and Evolution*, 157, 107068. <https://doi.org/10.1016/j.ympev.2021.107068>
- Ortiz, E. M. (2019). Vcf2phylo v2.0: Convert a VCF matrix into several matrix formats for phylogenetic analysis. *Zenodo*. <https://doi.org/10.5281/zenodo.2540861>
- Palma-Silva, C., & Fay, M. F. (2020). Bromeliaceae as a model group in understanding the evolution of Neotropical biota. *Botanical Journal of the Linnean Society*, 192(4), 569–586. <https://doi.org/10.1093/botlinnean/boaa003>
- Palma-Silva, C., Ferro, M., Bacci, M., & Turchetto-Zolet, A. C. (2016). De novo assembly and characterization of leaf and floral transcriptomes of the hybridizing bromeliad species (*Pitcairnia* spp.) adapted to Neotropical Inselbergs. *Molecular Ecology Resources*, 16(4), 1012–1022. <https://doi.org/10.1111/1755-0998.12504>
- 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 Biology*, 14(2), e1002379. <https://doi.org/10.1371/journal.pbio.1002379>
- Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: An efficient Swiss army knife for population genomic analyses in R. *Molecular Biology and Evolution*, 31(7), 1929–1936. <https://doi.org/10.1093/molbev/msu136>
- Picard toolkit (2019). Broad Institute, GitHub repository. <http://broadinstitute.github.io/picard/>
- Quattrini, A. M., Faircloth, B. C., Dueñas, L. F., Bridge, T. C. L., Brugler, M. R., Calixto-Botía, I. F., DeLeo, D. M., Forêt, S., Herrera, S., Lee, S. M. Y., Miller, D. J., Prada, C., Rádis-Baptista, G., Ramírez-Portilla, C., Sánchez, J. A., Rodríguez, E., & McFadden, C. S. (2018). Universal target-enrichment baits for anthozoan (Cnidaria) phylogenomics: New approaches to long-standing problems. *Molecular Ecology Resources*, 18(2), 281–295. <https://doi.org/10.1111/1755-0998.12736>
- Quezada, I. M., & Gianoli, E. (2011). Crassulacean acid metabolism photosynthesis in Bromeliaceae: An evolutionary key innovation. *Biological Journal of the Linnean Society*, 104(2), 480–486. <https://doi.org/10.1111/j.1095-8312.2011.01713.x>
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842. <https://doi.org/10.1093/bioinformatics/btq033>
- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Renaud, G., Stenzel, U., Maricic, T., Wiebe, V., & Kelso, J. (2015). deML: Robust demultiplexing of Illumina sequences using a likelihood-based approach. *Bioinformatics (Oxford, England)*, 31(5), 770–772. <https://doi.org/10.1093/bioinformatics/btu719>
- Roch, S., & Warnow, T. (2015). On the robustness to gene tree estimation error (or lack thereof) of coalescent-based species tree methods.

- Systematic Biology*, 64(4), 663–676. <https://doi.org/10.1093/sysbio/syv016>
- Salzburger, W. (2018). Understanding explosive diversification through cichlid fish genomics. *Nature Reviews Genetics*, 19(11), 705–717. <https://doi.org/10.1038/s41576-018-0043-9>
- Sanderson, B. J., DiFazio, S. P., Cronk, Q. C. B., Ma, T., & Olson, M. S. (2020). A targeted sequence capture array for phylogenetics and population genomics in the Salicaceae. *Applications in Plant Sciences*, 8(10), e11394. <https://doi.org/10.1002/aps3.11394>
- Sass, C., Iles, W. J. D., Barrett, C. F., Smith, S. Y., & Specht, C. D. (2016). Revisiting the Zingiberales: Using multiplexed exon capture to resolve ancient and recent phylogenetic splits in a charismatic plant lineage. *PeerJ*, 4, e1584. <https://doi.org/10.7717/peerj.1584>
- Sayyari, E., & Mirarab, S. (2016). Fast coalescent-based computation of local branch support from Quartet Frequencies. *Molecular Biology and Evolution*, 33(7), 1654–1668. <https://doi.org/10.1093/molbev/msw079>
- Schulte, K., Barfuss, M. H. J., & Zizka, G. (2009). Phylogeny of Bromelioideae (Bromeliaceae) inferred from nuclear and plastid DNA loci reveals the evolution of the tank habit within the subfamily. *Molecular Phylogenetics and Evolution*, 51(2), 327–339. <https://doi.org/10.1016/j.ympev.2009.02.003>
- Shah, T., Schneider, J. V., Zizka, G., Maurin, O., Baker, W., Forest, F., Brewer, G. E., Savolainen, V., Darbyshire, I., & Larridon, I. (2021). Joining forces in Ochnaceae phylogenomics: A tale of two targeted sequencing probe kits. *American Journal of Botany*, 108(7), 1201–1216. <https://doi.org/10.1002/ajb2.1682>
- Shee, Z. Q., Frodin, D. G., Cámara-Leret, R., & Pokorny, L. (2020). Reconstructing the complex evolutionary history of the Papuanian Schefflera radiation through Herbariomics. *Frontiers in Plant Science*, 11, 258. <https://doi.org/10.3389/fpls.2020.00258>
- Silvera, K., Neubig, K. M., Whitten, W. M., Williams, N. H., Winter, K., & Cushman, J. C. (2010). Evolution along the crassulacean acid metabolism continuum. *Functional Plant Biology*, 37(11), 995–1010. <https://doi.org/10.1071/FP10084>
- Slimp, M., Williams, L. D., Hale, H., & Johnson, M. G. (2021). On the potential of Angiosperms353 for population genomics. *Applications in Plant Sciences*, 9(7), e11419. <https://doi.org/10.1002/aps3.11419>
- Soltis, P. S., Folk, R. A., & Soltis, D. E. (2019). Darwin review: Angiosperm phylogeny and evolutionary radiations. *Proceedings of the Royal Society B: Biological Sciences*, 286(1899), 20190099. <https://doi.org/10.1098/rspb.2019.0099>
- Soltis, P. S., & Soltis, D. E. (2004). The origin and diversification of angiosperms. *American Journal of Botany*, 91(10), 1614–1626. <https://doi.org/10.3732/ajb.91.10.1614>
- Soto-Gomez, M. S., Pokorny, L., Kantar, M. B., Forest, F., Leitch, I. J., Gravendeel, B., & Viruel, J. (2019). A customized nuclear target enrichment approach for developing a phylogenomic baseline for *Dioscorea* yams (Dioscoreaceae). *Applications in Plant Sciences*, 7(6), e11254. <https://doi.org/10.1002/aps3.11254>
- Straub, S. C. K., Moore, M. J., Soltis, P. S., Soltis, D. E., Liston, A., & Livshultz, T. (2014). Phylogenetic signal detection from an ancient rapid radiation: Effects of noise reduction, long-branch attraction, and model selection in crown clade Apocynaceae. *Molecular Phylogenetics and Evolution*, 80, 169–185. <https://doi.org/10.1016/j.ympev.2014.07.020>
- Straub, S. C. K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R. C., & Liston, A. (2012). Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *American Journal of Botany*, 99(2), 349–364. <https://doi.org/10.3732/ajb.1100335>
- Streicher, J. W., Schulte, J. A., & Wiens, J. J. (2016). How should genes and taxa be sampled for phylogenomic analyses with missing data? An empirical study in Iguanian Lizards. *Systematic Biology*, 65(1), 128–145. <https://doi.org/10.1093/sysbio/syv058>
- Stroud, J. T., & Losos, J. B. (2020). Bridging the process-pattern divide to understand the origins and early stages of adaptive radiation: A review of approaches with insights from studies of *Anolis* lizards. *Journal of Heredity*, 111(1), 33–42. <https://doi.org/10.1093/jhered/esz055>
- Supple, M. A., & Shapiro, B. (2018). Conservation of biodiversity in the genomics era. *Genome Biology*, 19(1), 131. <https://doi.org/10.1186/s13059-018-1520-3>
- Thomas, A. E., Igea, J., Meudt, H. M., Albach, D. C., Lee, W. G., & Tanentzap, A. J. (2021). Using target sequence capture to improve the phylogenetic resolution of a rapid radiation in New Zealand *Veronica*. *American Journal of Botany*, 108(7), 1289–1306. <https://doi.org/10.1002/ajb2.1678>
- Ufimov, R., Zeisek, V., Pišová, S., Baker, W. J., Fér, T., Loo, M., Dobeš, C., & Schmickl, R. (2021). Relative performance of customized and universal probe sets in target enrichment: A case study in subtribe Malinae. *Applications in Plant Sciences*, 9(7), e11442. <https://doi.org/10.1002/aps3.11442>
- Van de Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nature Reviews Genetics*, 18(7), 411–424. <https://doi.org/10.1038/nrg.2017.26>
- Versieux, L. M., Barbará, T., Wanderley, M. D. G. L., Calvente, A., Fay, M. F., & Lexer, C. (2012). Molecular phylogenetics of the Brazilian giant bromeliads (*Alcantarea*, Bromeliaceae): Implications for morphological evolution and biogeography. *Molecular Phylogenetics and Evolution*, 64(1), 177–189. <https://doi.org/10.1016/j.ympev.2012.03.015>
- Villaverde, T., Pokorny, L., Olsson, S., Rincón-Barrado, M., Johnson, M. G., Gardner, E. M., Wickett, N. J., Molero, J., Riina, R., & Sanmartín, I. (2018). Bridging the micro- and macroevolutionary levels in phylogenomics: Hyb-Seq solves relationships from populations to species and above. *New Phytologist*, 220(2), 636–650. <https://doi.org/10.1111/nph.15312>
- Viruel, J., Conejero, M., Hidalgo, O., Pokorny, L., Powell, R. F., Forest, F., Kantar, M. B., Soto Gomez, M., Graham, S. W., Gravendeel, B., Wilkin, P., & Leitch, I. J. (2019). A target capture-based method to estimate ploidy from Herbarium specimens. *Frontiers in Plant Science*, 10, 937. <https://doi.org/10.3389/fpls.2019.00937>
- Wai, C. M., VanBuren, R., Zhang, J., Huang, L., Miao, W., Edger, P. P., & Ming, R. (2017). Temporal and spatial transcriptomic and microRNA dynamics of CAM photosynthesis in pineapple. *The Plant Journal*, 92(1), 19–30. <https://doi.org/10.1111/tj.13630>
- Weitemier, K., Straub, S. C. K., Cronn, R. C., Fishbein, M., Schmickl, R., McDonnell, A., & Liston, A. (2014). Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences*, 2(9), 1400042. <https://doi.org/10.3732/apps.1400042>
- Wendel, J. F., & Doyle, J. J. (1998). Phylogenetic incongruence: Window into genome history and molecular evolution. In D. E. Soltis, P. S. Soltis, & J. J. Doyle (Eds.), *Molecular systematics of plants II: DNA sequencing* (pp. 265–296). Springer US.
- Wöhrmann, T., Michalak, I., Zizka, G., & Weising, K. (2020). Strong genetic differentiation among populations of *Fosterella rusbyi* (Bromeliaceae) in Bolivia. *Botanical Journal of the Linnean Society*, 192(4), 744–759. <https://doi.org/10.1093/botlinnean/boz096>
- Xi, Z., Liu, L., & Davis, C. C. (2016). The impact of missing data on species tree estimation. *Molecular Biology and Evolution*, 33(3), 838–860. <https://doi.org/10.1093/molbev/msv266>
- Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, 19(6), 153. <https://doi.org/10.1186/s12859-018-2129-y>
- Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, 28(24), 3326–3328. <https://doi.org/10.1093/bioinformatics/bts606>
- Zimmer, E. A., & Wen, J. (2013). Reprint of: Using nuclear gene data for plant phylogenetics: Progress and prospects. *Molecular*

Phylogenetics and Evolution, 66(2), 539–550. <https://doi.org/10.1016/j.ympev.2013.01.005>

Zink, R. M., & Vázquez-Miranda, H. (2019). Species limits and phylogenomic relationships of Darwin's Finches remain unresolved: Potential consequences of a volatile ecological setting. *Systematic Biology*, 68(2), 347–357. <https://doi.org/10.1093/sysbio/syy073>

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Yardeni, G., Viruel, J., Paris, M., Hess, J., Groot Crego, C., de La Harpe, M., Rivera, N., Barfuss, M. H. J., Till, W., Guzmán-Jacob, V., Krömer, T., Lexer, C., Paun, O., & Leroy, T. (2022). Taxon-specific or universal? Using target capture to study the evolutionary history of rapid radiations. *Molecular Ecology Resources*, 22, 927–945. <https://doi.org/10.1111/1755-0998.13523>

Supporting information

including Supporting Figures, supporting information and supporting tables

MOLECULAR ECOLOGY

RESOURCES

Supplemental Information for:

Taxon-specific or universal? Using target capture to study the evolutionary history of a rapid radiation

Gil Yardeni, Juan Viruel, Margot Paris, Jaqueline Hess, Clara Groot Crego, Marylaure de La Harpe, Norma Rivera, Michael H. J. Barfuss, Walter Till, Valeria Guzmán-Jacob, Thorsten Krömer, Christian Lexer, Ovidiu Paun and Thibault Leroy

Table of Contents:

Supp. Figure S1	Page 2
Supp. Figure S2	Page 3
Supp. Figure S3	Page 4
Supp. Figure S4	Page 5
Supp. Figure S5	Page 6
Supp. Figure S6	Page 7
Supp. Figure S7	Page 8
Supp. Figure S8	Page 9
Supp. Figure S9	Page 10
Supporting information: ascertainment bias	Page 11

MOLECULAR ECOLOGY RESOURCES



Figure S1 Map of sampling locations for *Tillandsia* subgenus *Tillandsia* accessions within Mexico.

MOLECULAR ECOLOGY RESOURCES

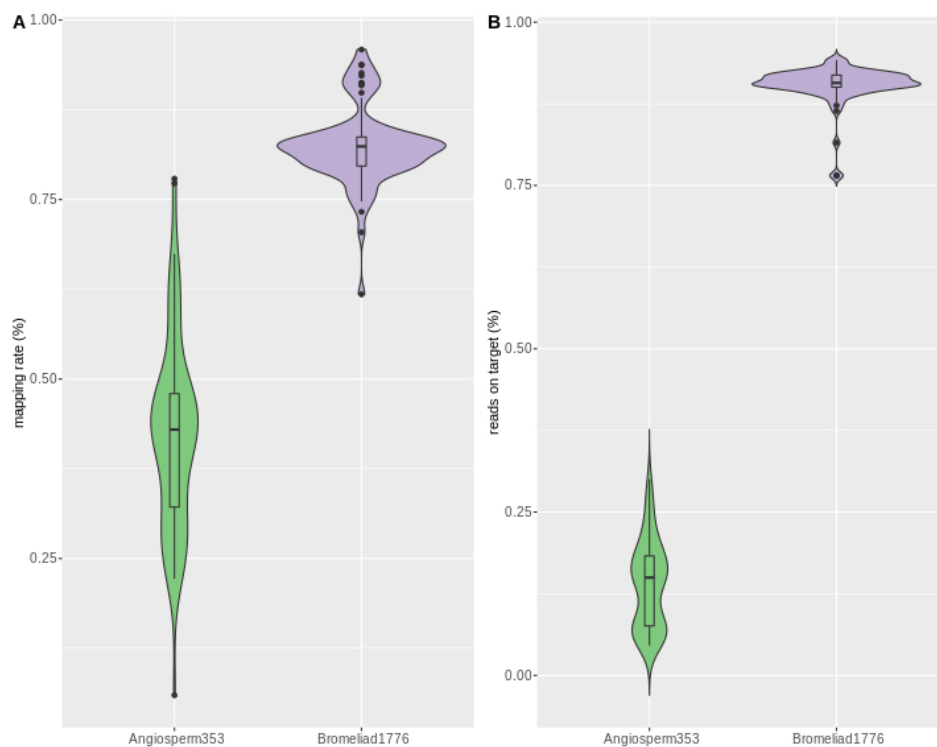


Figure S2 Mapping rates (A) and percentage of reads matching bait sequences (B) for Bromeliad samples enriched with one of two bait sets: Angiosperms353 and Bromeliad1776. Reads were mapped against *Ananas comosus* reference for both bait sets. Targets were defined as bait locations and flanking 500 base-pairs. Bromeliad1776 targets were defined as the regions used for bait design and Angiosperms353 targets were defined as *A. comosus* orthologous regions matching the genes used for bait design.

MOLECULAR ECOLOGY RESOURCES

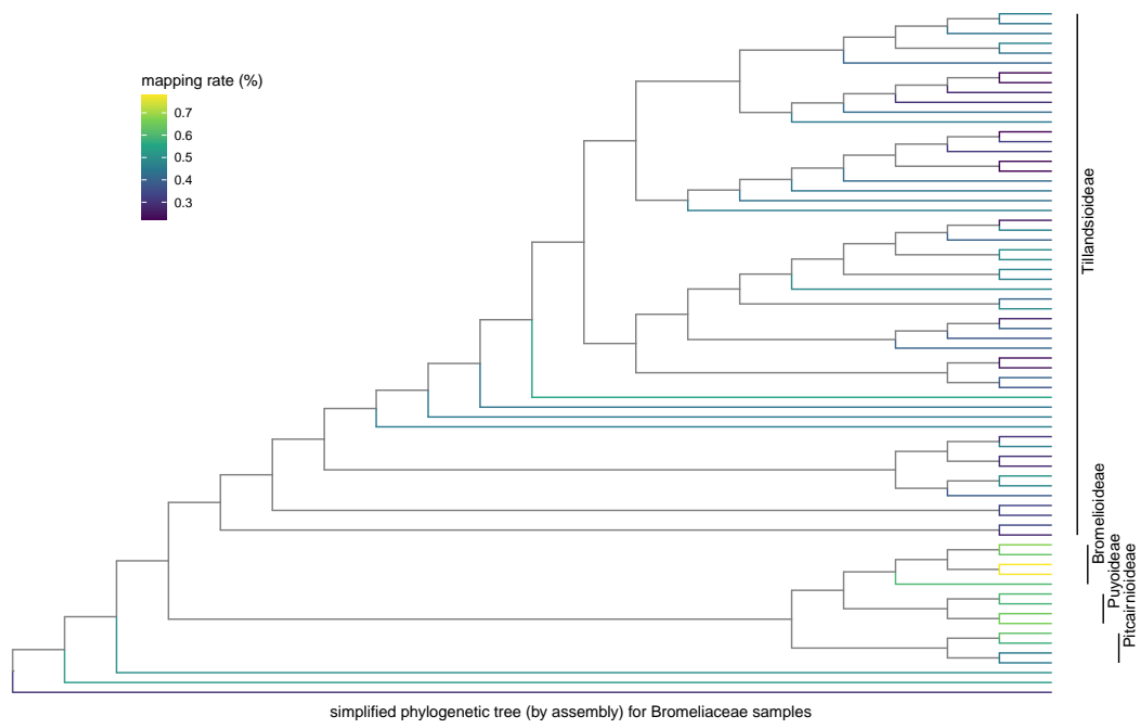


Figure S3 A simplified phylogenetic tree, with branches colored according to read mapping percentage for samples enriched with Angiosperms353.

MOLECULAR ECOLOGY RESOURCES

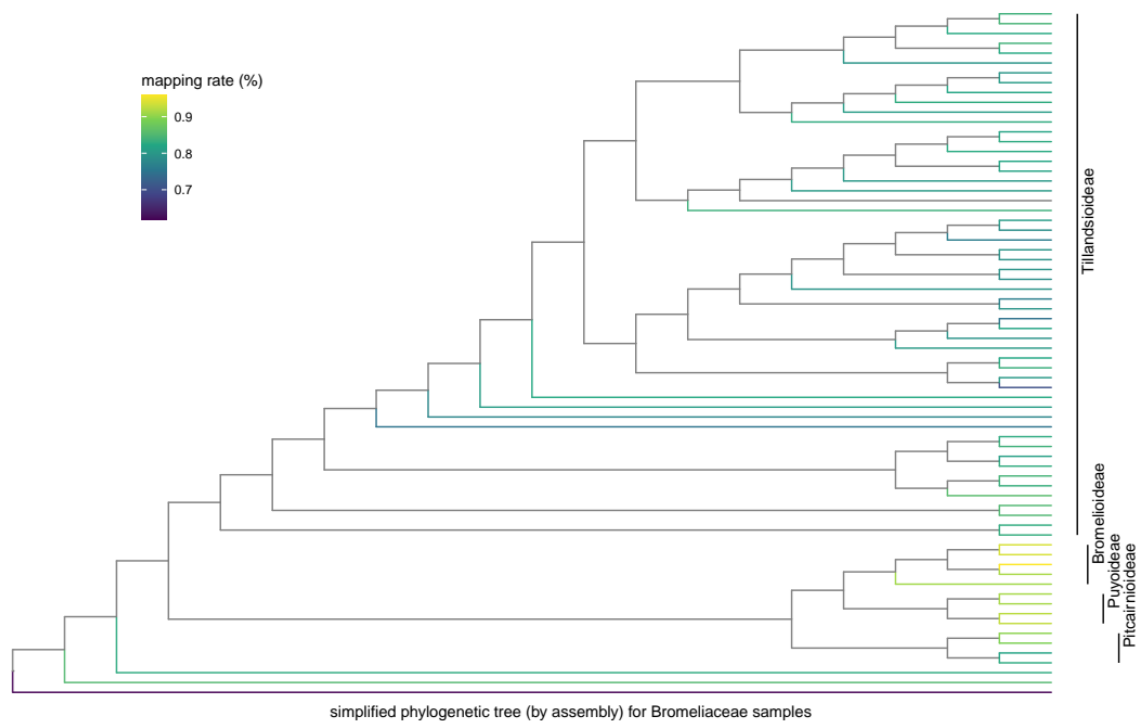


Figure S4 A simplified phylogenetic tree, with branches colored according to read mapping percentage for samples enriched with Bromeliad1776.

MOLECULAR ECOLOGY RESOURCES

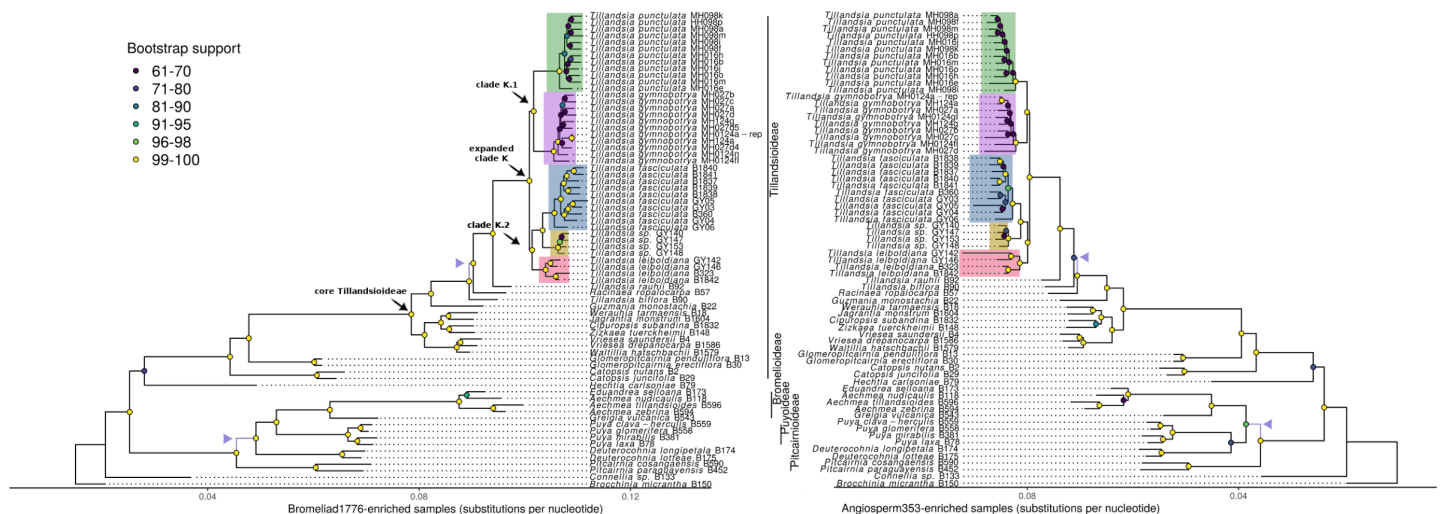


Figure S5 Maximum-likelihood (ML) phylogenetic tree inferred with RAXML-NG, based on variants called for data sets enriched with Bromeliad1776 bait set (left) and Angiosperms353 bait set (right, flipped for mirroring). Branch lengths were calculated by number of substitutions per site. Internal nodes are marked and colored according to bootstrap support. Nodes which differed among trees are colored purple and have been marked by an arrow.

MOLECULAR ECOLOGY RESOURCES

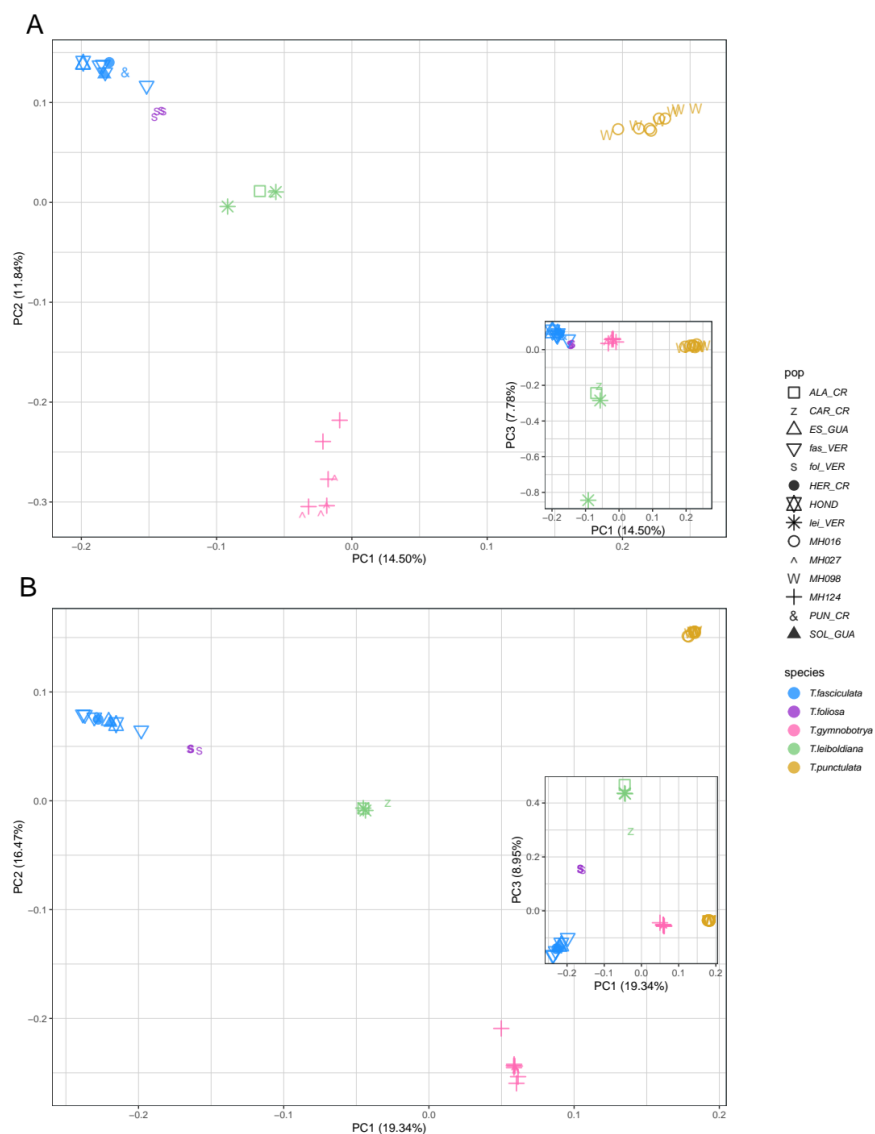


Figure S6 Principal Component Analysis (PCA) plot for samples of *Tillandsia* subgenus *Tillandsia* enriched with two bait sets: A. Angiosperms353 (1,025 variants after LD-pruning) B. Bromeliad1776

MOLECULAR ECOLOGY RESOURCES

(32,941 variants after LD-pruning). Colors indicate different species (following the scheme in Supporting Figure S5) and shapes represent different geographic origins (populations).

MOLECULAR ECOLOGY RESOURCES

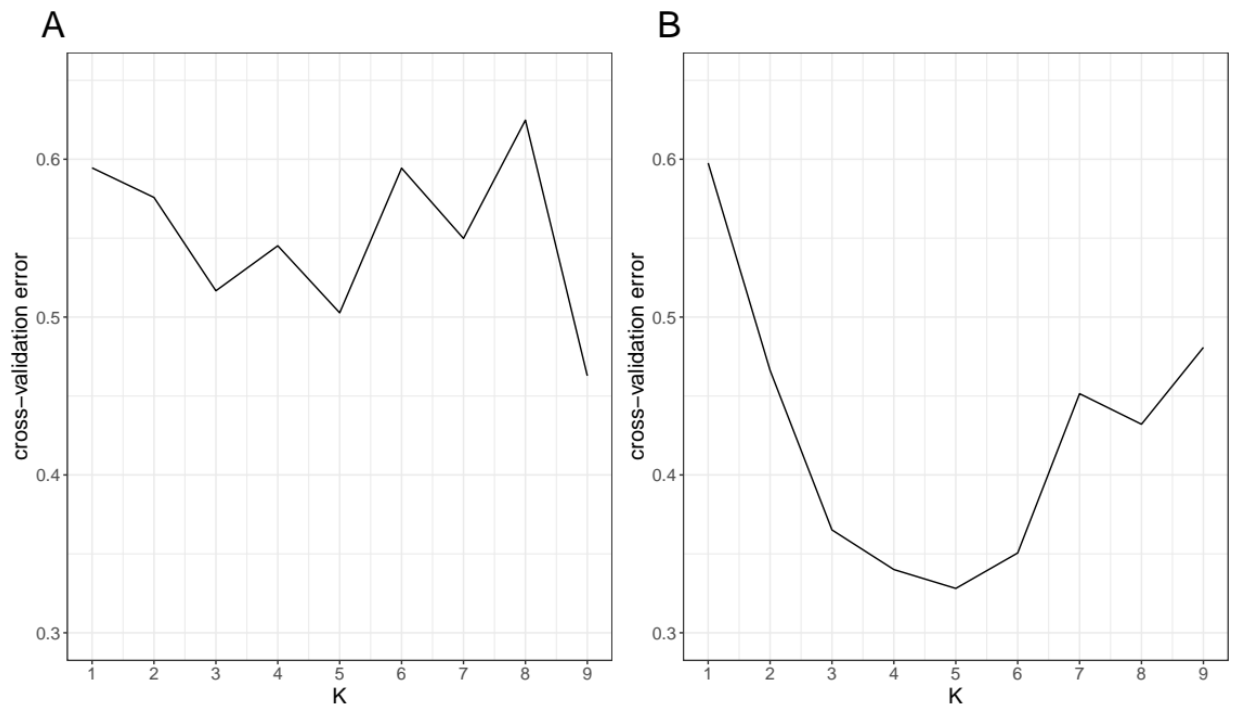


Figure S7 Admixture cross-validation errors (top) detected for values of K between 2 and 9 for A. Angiosperms353 data set and B. Bromeliad1776 data set.

MOLECULAR ECOLOGY RESOURCES

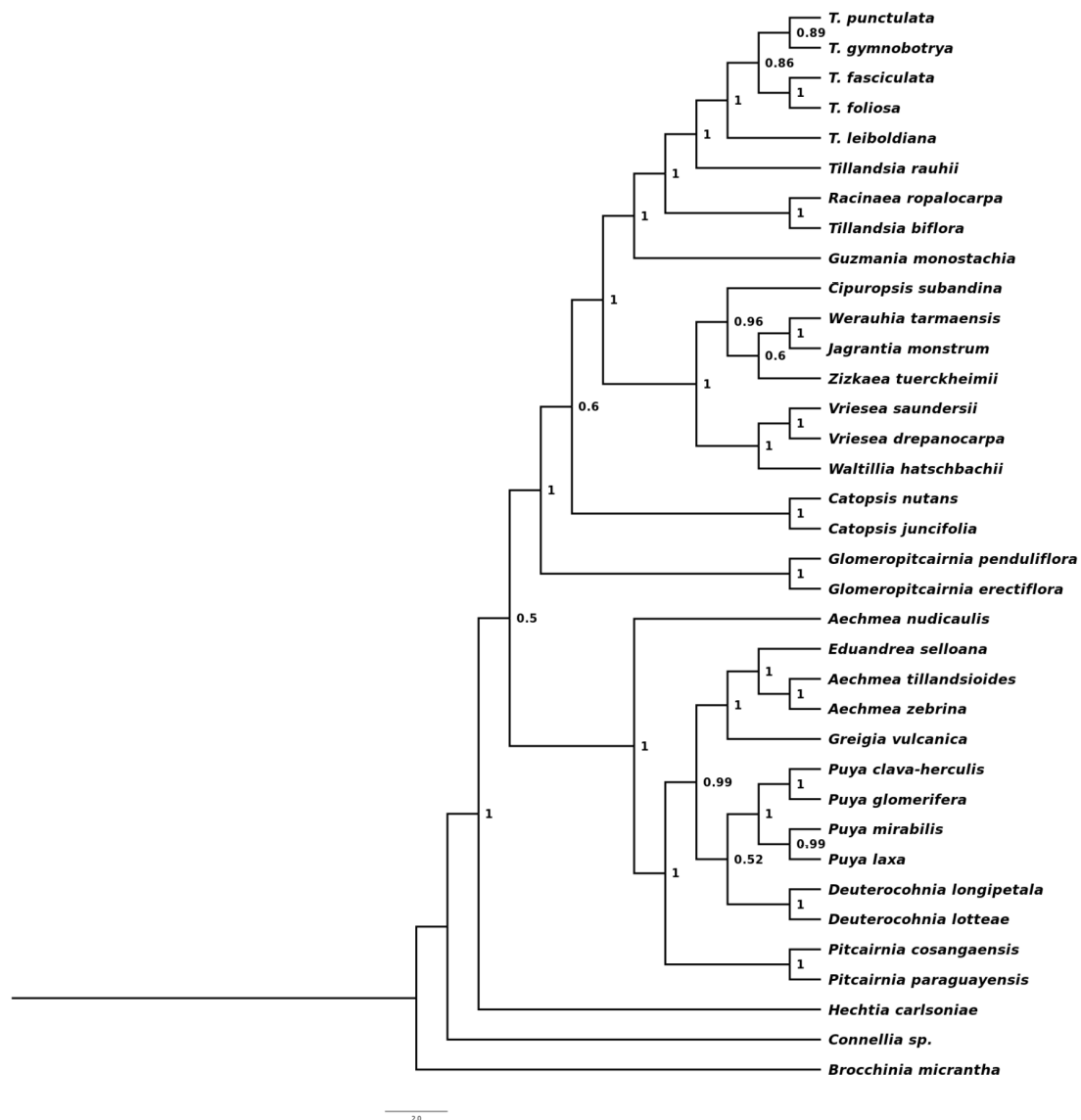


Figure S8 Coalescent-based species trees generated ASTRAL-III for samples enriched with Angiosperms353 using 269 genes. Node values represent local posterior probabilities (pp) for the main topology.

MOLECULAR ECOLOGY RESOURCES

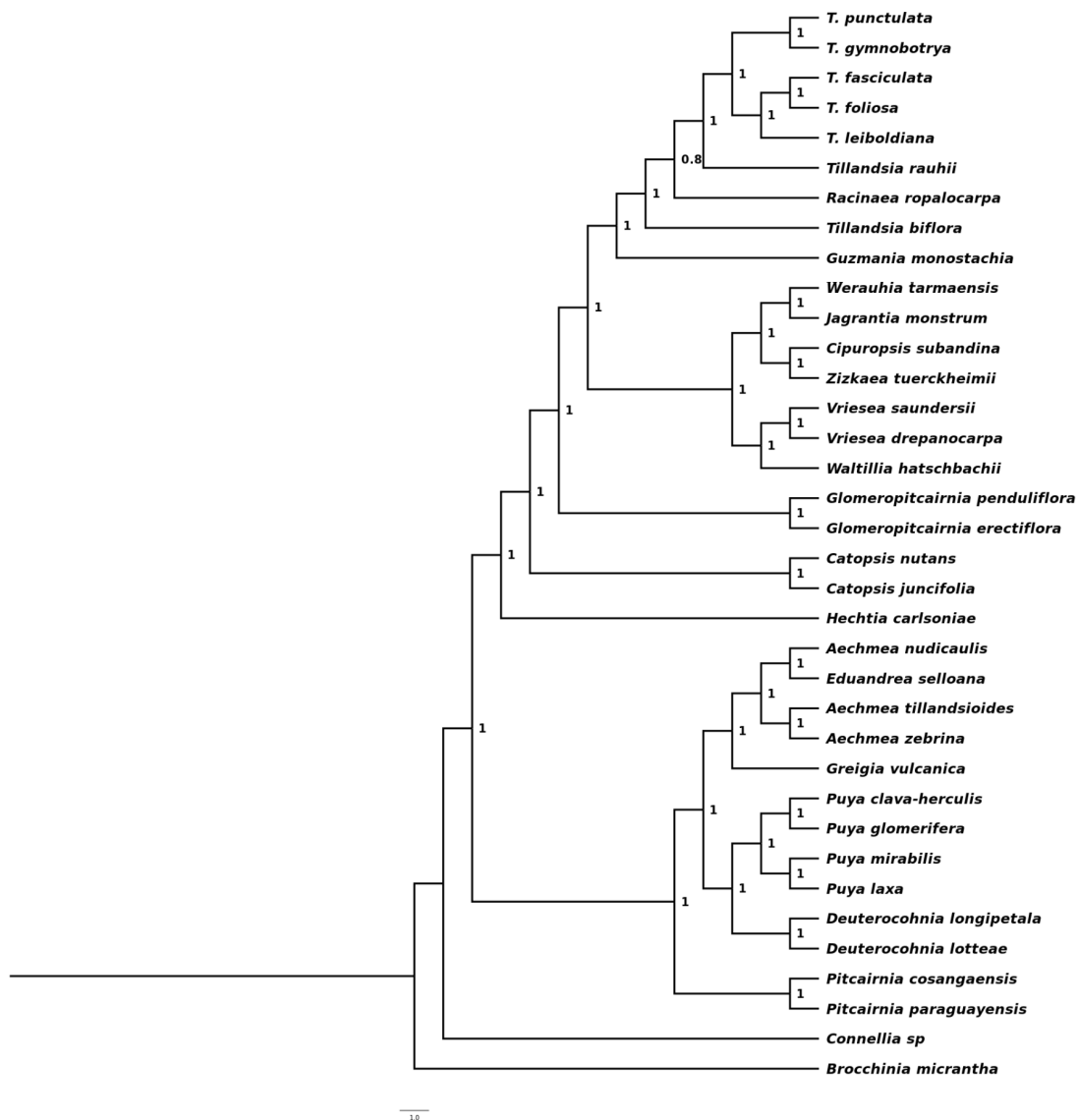


Figure S9 Coalescent-based species trees generated ASTRAL-III for samples enriched with Bromeliad1776 using 1600 genes. Node values represent local posterior probabilities (pp) for the main topology.

MOLECULAR ECOLOGY

RESOURCES

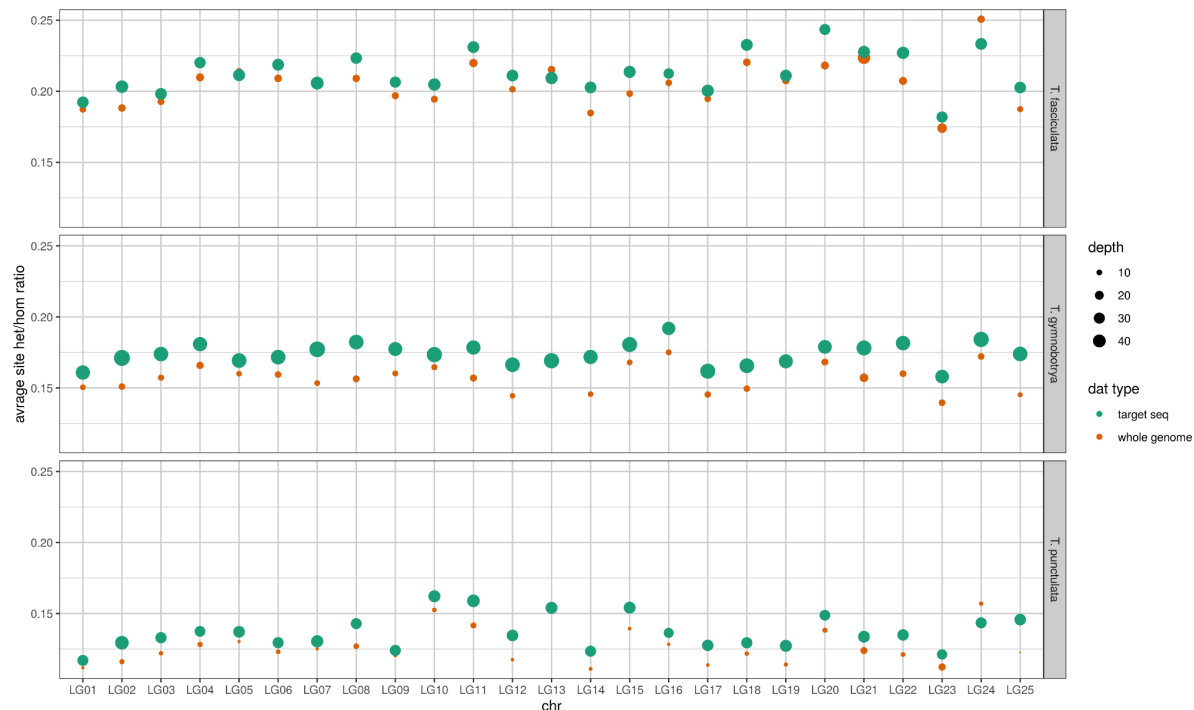
Supporting information: ascertainment bias

The non-random selection of markers for target capture may introduce ascertainment bias, with implications for population genomics methods (Heslot et al., 2013). An absence or misrepresentation of rare SNPs is an indication of such bias and may further result in loss of valuable information and misinterpretation of values, such as F_{st} and heterozygosity (Lachance & Tishkoff, 2013). To assess the extent of ascertainment bias that may rise in the use of Bromeliad1776, we assessed differences in SNP discovery by comparing whole genome data and target capture data in three *Tillandsia* samples, using the ratio of heterozygous to homozygous sites per LG as a measure to examine differences in SNP discovery.

Briefly, for each sample we obtained both whole-genome sequencing data and targeted sequencing data, produced with the bromeliad1776 bait-set. All accessions were reference mapped to the *A. comosus* v.3 reference (Ming, 2015). Variants were then called for each data type separately using freebayes v1.3.2-dirty (Garrison & Marth, 2012) - one variant call file for each data type, containing all three samples. Each vcf file was filtered to retain mean genotype depth > 4 and sites with 70% missing data using vcftools (Danecek et al., 2011). The ratio of heterozygous to homozygous sites was calculated for each individual and each LG (chr) using vcflib (Garrison, 2012).

The ratio heterozygous to homozygous sites was higher for target capture data sequencing in almost all species and loci, indicating high discovery of heterozygous sites. However, the differences in ratio were strongly correlated with sequencing depth ($P < 0.003$) and species ($P < 0.003$) and differences among species were more prominent than variation among data-types, indicating the former as causes for dissimilarities in SNP discovery, rather than marker selection (see figure below).

MOLECULAR ECOLOGY RESOURCES



Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156-2158.

Garrison, E. (2012). Vcflib: A C++ library for parsing and manipulating VCF files. GitHub <https://github.com/ekg/vcflib>.

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv:1207.3907 [q-bio]*.

Heslot, N., Rutkoski, J., Poland, J., Jannink, J.-L., & Sorrells, M. E. (2013). Impact of Marker Ascertainment Bias on Genomic Selection Accuracy and Estimates of Genetic Diversity. *PLOS ONE*, 8 (9), e74612. doi: 10.1371/journal.pone.0074612

Lachance, J., & Tishkoff, S. A. (2013). SNP ascertainment bias in population genetic analyses: Why it is important, and how to correct it. *BioEssays*, 35 (9), 780-786. doi: 10.1002/bies.201300014

Ming, R., VanBuren, R., Wai, C. M., Tang, H., Schatz, M. C., Bowers, J. E., . . . Yu, Q. (2015). The pineapple genome and the evolution of CAM photosynthesis. *Nature Genetics*, 47 (12), 1435-1442. doi: 10.1038/ng.3435

Table S1 Genes included in the Bromeliad1776 bait design, with identifiers as annotated in *Ananas comosus* genome v.3 (Ming et al., 2015). The table includes details about exon composition, copy number and putatively associated pathways. See legend in next sheet

gene	seqname	Single copy Broc. A	Low copy Broc. An	High copy Broc. An	Diff. expressed CA	pos. selection CAM	gene_families_CAM	Angiosperm_gene	CAM-related_Acomosus	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolites	malate transferase	Assimilation_Carbon	hydrates	flavonoid_anthocya	pos_sel_portulugo	glucogenesis_glycol	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc000037	LG16	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	593.250	4	2373	116
Acc000127	LG12	no	yes	no	yes	NA	NA	NA										NA	NA	NA			NA	299.000	2	598	254
Acc000152	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	381.000	4	1524	152
Acc000167	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	1268.000	1	1268	1268
Acc000168	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	2180.000	1	2180	2180
Acc000180	LG12	yes	no	no	NA	NA	NA	7067										NA	NA	NA			NA	136.333	9	1227	26
Acc000195	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	937.500	2	1875	514
Acc000210	LG12	no	yes	no	NA	NA	NA	NA										NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	226.857	7	1588	93
Acc000215	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	532.250	4	2129	196
Acc000229	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	1028.000	2	2056	459
Acc000232	LG12	no	yes	no	NA	NA	NA	NA										NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	564.500	2	1129	520
Acc000252	LG12	no	yes	no	NA	NA	NA	NA							yes			NA	NA	NA			00_#W_CircadianOscillator_flavin-binding_kelch-repeat_box1	1099.000	2	2198	410
Acc000269	LG12	yes	no	no	yes	NA	NA	NA										NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	426.800	5	2134	36
Acc000293	LG12	yes	no	no	NA	NA	NA	5922										NA	NA	NA			NA	181.636	11	1998	67
Acc000311	LG12	yes	no	no	NA	NA	NA	5335										NA	NA	NA			NA	121.500	16	1944	27
Acc000319	LG12	yes	no	no	NA	NA	NA	NA										STA_L_flavonol synthase	NA	NA	yes		11_Glycolysis_night_C4_malate_p athway	429.500	2	859	329
Acc000335	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	522.000	4	2088	154
Acc000337	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	215.100	10	2151	68
Acc000339	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	1592.000	1	1592	1592
Acc000352	LG12	yes	no	no	NA	NA	NA	NA							yes			NA	NA	NA			00_#W_ProtonPump_VHA-E	185.727	11	2043	23
Acc000385	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	367.800	5	1839	149
Acc000389	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	168.214	14	2355	23
Acc000396	LG12	no	yes	no	NA	NA	NA	NA										NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	882.000	3	2646	163
Acc000403	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	364.000	4	1456	214
Acc000413	LG12	no	no	yes	NA	NA	NA	NA										NA	NA	NA	yes		11_Glycolysis_night_C4_malate_p athway	136.929	14	1917	39
Acc000425	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	403.500	4	1614	303
Acc000456	LG12	yes	no	no	NA	NA	NA	NA										NA	NA	NA			NA	570.250	4	2281	248
Acc000482	LG12	NA	NA	NA	NA	NA	NA	NA				yes						NA	NA	NA			06_Drought_resistance_genes_LEA	269.000	1	269	269

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed CAM	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism	malate_transferrase	flavonoid_anthocyanin_Ps	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc000487	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	843.000	2	1696	572
Acc000491	LG12	no	yes	no	yes	NA	NA	NA									NA	NA			NA	713.000	2	1426	531
Acc000500	LG12	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	145.769	26	3790	27
Acc000507	LG12	NA	NA	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	217.667	6	1306	62
Acc000509	LG12	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	188.429	7	1319	36
Acc000523	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	346.500	4	1386	122
Acc000551	LG12	no	no	yes	NA	NA	NA	NA									ALB_F_leucoanthocyanidin dioxygenase	NA			NA	294.500	4	1178	78
Acc000552	LG12	yes	no	no	NA	NA	NA	5163									NA	NA			NA	132.000	14	1848	62
Acc000553	LG12	yes	no	no	NA	NA	NA	6924									NA	NA			NA	298.143	7	2087	162
Acc000572	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	301.167	6	1807	124
Acc000575	LG12	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	138.417	12	1661	34
Acc000577	LG12	NA	NA	NA	NA	NA	NA	NA								yes	NA	NA			03_VATPase	129.200	5	646	99
Acc000579	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1328.000	1	1328	1328
Acc000597	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1322.000	1	1322	1322
Acc000598	LG12	yes	no	no	NA	NA	NA	6780									NA	NA			NA	146.800	5	734	86
Acc000599	LG12	yes	no	no	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	355.000	3	1065	257
Acc000606	LG12	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	329.500	6	1977	26
Acc000610	LG12	no	yes	no	NA	NA	NA	4893									NA	NA			NA	171.556	9	1544	80
Acc000618	LG12	yes	no	no	NA	NA	NA	5980									NA	NA	yes		10_Glucogenesis_specific_dayC3	125.615	13	1633	63
Acc000632	LG12	yes	no	no	NA	NA	NA	5913									NA	NA			NA	369.500	4	1478	185
Acc000637	LG12	yes	no	no	NA	NA	NA	NA									ALB_L_dihydroflavono 4-reductase	NA	yes		10_Glucogenesis_specific_dayC3	254.000	7	1778	42
Acc000648	LG12	no	yes	no	NA	yes	NA	NA									NA	NA			NA	2834.000	1	2834	2834
Acc000651	LG12	yes	no	no	NA	NA	NA	6051									NA	NA			NA	1541.000	1	1541	1541
Acc000653	LG12	yes	no	no	yes	NA	NA	NA									NA	NA			NA	177.500	8	1420	52
Acc000666	LG12	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	257.286	7	1801	118
Acc000674	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1631.000	1	1631	1631
Acc000678	LG12	yes	no	no	NA	NA	NA	7628									NA	NA			NA	112.800	5	564	22
Acc000695	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1023.000	2	2046	482
Acc000707	LG12	no	yes	no	NA	NA	NA	NA									PREDICTED: Beta vulgaris subsp. vulgaris glucose-1-phosphate adenylyltransferase large subunit 3_chloroplast/camyloplastic (LOC104883901)_transcript variant X1_mRNA	NA	yes		10_Glucogenesis_specific_dayC3	172.938	16	2767	40
Acc000717	LG02	yes	no	no	NA	NA	NA	NA									ALB_L_bifunctional dihydroflavono 4-reductase flavanone 4-reductase	NA			NA	179.429	7	1256	137

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilia	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Acomnosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohyd	flavonoid_anthocyanin_Picalirina_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc000722	LG02	yes	no	no	yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		NA	NA	581.000	1	581	581
Acc000724	LG02	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		NA	NA	709.000	2	1418	526
Acc000727	LG02	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	443.000	1	443	443
Acc000731	LG02	no	yes	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	yes		11_Glycolysis_night_C4_malate_pathway	122.968	31	3812	8
Acc000735	LG02	no	yes	no	NA	NA	NA	NA	6176									NA	NA		NA	NA	160.308	13	2084	48
Acc000736	LG02	yes	no	no	yes	NA	NA	NA	NA									NA	NA		NA	NA	438.500	2	877	367
Acc000746	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	1679.000	1	1679	1679
Acc000748	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	772.250	4	3089	153
Acc000753	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	482.000	3	1446	229
Acc000754	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	1916.000	1	1916	1916
Acc000759	LG02	yes	no	no	NA	NA	NA	NA	6175									NA	NA		NA	NA	232.364	11	2556	94
Acc000767	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	736.333	3	2209	256
Acc000778	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	522.000	4	2088	331
Acc000793	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	734.667	3	2204	149
Acc000794	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	287.750	8	2302	146
Acc000795	LG02	yes	no	no	NA	NA	NA	NA	NA						yes			NA	PREDICTED: Beta vulgaris subsp. vulgaris tonoplast dicarboxylate transporter (LOC104900882)_mRNA			00_#W_MalateTransporter_IDT_11_Glycolysis_night_C4_malate_pathway	364.833	6	2189	43
Acc000798	LG02	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	356.625	8	2853	115
Acc000809	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	390.000	6	2340	122
Acc000814	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	445.714	7	3120	113
Acc000816	LG02	no	no	yes	NA	NA	NA	NA	NA									NA	PREDICTED: Solanum tuberosum heat shock cognate protein 80-like (LOC102600006)_mRNA			NA	506.000	5	2530	20
Acc000822	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	1571.000	1	1571	1571
Acc000823	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	399.600	5	1998	117
Acc000834	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	1179.750	4	4719	330
Acc000836	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	360.500	4	1442	148
Acc000843	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA		NA	NA	556.500	2	1113	434
Acc000868	LG02	no	yes	no	NA	NA	NA	NA	NA						yes			NA	NA			00_#Ming2019_Sl_gene	1151.000	1	1151	1151
Acc000907	LG02	no	yes	no	NA	NA	NA	NA	NA									ALB_F_UDP-glycosyltransferase	NA			NA	564.000	3	1692	98
Acc000950	LG02	no	no	yes	yes	NA	NA	NA	NA									NA	NA			NA	630.600	5	3153	67
Acc000965	LG02	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	637.500	2	1275	576
Acc000985	LG02	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA	386.500	2	773	213

gene	seqname	Single copy Broc_A n_Tilla	Low copy Broc_A n_Tilla	High copy Broc_A n_Tilla	Diff_expressed_CA an_Tilla	pos_selection_CAM_MCS	gene_C3_families_CAM-shifts	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_A nas	circadian_metabolis m_Wal2017	malate_transferease Assimilation_Carboh ydrates	flavonoid_anthocya nin_Picalirria_PS	pos_sel_portulugo	glucogenesis_glycol ysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc000993	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	262.727	11	2890	132
Acc000997	LG02	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	222.875	8	1783	53
Acc001001	LG02	yes	no	no	NA	yes	NA	NA									NA	NA			NA	822.000	3	2466	497
Acc001003	LG02	yes	no	no	NA	yes	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_p athway	128.357	14	1797	44
Acc001020	LG02	no	no	yes	NA	NA	NA	NA									NA				NA	320.091	11	3521	51
Acc001042	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	426.000	3	1278	310
Acc001049	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	406.667	3	1220	278
Acc001055	LG02	yes	no	no	yes	NA	NA	NA									NA	NA			NA	285.857	7	2001	66
Acc001068	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	344.400	5	1722	133
Acc001089	LG02	yes	no	no	NA	NA	NA	NA								yes	NA	NA			05_Chloroplasts_activity	387.000	3	1101	325
Acc001100	LG02	no	yes	no	NA	NA	NA	NA						yes			NA	NA			00_#Ming2019_Sl_gene	237.500	4	950	155
Acc001106	LG02	no	yes	no	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	191.000	4	764	61
Acc001112	LG02	yes	no	no	NA	NA	NA	7572									NA	NA			NA	378.000	4	1512	159
Acc001128	LG02	yes	no	no	NA	NA	NA	4890									NA	NA			NA	134.625	8	1077	35
Acc001146	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1916.000	1	1916	1916
Acc001178	LG02	yes	no	no	NA	NA	NA	6363									NA	NA			NA	197.182	11	2169	59
Acc001181	LG02	yes	no	no	NA	NA	NA	NA								yes	NA	NA			05_Chloroplasts_activity	606.500	2	1213	403
Acc001199	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1367.000	1	1367	1367
Acc001201	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1212.000	2	2424	915
Acc001216	LG02	no	yes	no	NA	NA	NA	NA						yes			NA	NA			00_#Ming2019_Sl_gene	474.000	5	2370	54
Acc001220	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	368.750	4	1475	124
Acc001222	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	554.000	2	1108	271
Acc001246	LG02	yes	no	no	NA	NA	NA	6389									NA	NA			NA	84.563	16	1353	27
Acc001261	LG02	yes	no	no	NA	NA	NA	NA	yes								NA	NA			00_#16_calciumdependent_protK	458.250	4	1833	106
Acc001317	LG02	yes	no	no	NA	NA	NA	6401									NA	NA			NA	617.500	2	1235	428
Acc001326	LG02	yes	no	no	NA	NA	NA	NA		yes							NA				00_#W_SomatolMovement_ABA_r esponsive_binding element_F101b_Absciscic_acid_sig nal	454.000	5	2270	26
Acc001329	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	304.778	9	2743	113
Acc001339	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	519.667	3	1559	228
Acc001350	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	386.800	5	1934	146
Acc001352	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	271.000	5	1355	113

gene	seqname	Single copy Broc_A n_Tilia.	Low copy Broc_Ana n_Tilia.	High copy Broc_Ana n_Tilia.	Diff_expressed CAM3	pos_selection_CAM3-shifts	gene_C3_families_CAM	Angiosperm_gene	CAM-related Mim2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana nas	circadian_metabolism_Wal2017	malate_transferrase Assimilation_Carbohyd rates	flavonoid_anthocyanin_Psicacitraria_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc001366	LG02	no	no	yes	NA	NA	NA	NA									NA	PREDICTED: Oryza sativa Japonica Group homeobox-leucine zipper protein ROC7 (LOC4344614)_mRNA			NA	204.636	11	2251	89
Acc001384	LG05	yes	no	no	NA	NA	NA	NA									NA				NA	560.500	2	1121	372
Acc001395	LG18	no	yes	no	NA	NA	NA	5942									NA				NA	652.000	2	1304	313
Acc001399	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	281.400	5	1407	180
Acc001415	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	381.667	3	1145	342
Acc001416	LG18	yes	no	no	yes	NA	NA	NA									NA				NA	259.333	3	778	132
Acc001430	LG18	no	yes	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	216.143	7	1513	44
Acc001431	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	1811.000	1	1811	1811
Acc001432	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	1379.000	1	1379	1379
Acc001435	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	1103.000	1	1103	1103
Acc001437	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	280.600	5	1403	111
Acc001473	LG18	no	yes	no	yes	NA	NA	NA									NA				NA	491.000	1	491	491
Acc001478	LG18	yes	no	no	NA	NA	NA	5366									NA				NA	108.529	17	1845	44
Acc001482	LG18	yes	no	no	yes	NA	NA	NA									NA				NA	288.500	2	577	254
Acc001508	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	424.667	3	1274	245
Acc001521	LG18	yes	no	no	NA	yes	NA	NA									NA				NA	104.500	12	1254	35
Acc001534	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	692.000	2	1384	690
Acc001542	LG18	yes	no	no	yes	NA	NA	NA									NA				NA	105.400	15	1581	27
Acc001543	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	2345.000	1	2345	2345
Acc001545	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	388.000	3	1164	210
Acc001553	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	652.000	2	1304	516
Acc001560	LG18	no	no	yes	NA	NA	NA	NA									NA		yes		11_Glycolysis_night_C4_malate_pathway	421.000	3	1263	330
Acc001566	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	526.000	3	1578	434
Acc001571	LG18	no	yes	no	NA	NA	NA	NA			yes						NA				06_Drought_resistance_genes_LEA	132.000	3	396	114
Acc001579	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	474.800	5	2374	170
Acc001580	LG18	no	yes	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	266.556	9	2399	4
Acc001608	LG18	NA	NA	NA	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	374.667	3	1124	272
Acc001618	LG18	no	no	yes	yes	NA	NA	NA									NA				NA	372.333	6	2234	191
Acc001621	LG18	yes	no	no	NA	NA	NA	NA									NA				NA	482.333	3	1447	170
Acc001637	LG18	no	yes	no	NA	NA	NA	NA		yes							NA				01d_Auxin_reponse_factor	247.692	13	3220	49
Acc001647	LG18	yes	no	no	yes	NA	NA	NA									NA				NA	377.000	1	377	377

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	Assimilation_Carbohyd	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portuligo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc001651	LG18	yes	no	no	yes	NA	NA	NA										NA	NA			NA	190.200	10	1902	59
Acc001675	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	299.500	4	1198	164
Acc001686	LG18	no	no	yes	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	298.571	7	2090	105
Acc001691	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	383.250	4	1533	196
Acc001693	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	2918.000	1	2918	2918
Acc001703	LG18	yes	no	no	NA	NA	NA	6550										NA	NA			NA	223.750	4	895	83
Acc001707	LG18	NA	NA	NA	NA	NA	NA	NA		yes								NA	NA			01b_Abscissic_acid_signal	290.333	3	871	71
Acc001718	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	612.000	2	1224	551
Acc001742	LG18	yes	no	no	NA	NA	NA	5357										NA	NA			NA	287.000	3	861	127
Acc001744	LG18	no	yes	no	NA	NA	yes	NA										NA	NA			NA	384.750	4	1539	134
Acc001767	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	221.600	5	1108	124
Acc001769	LG18	no	yes	no	NA	NA	NA	NA				yes						NA	NA			00_#MB_PHYE/B	852.400	5	4262	277
Acc001770	LG18	no	yes	no	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	146.588	17	2492	16
Acc001780	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	615.000	2	1230	518
Acc001783	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	656.500	2	1313	323
Acc001791	LG18	yes	no	no	NA	NA	NA	6450										NA	NA			NA	264.750	8	2118	104
Acc001803	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	1112.000	1	1112	1112
Acc001826	LG18	yes	no	no	NA	NA	NA	6034										NA	NA			NA	96.250	16	1540	48
Acc001830	LG18	no	yes	no	NA	NA	NA	5858										NA	NA			NA	261.167	6	1567	70
Acc001848	LG18	yes	no	no	yes	NA	NA	NA			yes							NA	NA			06_Drought_resistance_genes_LEA	518.000	2	1036	243
Acc001853	LG18	no	yes	no	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	228.875	8	1831	32
Acc001855	LG18	yes	no	no	NA	NA	NA	5138										NA	NA			NA	200.714	7	1405	79
Acc001857	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	633.000	2	1266	367
Acc001858	LG18	no	yes	no	NA	NA	NA	NA										ALB_F_acylglucose-dependent anthocyanin 7-o-glucosyltransferase	NA			NA	115.750	12	1389	25
Acc001863	LG18	no	yes	no	NA	NA	NA	5945										NA	NA			NA	99.222	18	1786	35
Acc001881	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	1763.000	1	1763	1763
Acc001894	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	2108.000	1	2108	2108
Acc001900	LG18	yes	no	no	yes	NA	NA	NA							yes			00_#W_SucroseTransporter_SWE ET11J12	NA			00_#W_SucroseTransporter_SWE ET11J12	170.000	7	1190	36
Acc001903	LG18	yes	no	no	NA	NA	NA	6717										NA	NA			NA	305.000	7	2135	90
Acc001926	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	278.000	5	1390	130
Acc001934	LG18	yes	no	no	NA	NA	NA	NA										NA	NA			NA	1235.000	1	1235	1235

gene	seqname	Single_copy_Broc_A an_Tilia.	Low_copy_Broc_Ana n_Tilia.	High_copy_Broc_An an_Tilia.	Diff_expressed_CA MC3	pos_selection_CAM _shifts	gene_C3 families_CAM	Angiosperm_gene	CAM- related_Acomosus- Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana nas	circadian_metabolis m_Walz2017	malate_transferase Assimilation_Carbohy drates	flavonoid_anthocya nin_Picalirnia_PS	pos_sel_portulungo	glucogenesis_glycol ysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc001941	LG18	no	yes	no	NA	NA	NA	NA		yes							NA	NA			01b_Absciscic_acid_signal	171.500	14	2401	46
Acc001945	LG18	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	288.000	3	864	43
Acc001954	LG18	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	140.667	24	3376	38
Acc001958	LG18	yes	no	no	NA	NA	NA	NA									NA	NA			NA	928.000	3	2784	341
Acc001967	LG18	yes	no	no	NA	NA	NA	NA									NA	NA			NA	361.400	5	1807	174
Acc001976	LG18	yes	no	no	NA	NA	NA	6483									NA	NA	yes		10_Glucogenesis_specific_dayC3	148.889	9	1340	53
Acc001994	LG18	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1919.000	1	1919	1919
Acc002015	LG04	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_p athway	388.000	3	1104	315
Acc002037	LG04	no	yes	no	yes	NA	NA	NA									NA	NA			NA	320.429	7	2243	15
Acc002055	LG04	no	yes	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	647.000	1	647	647
Acc002071	LG04	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	169.625	8	1357	51
Acc002118	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	643.000	2	1286	579
Acc002171	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	741.333	3	2224	175
Acc002187	LG04	yes	no	no	NA	yes	NA	NA									NA	NA			NA	291.357	14	4079	61
Acc002196	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1722.500	2	3445	497
Acc002227	LG04	no	yes	no	NA	NA	NA	NA								yes	NA	NA			04_Oxygen_evolving_enhancer_O EE	259.500	4	1038	50
Acc002247	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	576.667	3	1730	225
Acc002248	LG04	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	282.571	7	1978	87
Acc002253	LG04	no	yes	no	yes	NA	NA	NA									NA	NA			NA	428.375	8	3427	83
Acc002268	LG04	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	830.000	1	830	830
Acc002270	LG04	yes	no	no	NA	NA	NA	5469									NA	NA			NA	248.000	5	1240	127
Acc002284	LG04	yes	no	no	NA	yes	NA	NA									NA	NA			NA	475.500	4	1902	253
Acc002286	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	291.667	6	1750	142
Acc002296	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	386.000	5	1930	112
Acc002305	LG04	yes	no	no	yes	NA	NA	NA									NA	NA			NA	709.667	3	2129	422
Acc002324	LG04	no	yes	no	yes	NA	NA	NA									NA	NA			NA	384.000	5	1920	39
Acc002356	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	406.333	3	1219	264
Acc002358	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	818.000	3	2454	300
Acc002366	LG04	no	yes	no	yes	NA	NA	NA									NA	NA			NA	297.500	2	595	14
Acc002378	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	267.500	6	1605	180
Acc002400	LG04	yes	no	no	yes	NA	NA	NA									NA	NA			NA	123.667	9	1113	11

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed CAM	pos_selection_CAM	gene_families_CAM	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohyd	flavonoid_anthocyanin_Psic	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc002415	LG04	no	yes	no	NA	NA	NA	NA				yes					NA	Arabiopsis thaliana Nucleoporin interacting component (Nup93/Nie96-like) family protein mRNA		NA	NA	177.647	17	3020	53
Acc002475	LG04	NA	NA	NA	NA	NA	NA	NA			yes						NA				06_Drought_resistance_genes_LEA	372.000	2	744	177
Acc002476	LG04	yes	no	no	NA	NA	NA	NA						yes			NA				00_#W_FructoseTransporter_SWIE	214.167	6	1295	36
Acc002484	LG04	yes	no	no	NA	NA	NA	NA									NA				NA	896.000	2	1792	484
Acc002493	LG04	yes	no	no	NA	NA	NA	NA									NA				NA	513.000	3	1539	294
Acc002499	LG04	no	yes	no	NA	NA	NA	NA							yes		NA				00_#W_StomatalMovement_ethylene_insensitive1	362.600	5	1813	102
Acc002508	LG04	no	yes	no	NA	yes	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	784.000	3	2352	306
Acc002524	LG04	yes	no	no	NA	NA	NA	NA									NA				NA	1112.000	1	1112	1112
Acc002530	LG04	no	no	yes	yes	NA	NA	NA									NA				NA	232.250	4	929	55
Acc002531	LG04	yes	no	no	NA	NA	NA	NA									NA		yes		NA	820.000	3	2460	395
Acc002546	LG04	yes	no	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	398.000	1	398	398
Acc002547	LG04	yes	no	no	NA	NA	NA	NA									NA				NA	465.000	3	1395	339
Acc002572	LG04	no	yes	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	611.667	3	1835	449
Acc002583	LG04	no	yes	no	NA	NA	yes	NA									NA				NA	170.000	6	1020	15
Acc002605	LG06	no	yes	no	NA	yes	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	339.333	9	3054	53
Acc002616	LG06	yes	no	no	yes	NA	NA	NA									NA				NA	443.000	1	443	443
Acc002647	LG06	no	yes	no	NA	NA	NA	NA				yes					NA				00_#MB_NIA	525.200	5	2626	140
Acc002667	LG06	NA	NA	NA	yes	NA	NA	NA									NA				NA	222.429	7	1557	59
Acc002669	LG06	yes	no	no	NA	NA	NA	NA									NA				NA	424.000	3	1272	236
Acc002671	LG06	yes	no	no	NA	NA	NA	6379									NA				NA	123.000	3	369	77
Acc002677	LG06	yes	no	no	NA	yes	NA	NA									ALB_F_flavonol synthase				NA	914.000	2	1828	409
Acc002683	LG06	yes	no	no	NA	NA	NA	NA									NA				NA	459.000	3	1377	327
Acc002697	LG06	yes	no	no	yes	NA	NA	NA									NA				NA	241.500	4	966	60
Acc002732	LG06	yes	no	no	NA	NA	NA	NA	yes								NA				00_#04_carbonic_anhydrase	165.000	8	1320	48
Acc002736	LG06	NA	NA	NA	NA	NA	NA	NA									NA				NA	986.000	1	986	986
Acc002739	LG06	no	yes	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	167.643	14	2347	52
Acc002744	LG06	no	yes	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	128.000	9	1152	11
Acc002760	LG06	yes	no	no	NA	NA	NA	NA									NA				NA	279.500	6	1677	18
Acc002775	LG06	yes	no	no	NA	NA	NA	NA									NA				NA	444.333	3	1333	191
Acc002797	LG06	yes	no	no	NA	NA	NA	NA									NA				NA	934.250	4	3737	129

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilla	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohyd	flavonoid_anthocyanin_Picalitima_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc002844	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	937.000	2	1874	452
Acc002868	LG06	yes	no	no	NA	NA	NA	NA	NA									PREDICTED: Glycine max cullin-3A-like (LOC100794520)_mRNA	NA			NA	887.000	3	2661	65
Acc002869	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	227.750	8	1822	155
Acc002875	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1419.500	2	2839	580
Acc002885	LG06	NA	NA	NA	NA	NA	NA	NA	NA	yes								NA	00_#20_hypothetical_protein			00_#20_hypothetical_protein	941.000	2	1882	240
Acc002887	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	323.000	5	1615	169
Acc002912	LG06	yes	no	no	NA	NA	NA	NA	5343									NA	NA			NA	417.500	2	835	329
Acc002939	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	2510.000	1	2510	2510
Acc002940	LG06	no	yes	no	NA	NA	NA	NA	NA									PREDICTED: Beta vulgaris subsp. vulgaris alanine-tRNA ligase chloroplastic/mitochondrial (LOC104903847)_transcript variant X1_mRNA	NA			NA	293.000	13	3809	71
Acc002942	LG06	no	yes	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	196.625	8	1573	53
Acc002964	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1646.000	2	3292	125
Acc002968	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	545.250	4	2181	151
Acc002978	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	332.800	5	1664	166
Acc002990	LG06	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	111.205	44	4893	46
Acc003023	LG06	yes	no	no	NA	NA	NA	NA	NA							yes		NA	NA			00_#W_MalateTransporter_ALMT-10_Glucogenesis_specific_dayC3	429.667	6	2578	136
Acc003024	LG06	no	yes	no	NA	NA	NA	NA	5260									NA	NA			NA	208.833	6	1253	60
Acc003031	LG06	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_p	205.833	12	2470	61
Acc003034	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			11_Glycolysis_night_C4_malate_p	232.833	6	1397	131
Acc003091	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	322.667	6	1936	155
Acc003094	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	392.333	3	1177	379
Acc003097	LG06	no	no	yes	NA	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_p	591.667	3	1775	333
Acc003098	LG06	yes	no	no	NA	NA	NA	NA	5477									NA	NA			NA	178.833	6	1073	65
Acc003126	LG06	no	yes	no	NA	NA	NA	NA	NA	yes								NA	NA			01b_Absciscic_acid_signal	109.600	20	2192	20
Acc003131	LG06	no	no	yes	NA	NA	NA	NA	NA									NA	NA			NA	187.364	11	2061	3
Acc003135	LG06	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	986.000	4	3984	119
Acc003139	LG06	no	yes	no	NA	NA	NA	NA	NA				yes					NA	00_#Zizka_AGT1_#04_carbonic_a			00_#Zizka_AGT1_#04_carbonic_a	248.429	7	1739	48
Acc003181	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1334.000	1	1334	1334
Acc003188	LG17	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	253.000	5	1265	65
Acc003190	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			01b_Absciscic_acid_signal	362.500	2	725	229
Acc003191	LG17	yes	no	no	yes	NA	NA	NA	NA									NA	NA			01b_Absciscic_acid_signal	518.500	2	1037	445

gene	seqname	Single copy Broc_A nan_Tilla	Low copy Broc_A nan_Tilla	High copy Broc_A nan_Tilla	Diff_expressed CAM3	pos_selection_CAM-shifts	gene_C3_families_CAM	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Anas	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carboxhydrates	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac00030192	LG17	no	yes	no	NA	NA	yes	NA									NA	NA			NA	175.167	6	1051	88
Ac0003218	LG17	no	yes	no	yes	NA	NA	NA									NA	NA			NA	220.846	13	2871	73
Ac0003220	LG17	no	yes	no	NA	NA	NA	NA								yes	NA	NA			03_VATPase	405.000	3	1215	285
Ac0003224	LG17	yes	no	no	NA	yes	NA	NA									NA	NA			NA	300.000	6	1800	67
Ac0003228	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1529.000	1	1529	1529
Ac0003251	LG17	NA	NA	NA	yes	NA	NA	NA									NA	NA			NA	260.000	1	260	260
Ac0003266	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	633.500	2	1267	314
Ac0003271	LG17	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	181.750	8	1454	58
Ac0003277	LG17	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_pathway	178.900	10	1788	68
Ac0003279	LG17	no	no	yes	NA	yes	NA	NA									NA	NA			NA	356.889	9	3212	110
Ac0003284	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	862.500	2	1725	316
Ac0003298	LG17	no	yes	no	NA	NA	NA	NA									NA	Oryza sativa Japonica Group cDNA clone:JO23046C06_full insert sequence			NA	325.571	7	2279	178
Ac0003301	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2087.000	1	2087	2087
Ac0003312	LG17	yes	no	no	NA	NA	yes	NA									NA	NA			NA	402.500	2	805	348
Ac0003313	LG17	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Vitis vinifera transcriptional corepressor SEUSS (LOC100241984)_transcript variant X1_mRNA			NA	294.700	10	2947	48
Ac0003317	LG17	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	395.333	6	2372	127
Ac0003321	LG17	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	178.100	10	1781	46
Ac0003326	LG17	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	165.933	30	4978	54
Ac0003337	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	369.000	3	1107	134
Ac0003341	LG17	yes	no	no	NA	NA	NA	7331									NA	NA			NA	224.833	12	2698	53
Ac0003358	LG17	no	yes	no	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_reponse_factor	226.417	12	2717	56
Ac0003363	LG17	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	130.889	9	1178	60
Ac0003393	LG17	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	198.000	7	1386	36
Ac0003414	LG17	yes	no	no	yes	NA	NA	NA									NA	NA			NA	1046.000	1	1046	1046
Ac0003425	LG17	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	360.000	8	2880	76
Ac0003430	LG17	no	no	yes	yes	NA	NA	NA									NA	NA			NA	329.000	4	1316	120
Ac0003453	LG17	yes	no	no	NA	NA	NA	6507									NA	NA			NA	207.000	10	2070	67
Ac0003454	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	472.000	3	1416	308
Ac0003456	LG17	yes	no	no	yes	NA	NA	NA									NA	NA			NA	372.500	2	745	71
Ac0003459	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1370.000	1	1370	1370
Ac0003463	LG17	yes	no	no	NA	NA	NA	4796									NA	NA			NA	139.636	11	1536	44

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilia	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate transferase Assimilation_Carbohyd	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0003472	LG17	yes	no	no	NA	NA	NA	NA	6995									NA	NA			NA	178.333	6	1070	29
Ac0003480	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	386.000	3	1158	282
Ac0003503	LG17	yes	no	no	NA	NA	NA	NA	6407									NA	NA			NA	171.000	3	513	87
Ac0003507	LG17	no	yes	no	yes	NA	NA	NA	NA									NA	NA			NA	695.000	2	1390	65
Ac0003515	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1307.500	2	2615	1129
Ac0003517	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	743.500	2	1487	215
Ac0003518	LG17	yes	no	no	NA	NA	NA	NA	5434									NA	NA			NA	458.000	3	1374	101
Ac0003538	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	547.000	3	1641	494
Ac0003547	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	2261.000	1	2261	2261
Ac0003556	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	240.200	5	1201	133
Ac0003571	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	391.857	7	2743	161
Ac0003578	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	728.333	3	2185	383
Ac0003605	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	683.333	3	2050	191
Ac0003633	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	688.000	2	1376	290
Ac0003638	LG17	no	yes	no	yes	NA	NA	NA	NA									NA	NA			NA	331.667	6	1990	7
Ac0003639	LG17	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA	169.000	10	1690	55
Ac0003654	LG17	yes	no	no	NA	NA	NA	NA	7577									NA	NA			NA	257.600	5	1288	64
Ac0003687	LG17	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	981.500	2	1963	926
Ac0003688	LG17	no	yes	no	NA	NA	NA	NA	NA									10_Glucogenesis_specific_dayC3		yes			1481.000	1	1481	1481
Ac0003689	LG17	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA	134.778	9	1213	57
Ac0003757	LG15	no	no	yes	NA	NA	NA	NA	NA									PREDICTED: Glycine max histidine kinase 2-like (LOC100781655)_transcript variant X3_mRNA				NA	304.786	14	4267	51
Ac0003760	LG15	NA	NA	NA	yes	NA	NA	NA	NA									NA	NA			NA	128.167	6	769	35
Ac0003761	LG15	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1190.000	1	1190	1190
Ac0003762	LG15	yes	no	no	NA	NA	NA	NA	5168									NA	NA			NA	149.625	8	1197	56
Ac0003785	LG15	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	827.000	2	1654	666
Ac0003793	LG15	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	380.000	4	1520	159
Ac0003798	LG15	yes	no	no	NA	NA	NA	NA	5264									NA	NA			NA	219.250	12	2631	62
Ac0003802	LG15	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	2384.000	1	2384	2384
Ac0003806	LG15	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	599.000	2	1198	439
Ac0003826	LG15	no	yes	no	NA	NA	NA	NA	NA									11_Glycolysis_right_C4_malate_p		yes			118.500	8	948	65
Ac0003857	LG15	no	yes	no	NA	NA	NA	NA	NA									ALB_Flavanoid 3- hydroxylase				NA	400.333	6	2402	165

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilia	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohyd	flavonoid_anthocyanin_PScitral	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc003860	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	899.333	3	2518	210
Acc003871	LG15	no	yes	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	yes	NA	10_Glucogenesis_specific_dayC3	899.000	2	1798	350
Acc003873	LG15	no	yes	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	yes	NA	NA	NA	NA	NA	NA	00_Ming2019_Anthocyanin_biosynthesis	230.000	9	2070	65
Acc003885	LG15	no	yes	no	yes	NA	NA	NA	NA	NA	NA	NA	NA	yes	yes	NA	ALB_F_Flavonoid 3-hydroxylase	NA	NA	NA	NA	00_Ming2019_Anthocyanin_biosynthesis	313.400	5	1567	11
Acc003924	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	390.000	4	1560	162
Acc003927	LG15	yes	no	no	NA	NA	NA	NA	5449	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	222.353	17	3780	44
Acc003928	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	2099.000	1	2099	2099
Acc003937	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	yes	NA	10_Glucogenesis_specific_dayC3	650.000	1	650	650
Acc003946	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	862.500	2	1725	745
Acc003984	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1331.000	1	1331	1331
Acc004017	LG15	yes	no	no	NA	NA	NA	NA	5802	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	183.000	6	1098	106
Acc004039	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	583.000	2	1166	560
Acc004109	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	509.667	6	3058	151
Acc004131	LG15	no	yes	no	yes	NA	NA	NA	NA	NA	NA	NA	NA	yes	NA	NA	NA	NA	NA	NA	NA	00_WW_SucroseTransporter_10_Glucogenesis_specific_dayC3	164.643	14	2305	33
Acc004135	LG15	no	yes	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	yes	NA	NA	NA	NA	NA	NA	NA	00_WW_SucroseTransporter_SUT1_10_Glucogenesis_specific_dayC3	160.286	14	2244	33
Acc004155	LG15	yes	no	no	NA	NA	yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	3872.000	1	3872	3872
Acc004156	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	317.250	4	1269	164
Acc004162	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	yes	NA	10_Glucogenesis_specific_dayC3	152.000	8	1216	56
Acc004170	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1253.000	1	1253	1253
Acc004181	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1415.000	1	1415	1415
Acc004198	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1161.000	2	2322	166
Acc004199	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	858.000	2	1716	497
Acc004202	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	PREDICTED: Beta vulgaris subsp. vulgaris CWF19-like protein 2 (LOC104905381)_transcript variant X1_mRNA	NA	NA	NA	351.600	10	3516	73
Acc004218	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	yes	NA	10_Glucogenesis_specific_dayC3	189.800	5	949	57
Acc004222	LG15	NA	NA	NA	yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	90.500	2	181	32
Acc004236	LG15	no	no	yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	PREDICTED: Glycine max protein argonaute 1-like (LOC100778754)_mRNA	NA	NA	NA	150.083	24	3602	31
Acc004250	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1510.000	2	3020	296
Acc004281	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	1991.000	1	1991	1991
Acc004283	LG15	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	384.000	3	1152	194
Acc004290	LG15	no	yes	no	NA	NA	NA	NA	NA	NA	yes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	01b_Abscisic_acid_signal	1787.000	1	1787	1787

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilla	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Wal2017	malate_transferrase_Assimilation_Carbohyd	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0004297	LG05	no	yes	no	NA	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	282.800	5	1414	61
Acc0004304	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	2198.000	1	2198	2198
Acc0004321	LG05	no	yes	no	NA	NA	yes	NA	NA									NA	PREDICTED: Beta vulgaris subsp. vulgaris mitochondrial dicarboxylate/tricarboxylate transporter DTC (LOC104906718)_mRNA			NA	230.286	7	1612	13
Acc0004349	LG05	no	yes	no	NA	NA	NA	NA	NA	yes								NA	NA			00_#21_Malate_dehydrogenase	199.286	7	1395	61
Acc0004370	LG05	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	653.000	3	1959	138
Acc0004374	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	680.500	2	1361	447
Acc0004375	LG05	no	yes	no	NA	NA	NA	NA	NA		yes							NA	NA			01b_Abscisic_acid_signal	409.667	3	1229	215
Acc0004393	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	575.000	2	1150	505
Acc0004398	LG05	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA	685.000	2	1370	415
Acc0004401	LG05	NA	NA	NA	NA	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	309.000	2	618	304
Acc0004406	LG05	yes	no	no	NA	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	689.000	1	689	689
Acc0004430	LG05	no	yes	no	NA	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	364.000	3	1092	247
Acc0004447	LG05	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA	145.000	8	1160	52
Acc0004448	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	340.000	6	2040	168
Acc0004452	LG05	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA	1257.500	2	2515	316
Acc0004465	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	480.500	4	1922	116
Acc0004471	LG05	no	no	yes	yes	NA	NA	NA	NA									NA	NA			NA	685.000	2	1330	602
Acc0004476	LG05	yes	no	no	NA	NA	NA	NA	5893									NA	NA			NA	297.250	4	1189	25
Acc0004483	LG05	yes	no	no	NA	NA	NA	NA	5220									NA	NA			NA	312.000	4	1248	98
Acc0004488	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	739.500	2	1479	385
Acc0004489	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	287.250	4	1149	116
Acc0004512	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	701.000	2	1402	568
Acc0004532	LG05	no	yes	no	NA	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	635.000	1	635	635
Acc0004555	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	530.000	3	1590	254
Acc0004564	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	681.667	3	2045	315
Acc0004579	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1745.000	2	3490	687
Acc0004585	LG05	yes	no	no	NA	NA	NA	NA	6003									NA	NA			NA	409.000	2	818	224
Acc0004612	LG05	no	no	yes	NA	NA	NA	NA	NA									NA	PREDICTED: Glucine max histidine kinase 2-like (LOC100781655)_transcript variant X3_mRNA			NA	375.727	11	4133	51
Acc0004613	LG05	no	yes	no	yes	NA	NA	NA	NA									NA	NA			NA	250.857	7	1756	44
Acc0004638	LG05	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	154.533	15	2318	51

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc004644	LG05	yes	no	no	NA	NA	NA	NA									NA	NA			NA	238.714	7	1671	134
Acc004646	LG05	NA	NA	NA	NA	NA	NA	NA									NA	NA		yes	09_Circadian_clock	137.500	6	825	60
Acc004662	LG05	yes	no	no	NA	NA	NA	NA									NA	NA			NA	917.000	2	1834	167
Acc004672	LG05	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1115.500	2	2231	330
Acc004679	LG05	yes	no	no	NA	NA	NA	6383									NA	NA			NA	185.333	12	2224	29
Acc004700	LG05	yes	no	no	NA	NA	NA	NA									NA	NA			NA	512.800	5	2564	287
Acc004708	LG05	yes	no	no	NA	NA	NA	6506									NA	NA			NA	136.429	7	955	48
Acc004773	LG05	no	yes	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	306.400	5	1532	29
Acc004804	LG05	yes	no	no	yes	NA	NA	NA									NA	NA			NA	373.250	4	1483	232
Acc004808	LG05	yes	no	no	NA	NA	NA	NA									NA	NA			NA	372.000	3	1116	309
Acc004821	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	877.000	2	1754	739
Acc004827	LG07	no	yes	no	NA	NA	NA	NA				yes					NA	NA			00_#MB_PhyC	1026.200	5	5131	283
Acc004844	LG07	yes	no	no	yes	NA	NA	NA									NA	NA			NA	211.111	9	1900	78
Acc004853	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	582.500	2	1165	264
Acc004854	LG07	yes	no	no	yes	NA	NA	NA									NA	NA			NA	246.500	4	986	33
Acc004855	LG07	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	191.000	8	1528	74
Acc004857	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	491.333	3	1474	269
Acc004859	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1913.000	1	1913	1913
Acc004860	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1217.000	1	1217	1217
Acc004861	LG07	no	yes	no	NA	yes	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	1442.000	1	1442	1442
Acc004864	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1232.500	2	2465	757
Acc004888	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1078.000	2	2156	1073
Acc004899	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	385.000	3	1155	199
Acc004902	LG07	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_pathway	171.250	8	1370	65
Acc004919	LG07	no	no	yes	NA	yes	NA	6527									NA	NA			NA	258.048	21	5419	54
Acc004920	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	369.667	3	1109	172
Acc004926	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	539.400	5	2697	211
Acc004930	LG07	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	98.867	15	1483	50
Acc004946	LG07	yes	no	no	NA	NA	yes	6492									NA	NA			NA	161.111	9	1450	69
Acc004947	LG07	no	no	yes	NA	NA	NA	NA							yes		NA	NA			00_#W_SemataMovement_RootP	662.500	4	2650	69
Acc004960	LG07	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Beta vulgaris subsp. vulgaris PREDICTED: DETOXIFICATION 23 (LOC104889970) mRNA			NA	209.200	10	2092	56

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilia	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohyd	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc004989	LG07	yes	no	no	NA	NA	NA	NA	5395									NA	NA			NA	485.000	3	1455	298
Acc004996	LG07	no	yes	no	NA	NA	NA	NA	NA	yes								NA	NA			00_#29_Malate_dehydrogenase	161.444	9	1453	35
Acc005022	LG07	no	no	yes	yes	NA	NA	NA	NA									NA	NA			NA	285.333	6	1592	50
Acc005037	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1220.000	1	1220	1220
Acc005040	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	152.750	12	1833	58
Acc005041	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1991.000	1	1991	1991
Acc005042	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA		yes	10_Glucogenesis_specific_dayC3	178.167	6	1069	82
Acc005047	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	2012.000	1	2012	2012
Acc005049	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1700.000	1	1700	1700
Acc005054	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA		yes	10_Glucogenesis_specific_dayC3	229.750	8	1838	22
Acc005055	LG07	NA	NA	NA	yes	NA	NA	NA	NA									NA	NA			NA	318.000	2	636	220
Acc005078	LG07	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA	270.000	6	1620	56
Acc005083	LG07	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA	220.333	6	1322	44
Acc005101	LG07	no	yes	no	yes	NA	NA	NA	NA									NA	NA			NA	468.500	6	2811	70
Acc005103	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1271.000	1	1271	1271
Acc005104	LG07	yes	no	no	NA	NA	NA	NA	6533									NA	NA			NA	172.200	10	1722	55
Acc005108	LG07	no	yes	no	NA	NA	NA	NA	NA							yes		NA	NA			09_Circadian_clock	272.000	3	816	120
Acc005109	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA		yes	10_Glucogenesis_specific_dayC3	453.750	4	1815	43
Acc005110	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1220.000	1	1220	1220
Acc005112	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	318.000	4	1272	197
Acc005118	LG07	yes	no	no	NA	NA	NA	NA	6439									NA	NA			NA	210.800	5	1054	86
Acc005119	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	448.250	4	1793	188
Acc005125	LG07	yes	no	no	NA	NA	NA	NA	5866									NA	NA			NA	120.909	11	1330	4
Acc005127	LG07	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA	391.333	6	2348	56
Acc005134	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	462.667	3	1388	150
Acc005140	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	425.000	3	1275	310
Acc005155	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1937.000	1	1937	1937
Acc005164	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA		yes	10_Glucogenesis_specific_dayC3	135.667	12	1628	41
Acc005168	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	433.750	4	1735	131
Acc005173	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	345.500	4	1382	134
Acc005209	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	167.444	9	1507	50

gene	seqname	Single copy Broc_A n_Tilla.	Low copy Broc_A n_Tilla.	High copy Broc_A n_Tilla.	Diff_expressed CAM3	pos_selection_CAM-shifts	gene_C3_families_CAM	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Anas	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohydrites	flavonoid_anthocyanin_Picalirnia_PS	pos_sel_portuligo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc005229	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	390.000	3	1170	302
Acc005233	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	596.250	4	2385	111
Acc005239	LG07	yes	no	no	NA	NA	NA	NA									ALB_F_anthocyanidin 3-O-glucosyltransferase	NA			NA	1433.000	1	1433	1433
Acc005251	LG07	yes	no	no	NA	yes	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	219.875	8	1759	61
Acc005262	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	605.500	2	1211	395
Acc005283	LG07	no	yes	no	yes	NA	NA	NA									NA	NA			NA	729.000	4	2916	125
Acc005287	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	556.333	3	1669	385
Acc005291	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	276.286	7	1934	145
Acc005343	LG11	no	no	yes	yes	NA	NA	NA									NA	NA			NA	279.353	17	4749	6
Acc005353	LG11	yes	no	no	NA	NA	NA	5721									NA	NA			NA	337.167	6	2023	69
Acc005354	LG11	NA	NA	NA	NA	yes	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	800.000	2	1600	728
Acc005358	LG11	no	yes	no	yes	NA	NA	NA							yes		NA	NA			05_Chloroplasts_activity	584.000	2	1168	415
Acc005359	LG11	no	yes	no	yes	NA	NA	NA							yes		NA	NA			05_Chloroplasts_activity	680.000	2	1320	495
Acc005366	LG11	no	yes	no	yes	NA	NA	NA									NA	NA			NA	586.250	8	4530	65
Acc005373	LG11	no	yes	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	246.500	4	986	109
Acc005379	LG11	no	yes	no	NA	NA	NA	NA			yes						NA	NA			00_#W_HexoseTransporter GLUT 6/6-like_transporter	116.160	25	2904	16
Acc005384	LG11	yes	no	no	NA	NA	NA	5271									NA	NA			NA	160.571	7	1124	45
Acc005385	LG11	yes	no	no	NA	NA	NA	6026									NA	NA			NA	136.000	9	1224	38
Acc005390	LG11	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2282.000	1	2282	2282
Acc005393	LG11	no	yes	no	NA	NA	NA	NA									ALB_F_anthocyanin 3-o-beta-glucosyltransferase	NA			NA	729.500	6	4377	39
Acc005397	LG11	yes	no	no	NA	NA	NA	NA									ALB_F_dihydroflavonol 4-reductase	NA			NA	288.000	5	1440	153
Acc005399	LG11	yes	no	no	NA	yes	NA	NA									NA	NA			NA	926.000	1	926	926
Acc005402	LG11	no	yes	no	NA	NA	NA	NA							yes		NA	NA			00_#H06_carbonic_anhydrase	144.750	8	1158	48
Acc005409	LG11	yes	no	no	NA	NA	NA	NA									NA	NA			NA	866.000	2	1732	401
Acc005439	LG11	yes	no	no	NA	NA	NA	NA									NA	NA			NA	508.333	3	1525	223
Acc005444	LG11	yes	no	no	NA	NA	NA	NA									NA	NA			NA	455.000	3	1365	321
Acc005457	LG11	yes	no	no	yes	NA	NA	NA									NA	NA			NA	1235.000	1	1235	1235
Acc005478	LG11	yes	no	no	NA	NA	NA	NA									NA	NA			NA	203.500	6	1221	111
Acc005505	LG11	no	no	yes	NA	yes	NA	NA									NA	NA			NA	134.158	19	2549	48
Acc005508	LG11	yes	no	no	NA	NA	NA	NA									NA	NA			NA	339.750	4	1359	127
Acc005513	LG11	NA	NA	NA	yes	NA	NA	NA		yes							NA	NA			01b_Abscisic_acid_signal	565.500	2	1131	476

gene	seqname	Single copy Broc_A	Low copy Broc_A	High copy Broc_A	Diff_expressed_CA	CAM-related Mim2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolics	malate transferase m_Wal2017	Assimilation_Carbohydrtates	flavonoid_anthocyanin_Picalirria_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc005520	LG11	yes	no	no	NA	NA									NA	NA			NA	455.333	3	1366	128
Acc005552	LG11	no	yes	no	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	256.500	6	1539	139
Acc005561	LG11	yes	no	no	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	710.000	1	710	710
Acc005562	LG11	no	yes	no	NA	NA									ALB_L_anthocyanidin 3-O-glucosyltransferase	NA		NA	1481.000	1	1481	1481	
Acc005564	LG11	NA	NA	NA	NA	NA									ALB_F_isoflavone reductase	NA		NA	160.000	6	960	18	
Acc005567	LG11	NA	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	81.500	4	326	23
Acc005570	LG11	yes	no	no	NA	NA									NA	NA			NA	1952.000	1	1952	1952
Acc005587	LG11	yes	no	no	NA	NA									NA	NA			NA	381.000	3	1143	152
Acc005599	LG11	yes	no	no	NA	NA									NA	NA			NA	2074.000	2	4148	611
Acc005614	LG11	yes	no	no	NA	NA									NA	NA			NA	389.000	3	1167	225
Acc005616	LG11	yes	no	no	NA	NA									NA	NA			NA	216.000	8	1728	116
Acc005623	LG11	yes	no	no	NA	NA									NA	NA			NA	597.500	4	2390	119
Acc005631	LG11	no	yes	no	NA	NA	yes								NA	PREDICTED: Beta vulgaris subsp. vulgaris NADP-dependent malic enzyme (LOC104883143)_mRNA			00_#33_Malic_enzyme	113.125	16	1810	47
Acc005646	LG11	no	no	yes	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_pathway	139.200	15	2088	39
Acc005654	LG11	no	yes	no	NA	NA	6958								NA	NA			NA	151.500	8	1212	82
Acc005657	LG11	no	yes	no	NA	NA									NA	NA			NA	537.500	2	1075	427
Acc005667	LG11	yes	no	no	NA	NA									NA	NA			NA	567.500	2	1135	279
Acc005669	LG11	yes	no	no	NA	NA									NA	NA			NA	462.000	3	1386	267
Acc005671	LG11	yes	no	no	NA	NA									NA	NA			NA	599.000	2	1198	503
Acc005673	LG11	yes	no	no	NA	NA									NA	NA			NA	717.500	2	1435	570
Acc005684	LG11	yes	no	no	NA	NA									NA	NA			NA	390.667	3	1172	129
Acc005715	LG11	yes	no	no	yes	NA							yes		NA	NA			03_VATPase	167.900	10	1679	48
Acc005746	LG11	yes	no	no	NA	NA									NA	NA			NA	2411.000	1	2411	2411
Acc005794	LG11	yes	no	no	NA	NA									NA	NA			NA	199.900	10	1999	32
Acc005821	LG16	yes	no	no	NA	NA									NA	NA			NA	439.000	5	2195	231
Acc005826	LG16	no	yes	no	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	373.000	3	1119	182
Acc005867	LG16	yes	no	no	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	125.667	12	1508	52
Acc005868	LG16	yes	no	no	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	794.000	1	794	794
Acc005891	LG16	yes	no	no	NA	NA									NA	NA			NA	183.900	10	1839	39
Acc005908	LG16	no	yes	no	NA	NA							yes		NA	NA			03_VATPase	160.143	7	1121	19
Acc005915	LG16	yes	no	no	yes	NA									NA	NA			NA	220.667	6	1324	67

gene	seqname	Single_copy_Broc_An	Low_copy_Broc_An	High_copy_Broc_An	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_lunc	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolis	malate_transferase	flavonoid_anthocya	pos_sel_portulugo	glucogenesis_glycol	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0059930	LG16	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_p athway	192.444	9	1732	61
Acc0059931	LG16	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_p athway	211.778	9	1906	61
Acc0059947	LG16	yes	no	no	yes	NA	NA	NA									NA	NA			NA	585.000	2	1170	570
Acc0059956	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	657.500	2	1315	451
Acc0059968	LG16	no	no	yes	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_p athway	149.200	15	2238	39
Acc0059975	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	334.600	5	1673	134
Acc0059989	LG16	no	yes	no	NA	NA	NA	NA	yes								NA	PREDICTED: Beta vulgaris subsp. vulgaris NADP-dependent malic enzyme (LOC104883143)_ mRNA			00_#34_Malic_enzyme	86.529	17	1471	28
Acc0059991	LG16	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Glycine max kinesis-4-like (LOC100777089)_ transcript variant X4_ mRNA			NA	174.095	21	3656	29
Acc0060015	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	312.875	8	2503	186
Acc0060030	LG16	no	yes	no	NA	yes	NA	5206									NA	NA			NA	282.500	4	1130	170
Acc0060043	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	273.455	11	3008	116
Acc0060056	LG16	yes	no	no	yes	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	451.500	2	903	386
Acc0060063	LG16	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_p athway	286.333	9	2397	31
Acc0060064	LG16	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_p athway	252.800	10	2528	31
Acc0060068	LG16	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	194.091	11	2135	62
Acc0060084	LG16	no	yes	no	yes	NA	NA	NA									NA	NA			NA	626.333	3	1879	349
Acc0060096	LG16	yes	no	no	yes	yes	NA	NA									NA	NA			NA	571.000	3	1713	406
Acc0060122	LG16	no	yes	no	NA	NA	NA	NA	yes								NA	NA			00_#17_Malate_dehydrogenase	229.143	7	1604	61
Acc0060129	LG16	yes	no	no	NA	NA	NA	6923									NA	NA			NA	167.909	11	1847	64
Acc0060140	LG16	yes	no	no	yes	NA	NA	NA									NA	NA			NA	2138.000	1	2138	2138
Acc0060144	LG16	yes	no	no	NA	NA	NA	NA									NA	PREDICTED: Beta vulgaris subsp. vulgaris shikimate dehydrogenase BAS2_ chloroplastic (LOC104902128)_ mRNA			NA	130.385	13	1695	41
Acc0060150	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1136.000	1	1136	1136
Acc0060157	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	384.571	7	2692	123
Acc0060164	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1325.000	1	1325	1325
Acc0060172	LG16	yes	no	no	NA	yes	NA	7141									NA	NA			NA	405.667	3	1217	164
Acc0060181	LG16	yes	no	no	NA	NA	NA	NA	yes								NA	PREDICTED: Beta vulgaris subsp. vulgaris carbonic anhydrase_chloroplastic-like (LOC104902037)_ mRNA			00_#05_carbonic_anhydrase	146.250	8	1170	48
Acc0060186	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	239.167	6	1435	122
Acc0060199	LG16	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Beta vulgaris subsp. vulgaris glucose-1-phosphate adenylyltransferase large subunit 3_chloroplastic/amyloplastic (LOC104883901)_ transcript variant X1_ mRNA	yes		10_Glucogenesis_specific_dayC3	134.063	16	2145	55
Acc0060225	LG16	yes	no	no	NA	NA	NA	6883									NA	NA			NA	154.143	7	1079	60

gene	seqname	Single_copy_Broc_A nan_Tilia.	Low_copy_Broc_Ana n_Tilia.	High_copy_Broc_An an_Tilia.	Diff_expressed_CA MC3	CAM- related_Acomosus_ Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana nas	circadian_metabolis m_Walz2017	malateTtransferase Assimilation_Carboh ydrates	flavonoid_anthocya nin_Pitcairnia_PS	pos_sel_portulugo	glucogenesis_glycol ysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size					
Acc006227	LG16	say	no	no	NA	NA							NA	NA	NA			NA	211.500	14	2961	54					
Acc006252	LG14	no	yes	no	NA	NA							NA	NA	NA			NA	163.000	3	489	55					
Acc006253	LG14	no	no	yes	NA	NA							NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	656.333	3	1969	356					
Acc006280	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	1055.000	2	2110	160					
Acc006311	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	184.667	6	1108	39					
Acc006312	LG14	say	no	no	NA	NA							NA	NA	NA			NA	1508.000	1	1508	1508					
Acc006319	LG14	no	yes	no	NA	NA							NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	436.667	6	2620	139					
Acc006345	LG14	yes	no	no	NA	yes							NA	NA	NA			NA	226.000	14	3164	119					
Acc006349	LG14	no	no	yes	NA	NA							NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	185.467	15	2782	77					
Acc006380	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	264.800	5	1324	104					
Acc006394	LG14	no	yes	no	yes	NA							NA	NA	NA			NA	578.500	2	1157	349					
Acc006403	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	122.455	11	1347	18					
Acc006406	LG14	no	yes	no	NA	NA							NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	284.500	10	2845	41					
Acc006409	LG14	say	no	no	NA	NA							NA	NA	NA			NA	2159.000	1	2159	2159					
Acc006416	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	543.500	2	1087	339					
Acc006436	LG14	no	yes	no	NA	yes							NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	155.000	9	1395	12					
Acc006438	LG14	no	yes	no	yes	NA							NA	NA	NA			NA	626.000	1	626	626					
Acc006440	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	167.250	8	1338	68					
Acc006453	LG14	no	yes	no	NA	NA							NA	PREDICTED: Beta vulgaris subsp. vulgaris adenyate kinase 5, chloroplastic (LOC104901827)_transcript variant X1_ mRNA										97.000	18	1746	6
Acc006461	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	1118.000	1	1118	1118					
Acc006467	LG14	NA	NA	NA	NA	NA							NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	219.400	5	1097	154					
Acc006469	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	2582.000	1	2582	2582					
Acc006470	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	382.750	4	1571	158					
Acc006479	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	629.000	2	1258	409					
Acc006487	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	597.333	3	1792	318					
Acc006523	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	1988.000	1	1988	1988					
Acc006524	LG14	no	no	yes	NA	NA							NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	154.778	18	2786	36					
Acc006529	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	1229.000	1	1229	1229					
Acc006536	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	222.833	6	1337	123					
Acc006540	LG14	yes	no	no	NA	NA							NA	NA	NA			NA	374.000	2	748	333					

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilia	Diff_expressed CAM_Mc3	pos_selection_CAM-shifts	gene_families_CAM-C3	Angiosperm_gene	CAM-related_Acomnosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate transferase	yltrates	flavonoid_anthocyanin_Picairma_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc006549	LG14	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA						NA	NA	NA			NA	738.500	2	1477	360
Acc006583	LG14	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA						NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	118.846	13	1545	24
Acc006586	LG14	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA						NA	NA	NA	yes		11_Glycolysis_right_C4_malate_P	85.409	22	1879	18
Acc006591	LG14	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA						NA	NA	NA			NA	2375.000	1	2375	2375
Acc006594	LG14	yes	no	no	NA	NA	NA	NA	NA	NA	NA	NA						NA	NA	NA			NA	589.000	2	1178	449
Acc006606	LG14	no	no	yes	yes	NA	NA	NA	NA	NA	NA	NA						NA	NA	NA			NA	287.333	6	1784	77
Acc006629	LG01	yes	no	no	NA	NA	NA	NA	6947									NA	NA	NA			NA	256.600	10	2566	47
Acc006642	LG01	no	yes	no	NA	NA	NA	NA	5899									NA	NA	NA			NA	251.750	12	3021	99
Acc006659	LG01	yes	no	no	NA	NA	NA	yes	NA									NA	NA	NA			NA	1016.000	1	1016	1016
Acc006675	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	1005.500	2	2011	968
Acc006679	LG01	yes	no	no	NA	NA	NA	NA	6636									NA	NA	NA			NA	208.400	5	1042	70
Acc006692	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	Populus trichocarpa hypothetical protein (POPTR_0001s35130g) mRNA_complete cds	yes		10_Glucogenesis_specific_dayC3	209.000	9	1881	68
Acc006693	LG01	yes	no	no	NA	NA	yes	NA	5032									NA	NA	NA			NA	140.643	14	1969	31
Acc006698	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	257.800	5	1289	129
Acc006711	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	847.000	3	2541	148
Acc006719	LG01	yes	no	no	NA	NA	NA	NA	NA								yes	NA	NA	NA			04_Oxygen_evolution_enhancer_OEE	451.333	3	1354	293
Acc006733	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	230.000	7	1610	60
Acc006735	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	694.333	3	2083	200
Acc006741	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	PREDICTED: Beta vulgaris subsp. vulgaris aspartate aminotransferase_cytoplasmic (LOC104906385)_mRNA			NA	138.667	12	1664	61
Acc006754	LG01	yes	no	no	NA	NA	NA	yes	NA									NA	NA	NA			NA	142.750	4	571	80
Acc006755	LG01	yes	no	no	NA	NA	NA	NA	NA						yes			NA	NA	NA			NA	695.000	2	1270	210
Acc006769	LG01	yes	no	no	NA	NA	NA	NA	NA						yes			ALB_L_dihydroflavonol_4-reductase	NA	NA			00_#Ming2019_Anthocyanin_biosynthesis	226.500	6	1359	125
Acc006847	LG01	no	yes	no	NA	NA	NA	NA	NA				yes					NA	NA	NA			06_Drought_resistance_genes_LEA	185.000	2	370	3
Acc006861	LG22	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	901.500	2	1803	356
Acc006874	LG22	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	201.000	7	1407	118
Acc006886	LG22	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	1229.000	1	1229	1229
Acc006923	LG22	no	yes	no	NA	NA	NA	NA	NA				yes					NA	NA	NA			06_Drought_resistance_genes_LEA	225.667	3	677	113
Acc006939	LG22	yes	no	no	NA	NA	NA	NA	6875									NA	NA	NA			NA	1475.000	1	1475	1475
Acc006949	LG22	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	617.250	4	2469	169
Acc006978	LG22	no	yes	no	NA	NA	NA	yes	NA									NA	NA	NA			NA	189.308	13	2461	42
Acc006993	LG22	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	615.333	3	1846	396

gene	seqname	Single-copy Broc_A	Low-copy Broc_Ana	High-copy Broc_Ana	an_Tilia	Diff_expressed CA	pos_selection_CAM	gene_families_CAM	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	flavonoid_anthocyanin_Ps	pos_sel_portulugo	glucogenesis_ghycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0007014	LG23	no	yes	no	yes	yes	yes	NA	NA									NA	NA			NA	282.833	12	3394	43
Ac0007033	LG23	yes	no	no	NA	NA	NA	NA	NA								yes	NA	NA			03_VATPase	173.412	17	2948	44
Ac0007035	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	351.429	7	2460	117
Ac0007041	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1075.500	2	2151	364
Ac0007046	LG23	no	no	yes	NA	NA	NA	NA	NA									ALB_F_leucoanthocyanidin dioxygenase	NA			NA	363.000	3	1089	257
Ac0007054	LG23	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	204.250	8	1634	65
Ac0007057	LG23	yes	no	no	NA	NA	NA	NA	6620									NA	NA			NA	564.500	2	1129	424
Ac0007065	LG23	no	yes	no	NA	NA	NA	NA	6318									NA	NA			NA	168.294	17	2861	60
Ac0007085	LG23	no	no	yes	yes	yes	yes	NA	NA									NA	NA			NA	693.000	3	2079	133
Ac0007096	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	381.400	5	1907	152
Ac0007098	LG23	yes	no	no	NA	NA	NA	NA	NA									NA				NA	221.042	24	5305	78
Ac0007115	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	426.800	5	2134	119
Ac0007119	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	624.333	3	1873	142
Ac0007198	LG23	NA	NA	NA	NA	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	330.000	3	990	4
Ac0007204	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	587.500	4	2350	329
Ac0007212	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	405.333	3	1216	124
Ac0007214	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	386.000	5	1830	159
Ac0007234	LG23	no	yes	no	NA	NA	NA	NA	6538									NA	NA			NA	314.667	3	944	247
Ac0007241	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	681.333	3	2044	442
Ac0007248	LG23	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	656.571	7	4596	139
Ac0007258	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	612.400	5	3062	126
Ac0007277	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	446.500	4	1786	158
Ac0007291	LG23	no	yes	no	NA	NA	NA	NA	NA		yes							NA	NA			01b_Absciscic_acid_signal	170.929	14	2393	46
Ac0007294	LG23	yes	no	no	NA	NA	NA	NA	6738									NA	NA			NA	300.500	2	601	260
Ac0007304	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	960.000	2	1920	619
Ac0007306	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	208.889	9	1880	46
Ac0007307	LG23	no	yes	no	NA	NA	NA	NA	4992									NA	NA			NA	179.136	22	3941	92
Ac0007319	LG23	yes	no	no	NA	NA	NA	NA	5348									NA	NA			NA	186.571	7	1306	75
Ac0007326	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1700.000	1	1700	1700
Ac0007376	LG23	yes	no	no	NA	NA	NA	NA	6016									NA	NA			NA	523.000	6	3138	97
Ac0007384	LG23	yes	no	no	NA	NA	NA	NA	6048									NA	NA			NA	131.500	8	1052	37

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related Mim2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Wal2017	malate transferase Assimilation_Carbohyd	flavonoid_anthocyanin_Psicaria	pos_sel_portuligo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc007400	LG12	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	525.000	3	1575	132
Acc007413	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2126.000	1	2126	2126
Acc007423	LG12	no	no	yes	NA	NA	NA	NA									NA	PREDICTED: Oriza sativa Japonica Group homeobox-leucine zipper protein ROC7 (LOC4344614). mRNA			NA	262.000	12	3144	95
Acc007440	LG12	no	yes	no	yes	NA	NA	NA									NA	NA			NA	537.000	5	2685	181
Acc007441	LG12	no	no	yes	NA	NA	NA	NA							yes		NA	00_#W_StomatalMovement_RootP_heterodispin2			00_#W_StomatalMovement_RootP_heterodispin2	456.800	5	2284	69
Acc007443	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	223.714	7	1566	128
Acc007445	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	626.500	2	1253	166
Acc007457	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1316.000	1	1316	1316
Acc007465	LG12	no	yes	no	yes	NA	NA	NA									NA	NA			NA	100.529	17	1709	38
Acc007469	LG12	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	128.538	13	1671	38
Acc007483	LG12	NA	NA	NA	NA	yes	NA	NA									NA	NA			NA	117.105	19	2225	22
Acc007490	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	228.000	5	1140	124
Acc007494	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	519.000	3	1557	486
Acc007498	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	286.250	4	1145	132
Acc007502	LG12	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	134.381	21	2822	32
Acc007508	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	587.667	3	1763	144
Acc007509	LG12	NA	NA	NA	yes	NA	NA	NA									NA	NA			NA	446.000	1	446	446
Acc007534	LG08	yes	no	no	NA	NA	NA	NA									NA	NA			NA	702.500	2	1405	347
Acc007539	LG08	no	no	yes	NA	yes	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	307.400	5	1537	140
Acc007548	LG08	yes	no	no	yes	NA	NA	NA									NA	NA			NA	1484.000	1	1484	1484
Acc007561	LG08	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	287.636	11	3164	37
Acc007570	LG08	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_pathway	180.800	10	1808	63
Acc007575	LG08	no	yes	no	yes	NA	NA	NA									NA	NA			NA	152.900	10	1529	62
Acc007580	LG08	yes	no	no	NA	NA	NA	NA									NA	NA			NA	778.000	2	1556	586
Acc007622	LG08	no	yes	no	NA	NA	NA	NA	yes								NA	PREDICTED: Beta vulgaris subsp. vulgaris NAD-dependent malic enzyme 59 kDa isoform_mitochondrial (LOC10488560). mRNA			00_#36_Malic_enzyme	130.765	17	2223	54
Acc007629	LG08	no	yes	no	NA	NA	NA	4691									NA	NA			NA	207.250	4	829	142
Acc007645	LG08	yes	no	no	yes	NA	NA	NA									NA	NA			NA	3146.000	1	3146	3146
Acc007647	LG08	yes	no	no	NA	NA	NA	NA									NA	NA			NA	220.750	12	2649	111
Acc007652	LG08	yes	no	no	yes	NA	NA	NA									NA	NA			NA	132.160	25	3304	45
Acc007659	LG08	no	yes	no	yes	NA	NA	NA									NA	NA			NA	266.167	6	1597	85

gene	seqname	Single copy Broc_A n_Tilla.	Low copy Broc_A n_Tilla.	High copy Broc_A n_Tilla.	Diff_expressed CA n_Tilla.	pos_selection_CAM_MC3	gene_C3_families_CAM_shifts	Angiosperm_gene	CAM-related Mim2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Anas	circadian_metabolism_Walz2017	malate_transferase_Assimilation_Carbohydrites	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc007664	LG08	yes	no	no	yes	NA	NA	NA									NA	NA			NA	760.000	3	2290	396
Acc007670	LG08	no	yes	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	785.000	1	785	785
Acc007671	LG08	yes	no	no	NA	NA	NA	NA									NA	NA			NA	636.333	3	1909	166
Acc007678	LG08	yes	no	no	yes	NA	NA	NA									NA	NA			NA	809.000	1	809	809
Acc007684	LG08	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2558.000	1	2558	2558
Acc007686	LG08	yes	no	no	NA	NA	NA	NA									NA	PREDICTED: Glycine max actin-related protein 4-like (LOC100795390)_mRNA			NA	92.550	20	1851	49
Acc007699	LG08	yes	no	no	NA	NA	NA	NA						yes			NA	00_Ming2019_Anthocyanin_biosynthesis			00_Ming2019_Anthocyanin_biosynthesis	593.000	2	1186	198
Acc007718	LG08	no	yes	no	NA	NA	NA	NA			yes						NA	02_Aquaporins			02_Aquaporins	242.600	5	1213	41
Acc007730	LG08	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1301.000	1	1301	1301
Acc007734	LG08	no	yes	no	NA	NA	NA	NA	yes								NA	Beta vulgaris susp. vulgaris malate dehydrogenase, cytoplasmic (LOC104887930)_mRNA			00_#17_Malate_dehydrogenase	217.143	7	1520	111
Acc007753	LG08	yes	no	no	NA	NA	NA	NA									NA	NA			NA	350.750	4	1403	155
Acc007762	LG21	no	yes	no	yes	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	126.813	16	2029	59
Acc007777	LG21	yes	no	no	yes	NA	NA	NA									NA	NA			NA	247.667	3	743	29
Acc007782	LG21	no	no	yes	NA	NA	NA	NA									NA	NA			NA	358.429	7	2509	48
Acc007800	LG21	yes	no	no	NA	NA	NA	7111									NA	NA			NA	123.500	8	988	70
Acc007803	LG21	yes	no	no	yes	yes	NA	NA	yes								NA	00_#01_carbonic_anhydrase			00_#01_carbonic_anhydrase	202.167	6	1213	102
Acc007818	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	564.000	4	2256	122
Acc007825	LG21	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	377.714	7	2644	32
Acc007828	LG21	yes	no	no	NA	NA	NA	NA				yes					NA	06_Drought_resistance_genes_LEA			06_Drought_resistance_genes_LEA	432.500	2	865	353
Acc007832	LG21	yes	no	no	NA	NA	NA	6447									NA	NA			NA	303.000	3	909	200
Acc007857	LG21	yes	no	no	NA	NA	NA	NA		yes							NA	01b_Absciscic_acid_signal			01b_Absciscic_acid_signal	575.000	2	1150	514
Acc007859	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	418.250	4	1673	193
Acc007876	LG21	yes	no	no	NA	NA	yes	NA									NA	NA			NA	416.000	2	832	282
Acc007882	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	401.667	3	1205	227
Acc007890	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1391.000	1	1391	1391
Acc007901	LG21	no	yes	no	yes	NA	NA	NA									NA	NA			NA	544.111	9	4897	85
Acc007906	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	515.250	4	2061	210
Acc007911	LG21	yes	no	no	NA	NA	NA	6148									NA	NA			NA	244.333	3	733	152
Acc007924	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2132.000	1	2132	2132
Acc007928	LG21	yes	no	no	NA	NA	NA	NA									NA	10_Glucogenesis_specific_dayC3	yes		10_Glucogenesis_specific_dayC3	233.625	8	1869	99
Acc007953	LG21	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	178.882	17	3041	56

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc007973	LG21	no	yes	no	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_response_factor	272.385	13	3541	56
Acc007974	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2810.000	1	2810	2810
Acc007975	LG21	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	253.692	13	3298	53
Acc008032	LG21	no	yes	no	NA	NA	NA	5489									NA	NA			NA	149.067	15	2236	38
Acc008037	LG21	no	yes	no	NA	NA	yes	NA									NA	NA			NA	204.077	13	2653	44
Acc008046	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	329.250	4	1317	178
Acc008062	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	553.000	3	1659	350
Acc008066	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	442.200	5	2211	177
Acc008078	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1919.000	1	1919	1919
Acc008099	LG21	NA	NA	NA	NA	NA	NA	NA				yes					NA	NA			00_#MB_MS	422.000	4	1688	328
Acc008105	LG21	no	yes	no	yes	NA	NA	NA									NA	NA			NA	282.833	12	3394	16
Acc008126	LG21	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	125.563	16	2009	36
Acc008167	LG19	no	no	yes	NA	NA	NA	NA							yes		NA	NA			00_#W_ProtonPump_AVP1	369.625	8	2957	59
Acc008175	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1433.000	1	1433	1433
Acc008192	LG19	no	no	yes	NA	NA	NA	NA							yes		NA	NA			00_#W_StomatalMovement_AHA2	222.214	14	3111	96
Acc008206	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	399.000	3	1197	158
Acc008234	LG19	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	229.455	11	2524	107
Acc008236	LG19	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	233.000	2	466	172
Acc008246	LG19	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	190.583	12	2287	70
Acc008257	LG19	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	259.000	6	1554	60
Acc008287	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1580.000	1	1580	1580
Acc008292	LG19	yes	no	no	yes	NA	NA	6405									NA	NA			NA	162.500	12	1950	56
Acc008298	LG19	yes	no	no	yes	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	179.000	7	1253	50
Acc008300	LG19	yes	no	no	NA	NA	NA	NA							yes		NA	NA			00_#W_ProtonPump_VHA-A	125.050	20	2501	37
Acc008315	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	469.667	3	1409	127
Acc008321	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	396.250	4	1585	136
Acc008330	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	422.000	3	1266	396
Acc008331	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1967.000	1	1967	1967
Acc008348	LG19	yes	no	no	yes	NA	NA	NA									NA	NA			NA	184.250	8	1474	89
Acc008394	LG19	yes	no	no	NA	yes	NA	NA									NA	NA			NA	1616.000	1	1616	1616
Acc008399	LG19	no	yes	no	NA	NA	NA	4889									NA	NA			NA	114.850	20	2297	21

gene	seqname	Single copy Broc_A nan_Tilia.	Low copy Broc_Ana n_Tilia.	High copy Broc_Ana n_Tilia.	Diff_expressed_CAn_Tilia.	pos_selection_CAM-shifts	gene_C3_families_CAM	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Anas	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohydrites	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc008426	LG19	no	no	yes	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_response_factor	244.214	14	3419	56
Acc008429	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	702.500	2	1405	540
Acc008430	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	342.750	4	1371	168
Acc008431	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1538.000	1	1538	1538
Acc008441	LG19	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	191.158	19	3632	20
Acc008446	LG19	NA	NA	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	299.600	5	1498	92
Acc008452	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	287.200	5	1436	127
Acc008453	LG19	yes	no	no	NA	NA	NA	7324								yes	NA	NA			05_Chloroplasts_activity	226.600	5	1133	52
Acc008459	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	353.000	4	1412	113
Acc008460	LG19	no	no	yes	NA	NA	NA	NA									NA	PREDICTED: Oriza sativa Japonica Group homeobox-leucine zipper protein ROC7 (LOC4344614). mRNA			NA	170.250	16	2724	48
Acc008491	LG19	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	310.286	7	2172	93
Acc008510	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	371.000	4	1484	143
Acc008513	LG09	no	yes	no	yes	NA	NA	NA									NA	NA			NA	202.500	10	2025	36
Acc008564	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	778.500	2	1557	527
Acc008568	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2831.000	1	2831	2831
Acc008569	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	277.200	10	2772	128
Acc008573	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	307.000	4	1228	223
Acc008578	LG09	no	yes	no	yes	NA	NA	NA									NA	NA			NA	221.143	14	3096	59
Acc008626	LG09	no	yes	no	NA	NA	NA	NA	yes								NA	NA			00_#30_Malate_dehydrogenase	316.200	10	3162	35
Acc008682	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	238.857	14	3344	122
Acc008684	LG09	no	yes	no	NA	NA	NA	NA		yes							NA	04_Oxygen_evolving_enhancer_OEE			NA	179.600	5	898	10
Acc008685	LG09	NA	NA	no	yes	NA	NA	NA									NA	NA			NA	524.000	1	524	524
Acc008698	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	483.667	3	1451	245
Acc008715	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	398.143	7	2787	151
Acc008738	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1178.000	1	1178	1178
Acc008744	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	517.333	3	1552	308
Acc008752	LG09	yes	no	no	NA	yes	NA	NA									NA	NA			NA	508.000	9	4572	27
Acc008775	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	304.600	5	1523	121
Acc008777	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	974.500	2	1949	558
Acc008784	LG09	no	yes	no	NA	NA	NA	6859									NA	NA	yes		10_Glucogenesis_specific_dayC3	111.846	13	1454	35
Acc008787	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	543.000	3	1629	290

gene	seqname	Single copy Broc_A n.Tilla.	Low copy Broc_A n.Tilla.	High copy Broc_A n.Tilla.	Diff_expressed CAM3	pos_selection_CAM3-shifts	gene_families_CAM3	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_lunc	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Anas	circadian_metabolism_Walz2017	malate_transferase_Assimilation_Carbohydrites	flavonoid_anthocyanin_Picalirina_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0088789	LC09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2630.000	1	2630	2630
Acc0088802	LC09	yes	no	no	NA	yes	NA	NA									NA	NA			NA	250.182	11	2752	43
Acc0088830	LC09	no	yes	no	NA		NA	NA									NA	NA	yes			327.571	7	2293	90
Acc0088834	LC09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1652.000	1	1652	1652
Acc0088854	LC09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	368.200	5	1841	198
Acc0088863	LC09	no	no	yes	NA	NA	NA	NA							yes		NA	NA				330.909	11	3640	54
Acc0088872	LC09	yes	no	no	NA	NA	NA	NA					yes				NA	NA			00_Ming2019_Anthocyanin_biosynthesis	420.000	3	1260	204
Acc0088881	LC09	NA	NA	NA	NA	NA	NA	NA									STFA_F_coumaroyl-anthocyanidin 3-O-glucoside-6-O-coumaroyltransferase 2	NA			NA	554.000	2	1108	512
Acc0088882	LC09	no	yes	no	NA	NA	NA	NA									ALB_F_anthocyanin 5-aromatic acyltransferase	NA			NA	794.500	2	1589	506
Acc0088884	LC09	yes	no	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_pathway	225.500	2	451	200
Acc0088903	LC09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1430.000	1	1430	1430
Acc0088910	LC09	yes	no	no	NA	NA	NA	5958									NA	NA			NA	227.000	3	681	15
Acc0088925	LC09	NA	NA	NA	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	192.500	4	770	61
Acc0088929	LC09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2594.000	1	2594	2594
Acc0088932	LC09	no	yes	no	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	193.250	4	773	104
Acc0088947	LC09	NA	NA	NA	NA	NA	NA	NA							yes		NA	NA			04_Oxygen_evolution_enhancer_OEE	311.667	3	935	120
Acc0088949	LC09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	313.600	5	1568	125
Acc0088951	LC09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	383.750	4	1535	213
Acc0088952	LC09	no	yes	no	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	306.500	4	1226	140
Acc0088955	LC09	no	yes	no	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_response_factor	363.143	7	2542	39
Acc0088957	LC09	no	no	yes	yes	NA	NA	NA									NA	NA			NA	169.250	12	2031	50
Acc0088992	LC09	yes	no	no	NA	yes	NA	NA									NA	NA			NA	1130.000	1	1130	1130
Acc0090949	LC09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1928.000	1	1928	1928
Acc0090959	LC09	yes	no	no	NA	NA	NA	5744									NA	NA			NA	519.000	2	1038	475
Acc0090970	LC09	no	yes	no	NA	NA	yes	NA									NA	NA			NA	181.182	11	1983	83
Acc0090980	LC09	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	134.000	12	1608	57
Acc009124	LC09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1589.000	1	1589	1589
Acc009125	LC09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	221.800	5	1109	117
Acc009134	LC09	yes	no	no	NA	NA	NA	NA							yes		NA	NA			05_Chloroplasts_activity	464.000	3	1392	161
Acc009143	LC09	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	245.750	8	1966	65
Acc009185	LC22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	701.250	4	2805	394

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism	malate_transferrase	flavonoid_anthocyanin_Pitcairnia_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc009215	LG22	no	yes	no	NA	NA	NA	6216									NA	NA			NA	116.917	12	1403	30
Acc009217	LG22	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	185.217	23	4260	21
Acc009219	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	287.500	4	1150	140
Acc009231	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1313.000	2	2626	126
Acc009247	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1183.000	1	1183	1183
Acc009256	LG22	yes	no	no	NA	NA	NA	6913									NA	NA			NA	191.000	8	1528	74
Acc009272	LG22	yes	no	no	NA	NA	NA	6068									NA	NA			NA	218.167	6	1309	36
Acc009273	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	900.500	2	1801	481
Acc009274	LG22	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	682.000	3	2046	461
Acc009281	LG22	NA	NA	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	167.400	15	2511	33
Acc009282	LG22	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_pathway	181.125	8	1449	84
Acc009283	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	586.667	3	1760	262
Acc009289	LG22	yes	no	no	yes	NA	NA	NA									NA	NA			NA	125.385	13	1630	11
Acc009291	LG22	yes	no	no	yes	NA	yes	NA									NA	NA			NA	441.000	2	882	362
Acc009299	LG22	yes	no	no	NA	NA	NA	6393									NA	NA	yes		10_Glucogenesis_specific_dayC3	208.333	6	1250	123
Acc009301	LG22	NA	NA	NA	NA	NA	NA	NA								yes	NA	NA			03_VATPase	118.167	6	709	77
Acc009310	LG22	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Glycine max lysine-specific demethylase 1 (LOC10819576)_mRNA			NA	169.667	33	5599	45
Acc009322	LG22	yes	no	no	NA	NA	NA	NA									ALB_F flavone 3-O-methyltransferase 1	NA			NA	527.333	3	1582	64
Acc009346	LG22	yes	no	no	NA	NA	NA	NA		yes							NA	NA			01b_Absciscic_acid_signal	234.667	3	704	77
Acc009353	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	940.000	2	1880	603
Acc009366	LG22	yes	no	no	NA	yes	NA	NA									NA	NA			NA	311.400	5	1557	167
Acc009370	LG22	no	no	yes	NA	NA	NA	NA									NA	PREDICTED: Solanum tuberosum heat shock cognate protein 80-like (LOC102000008)_mRNA			NA	214.600	15	3219	80
Acc009371	LG22	yes	no	no	yes	NA	NA	NA									NA	NA			NA	530.000	1	530	530
Acc009385	LG22	yes	no	no	yes	NA	NA	NA									NA	NA			NA	341.500	2	683	254
Acc009388	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	332.250	4	1329	198
Acc009392	LG22	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	151.214	14	2117	44
Acc009409	LG22	yes	no	no	NA	NA	NA	6572									NA	NA	yes		10_Glucogenesis_specific_dayC3	101.077	13	1314	44
Acc009425	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1007.500	2	2015	966
Acc009446	LG14	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1077.000	2	2154	616
Acc009453	LG14	yes	no	no	NA	NA	NA	NA									NA	NA			NA	531.750	4	2127	365

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilla	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	Assimilation_Carbohyd	flavonoid_anthocyanin_Psicaria	pos_sel_portuligo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc009463	LG19	no	yes	no	yes	yes	NA	NA	NA									NA	NA	NA		NA	NA	323.000	5	1615	21
Acc009469	LG01	no	yes	no	NA	NA	NA	NA	4942									NA	NA	NA		NA	NA	205.375	8	1643	35
Acc009474	LG01	no	yes	no	NA	NA	NA	NA	6660									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	104.333	9	939	58
Acc009477	LG01	no	yes	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	737.000	1	737	737
Acc009484	LG01	no	yes	no	yes	NA	NA	NA	NA									NA	NA	NA		NA	NA	324.250	4	1297	67
Acc009499	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	685.000	2	1370	622
Acc009508	LG01	no	no	yes	NA	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	223.286	14	3126	116
Acc009511	LG01	yes	no	no	NA	NA	yes	NA	NA									NA	NA	NA		NA	NA	548.000	3	1644	238
Acc009513	LG01	no	yes	no	NA	NA	NA	NA	NA			yes						NA	NA	NA			06_Drought_resistance_genes_LEA	182.000	1	182	182
Acc009514	LG01	no	yes	no	NA	NA	NA	NA	NA			yes						NA	NA	NA			06_Drought_resistance_genes_LEA	125.000	2	250	13
Acc009532	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	286.833	6	1721	155
Acc009549	LG01	yes	no	no	NA	NA	yes	NA	NA									NA	NA	NA		NA	NA	2039.000	1	2039	2039
Acc009551	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	173.400	5	867	81
Acc009559	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	302.200	5	1511	127
Acc009566	LG01	yes	no	no	NA	NA	NA	NA	6110									NA	NA	NA		NA	NA	566.000	1	566	566
Acc009569	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	542.200	5	2711	162
Acc009577	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	368.429	7	2579	115
Acc009587	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	466.000	3	1398	238
Acc009588	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA	yes		11_Glycolysis_night_C4_malate_pathway	123.824	17	2105	31
Acc009592	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	794.000	2	1588	609
Acc009599	LG01	no	yes	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	191.750	4	767	16
Acc009600	LG01	no	yes	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	641.000	1	641	641
Acc009612	LG01	yes	no	no	NA	NA	NA	NA	NA							yes		NA	NA	NA			06_#W_CircadianOscillator_cycling_DOF_F1	862.500	2	1725	260
Acc009627	LG01	no	yes	no	NA	NA	NA	NA	NA			yes						NA	NA	NA			06_Drought_resistance_genes_LEA	428.000	2	856	162
Acc009639	LG01	no	yes	no	NA	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	1033.000	2	2066	685
Acc009644	LG01	no	yes	no	NA	NA	NA	NA	NA									NA			yes		10_Glucogenesis_specific_dayC3	129.600	15	1944	55
Acc009645	LG01	no	yes	no	yes	NA	NA	NA	NA									NA	NA	NA		NA	NA	353.167	24	8476	72
Acc009647	LG01	yes	no	no	NA	NA	NA	NA	6601									NA	NA	NA		NA	NA	93.800	10	938	14
Acc009654	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	388.800	5	1944	152
Acc009662	LG01	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA		NA	NA	404.500	4	1618	125

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Wal2017	malate_transferrase	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc009673	LG01	yes	no	no	NA	NA	NA	NA									NA	PREDICTED: Beta vulgaris subsp. vulgaris histone acetyltransferase GCN5 (LOC104894583). mRNA			NA	161.846	13	2104	43
Acc009686	LG01	yes	no	no	NA	NA	NA	NA									NA				NA	656.500	2	1313	344
Acc009689	LG01	yes	no	no	NA	NA	NA	NA								yes	NA				09_Circadian_clock	124.100	10	1241	42
Acc009713	LG01	yes	no	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	305.000	3	915	71
Acc009744	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	501.667	3	1505	131
Acc009779	LG10	NA	NA	NA	NA	NA	NA	NA									NA				01d_Auxin_reponse_factor	433.333	3	1300	138
Acc009780	LG10	no	no	yes	NA	NA	NA	NA									NA	PREDICTED: Glycine max protein argonaute 1-like (LOC100778754). mRNA			NA	156.000	20	3120	63
Acc009790	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	726.500	2	1453	488
Acc009794	LG10	yes	no	no	yes	NA	NA	NA									NA				NA	259.500	6	1557	161
Acc009801	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	554.667	3	1664	264
Acc009807	LG10	no	yes	no	NA	NA	NA	NA									STA_L_flavonoid_glucosyltransferase				NA	458.000	4	1832	113
Acc009813	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	529.500	6	3177	116
Acc009822	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	632.667	3	1898	175
Acc009889	LG10	yes	no	no	NA	NA	NA	6528									NA				NA	269.750	4	1079	112
Acc009903	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	388.333	3	1165	202
Acc009914	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	351.000	6	2106	198
Acc009924	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	596.500	4	2386	200
Acc009949	LG10	no	no	yes	NA	NA	NA	NA									NA		yes		11_Glycolysis_right_C4_malate_pathway	156.929	14	2197	39
Acc009967	LG10	NA	NA	NA	NA	NA	NA	NA	yes								NA				00_#32_Malic_enzyme	144.850	20	2897	47
Acc009980	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	1112.000	1	1112	1112
Acc009989	LG10	no	yes	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	1258.000	2	2516	796
Acc010002	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	1415.000	1	1415	1415
Acc010013	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	845.000	2	1690	640
Acc010025	LG10	no	yes	no	NA	NA	NA	NA				yes					NA	PREDICTED: Beta vulgaris subsp. vulgaris phosphoenolpyruvate carboxylase 1 (LOC104903486). mRNA			00_#MB_PEP_C10_Glucogenesis_specific_dayC3	348.000	10	3480	84
Acc010028	LG10	no	yes	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	173.000	7	1211	69
Acc010030	LG10	yes	no	no	NA	NA	NA	NA									NA		yes		11_Glycolysis_right_C4_malate_pathway	425.667	3	1277	345
Acc010034	LG10	yes	no	no	NA	NA	NA	NA									NA				NA	1460.000	1	1460	1460
Acc010086	LG25	NA	NA	NA	yes	NA	NA	NA									NA				NA	125.714	7	880	12
Acc010087	LG25	yes	no	no	NA	NA	NA	NA									NA				NA	1985.000	1	1985	1985
Acc010088	LG25	yes	no	no	NA	NA	NA	NA									NA				NA	2285.000	1	2285	2285

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohyd	flavonoid_anthocyanin_Picalirina_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0010113	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1433.000	1	1433	1433
Acc0010116	LG25	yes	no	no	NA	NA	6412										NA	NA			NA	234.500	4	938	126
Acc0010122	LG25	no	yes	no	NA	NA	yes	NA									NA	NA			NA	163.250	12	1959	83
Acc0010134	LG25	no	yes	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	286.667	3	860	107
Acc0010136	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	279.750	4	1119	155
Acc0010148	LG25	no	yes	no	NA	NA	NA	NA									NA	Populus trichocarpa hypothetical protein (POPTR_0002s25930g) mRNA_complete cds			NA	208.750	4	835	41
Acc0010160	LG25	no	yes	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	186.750	4	747	86
Acc0010172	LG25	yes	no	no	NA	NA	6458										NA	NA			NA	101.632	19	1931	20
Acc0010193	LG25	no	yes	no	NA	NA	7333										NA	NA			NA	704.000	1	704	704
Acc0010197	LG25	no	yes	no	NA	NA	5562										NA	NA			NA	285.750	4	1063	62
Acc0010215	LG25	yes	no	no	NA	NA	NA	NA									NA	11_Glycolysis_right_C4_malate_pathway	yes			313.750	4	1255	247
Acc0010217	LG25	yes	no	no	NA	NA	NA	NA									ALB_F_malonyl-coenzyme a:anthocyanin 3-o-glucoside-6-o-malonyltransferase	NA			NA	1358.000	1	1358	1358
Acc0010230	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1184.000	1	1184	1184
Acc0010232	LG25	yes	no	no	NA	NA	yes	NA									NA	PREDICTED: Beta vulgaris subsp. vulgaris malate dehydrogenase (NAOPL_chloroplastic (LOC104884197) mRNA			00_#428_Malate_dehydrogenase	117.333	15	1760	19
Acc0010251	LG25	no	yes	no	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	304.333	6	1826	30
Acc0010261	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1169.000	1	1169	1169
Acc0010278	LG25	yes	no	no	NA	NA	NA	NA						yes			NA	00_#W_CircadianOscillator_WD_repeat-containing_protein_RUP2				1253.000	1	1253	1253
Acc0010280	LG25	yes	no	no	NA	NA	NA	NA						yes			NA	00_#W_ProtonPump_VHA-F_03_VATPase				171.600	5	858	77
Acc0010296	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	867.500	2	1735	478
Acc0010298	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	282.250	4	1129	120
Acc0010300	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	456.667	3	1370	153
Acc0010310	LG25	no	no	yes	yes	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	557.500	4	2230	229
Acc0010312	LG25	no	no	yes	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	364.000	3	1092	276
Acc0010315	LG25	no	yes	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	369.500	2	739	353
Acc0010319	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	892.500	2	1785	644
Acc0010322	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	419.500	4	1678	162
Acc0010340	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	640.667	3	1922	119
Acc0010341	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	306.250	4	1225	203
Acc0010367	LG03	no	yes	no	yes	NA	NA	NA									NA	NA			NA	170.731	26	4439	51
Acc0010377	LG03	no	yes	no	yes	NA	NA	NA									NA	NA			NA	237.700	10	2377	80

gene	seqname	Single copy Broc_A nan_Tilia.	Low copy Broc_Ana n_Tilia.	High copy Broc_Ana an_Tilia.	Diff_expressed_CA an_Tilia.	pos_selection_CAM _shifts	gene_families_CAM C3	Angiosperm_gene	CAM- related_Acomosus- Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana nas	circadian_metabolics m_Walz2017	malate_transferase Assimilation_Carbohy drates	flavonoid_anthocya nin_Picalirnia_PS	pos_sel_portulugo	glucogenesis_glycol ysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc010381	LG03	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	107.250	12	1287	44
Acc010390	LG03	no	yes	no	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	399.333	3	1198	140
Acc010438	LG03	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_p athway	361.000	5	1805	51
Acc010466	LG03	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2201.000	1	2201	2201
Acc010544	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	436.333	6	2618	167
Acc010549	LG07	yes	no	no	NA	NA	NA	5974									NA	NA			NA	126.444	9	1138	39
Acc010557	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1203.000	2	2406	925
Acc010572	LG07	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	161.875	8	1295	83
Acc010579	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1979.000	1	1979	1979
Acc010590	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	562.000	2	1124	488
Acc010614	LG07	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	491.000	3	1473	177
Acc010619	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	458.250	4	1833	247
Acc010620	LG07	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	345.800	5	1729	58
Acc010624	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2342.000	1	2342	2342
Acc010634	LG07	no	no	yes	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	704.500	2	1409	420
Acc010638	LG07	no	yes	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	383.000	2	766	278
Acc010641	LG07	yes	no	no	NA	NA	yes	NA									NA	NA			NA	447.000	2	894	362
Acc010642	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1658.000	1	1658	1658
Acc010678	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	557.500	2	1115	239
Acc010696	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	350.400	5	1752	158
Acc010710	LG10	yes	no	no	NA	NA	NA	NA									ALB_F_dihydroflavonol 4- reductase	NA	yes		10_Glucogenesis_specific_dayC3	259.600	5	1298	169
Acc010716	LG10	no	yes	no	NA	NA	NA	NA									NA	NA			NA	113.250	12	1359	15
Acc010721	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	948.500	2	1897	351
Acc010725	LG10	no	yes	no	NA	NA	NA	NA							yes		NA	NA			00_#W_MalateTransporter_ALMT_10_Glucogenesis_specific_dayC3	275.167	6	1651	139
Acc010731	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	639.000	3	1917	125
Acc010740	LG10	no	yes	no	yes	NA	NA	NA									NA	NA			NA	501.000	2	1002	461
Acc010747	LG10	yes	no	no	NA	NA	NA	6320									NA	NA			NA	301.625	8	2413	125
Acc010748	LG10	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	237.667	6	1426	23
Acc010756	LG10	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_p athway	430.000	6	2580	31
Acc010770	LG10	no	yes	no	NA	NA	NA	6689									NA	NA			NA	126.182	11	1388	45
Acc010774	LG10	yes	no	no	NA	NA	NA	6303									NA	NA			NA	162.375	8	1299	20

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate transferase	flavonoid_anthocyanin_Psicartria_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0010779	LG10	no	yes	no	yes	NA	NA	NA									NA	NA			NA	295.500	2	591	286
Acc0010780	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	571.000	2	1142	523
Acc0010787	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	877.500	2	1755	654
Acc0010792	LG10	yes	no	no	NA	NA	NA	6038									NA	NA			NA	183.900	10	1839	79
Acc0010811	LG10	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	184.000	14	2576	60
Acc0010816	LG10	no	yes	no	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	349.667	3	1049	212
Acc0010827	LG10	no	yes	no	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	336.500	2	673	174
Acc0010835	LG10	yes	no	no	NA	NA	NA	5919									NA	NA			NA	404.750	4	1619	85
Acc0010847	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1133.000	1	1133	1133
Acc0010877	LG04	NA	NA	NA	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	362.333	3	1087	225
Acc0010980	LG04	yes	no	no	NA	NA	NA	NA							yes		NA	NA			04_Oxygen_evolution_enhancer_OEE	93.500	6	561	71
Acc0010982	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	596.667	3	1790	395
Acc0011000	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	686.000	2	1372	583
Acc0011010	LG04	yes	no	no	NA	NA	NA	5926									NA	NA			NA	118.778	9	1069	58
Acc0011011	LG04	NA	NA	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	226.286	7	1584	50
Acc0011012	LG04	no	yes	no	NA	NA	NA	NA							yes		NA	NA			00_#W_CircadianOscillator_Pseudoreponse_regulator7	254.636	11	2801	109
Acc0011025	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	737.500	2	1475	647
Acc0011032	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	349.667	6	2088	191
Acc0011047	LG04	no	yes	no	NA	NA	NA	NA							yes		NA	NA			00_#W_StomatalMovement_Phospholipase_C	200.294	17	3405	86
Acc0011051	LG04	no	no	yes	NA	NA	NA	NA									ALB_F_anthocyanidin synthase	NA			NA	256.250	4	1025	157
Acc0011056	LG04	yes	no	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_pathway	325.000	3	975	230
Acc0011067	LG04	no	yes	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	141.500	2	283	114
Acc0011075	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	741.000	2	1482	617
Acc0011098	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	845.000	2	1690	605
Acc0011105	LG04	no	yes	no	yes	NA	NA	NA									NA	NA			NA	221.600	10	2216	59
Acc0011162	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	293.500	6	1761	125
Acc0011172	LG04	yes	no	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	629.000	1	629	629
Acc0011192	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	823.500	2	1647	805
Acc0011198	LG01	yes	no	no	NA	NA	NA	NA	yes								NA	NA			01d_Auxin_response_factor	1018.500	2	2037	965
Acc0011204	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	536.333	3	1609	277
Acc0011238	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1625.000	1	1625	1625

gene	seqname	Single copy Broc-A	Low copy Broc_Ana	High copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_families_CAM	Angiosperm_gene	CAM-related_Acomus	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism	malate transferase	Assimilation_Carbohyd	flavonoid_anthocyanin_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0011239	LG01	yes	no	no	no	NA	NA	NA	6961										NA	NA			NA	156.667	9	1410	57
Ac0011246	LG01	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	458.333	3	1375	246
Ac0011261	LG01	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	437.667	3	1313	276
Ac0011262	LG01	no	yes	no	no	yes	NA	NA	NA										NA	NA			NA	1110.500	2	2221	981
Ac0011283	LG01	yes	no	no	no	NA	NA	NA	6979										NA	NA			NA	141.778	9	1276	50
Ac0011299	LG01	no	yes	no	no	NA	yes	NA	NA										NA	NA			NA	1670.000	1	1670	1670
Ac0011300	LG01	no	yes	no	no	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	252.000	6	1512	100
Ac0011306	LG01	yes	no	no	no	yes	NA	NA	NA										NA	NA			NA	368.000	5	1840	141
Ac0011332	LG01	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	722.000	2	1444	604
Ac0011338	LG01	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	491.333	3	1474	151
Ac0011346	LG01	yes	no	no	no	NA	NA	NA	5791										NA	NA			NA	104.333	12	1252	46
Ac0011349	LG01	yes	no	no	no	NA	NA	NA	5034										NA	NA			NA	205.000	6	1230	69
Ac0011352	LG01	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	367.000	3	1101	210
Ac0011371	LG01	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	279.800	5	1399	133
Ac0011378	LG01	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	1760.000	1	1760	1760
Ac0011401	LG01	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	671.500	4	2686	241
Ac0011462	LG06	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	421.333	3	1264	249
Ac0011464	LG06	no	no	yes	no	NA	NA	NA	NA		yes								NA	NA			01d_Auxin_response_factor	219.933	15	3299	56
Ac0011466	LG06	no	yes	no	no	NA	NA	NA	NA										NA	PREDICTED: Solanum tuberosum protein HUA ENHANCER 2-like (LOC102587942)_transcript variant X3. mRNA		NA	212.529	17	3613	68	
Ac0011496	LG06	yes	no	no	no	yes	NA	NA	NA								yes		NA	NA			03_VATPase	279.250	4	1117	77
Ac0011515	LG06	yes	no	no	no	NA	NA	NA	6404										NA	NA			NA	139.333	9	1254	49
Ac0011516	LG06	no	yes	no	no	yes	NA	NA	NA										NA	NA			NA	297.222	9	2675	48
Ac0011527	LG06	no	no	yes	no	NA	yes	NA	NA										NA	NA			NA	251.947	19	4787	5
Ac0011528	LG06	no	yes	no	no	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	223.000	8	1784	62
Ac0011530	LG06	yes	no	no	no	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	516.750	4	2067	221
Ac0011531	LG06	yes	no	no	no	NA	NA	NA	NA										NA	NA	yes		11_Glycolysis_right_C4_malate_p athway	320.000	4	1280	247
Ac0011560	LG06	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	1109.000	2	2218	344
Ac0011590	LG06	yes	no	no	no	NA	NA	NA	6406										NA	NA			NA	138.917	12	1667	59
Ac0011609	LG06	no	yes	no	no	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	128.231	13	1667	70
Ac0011644	LG08	yes	no	no	no	NA	NA	NA	NA										NA	NA			NA	334.500	4	1338	112
Ac0011651	LG08	no	yes	no	no	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	203.556	9	1832	74

gene	seqname	Single copy Broc_A	Low copy Broc_A	High copy Broc_A	Highly expressed	Diff. expressed	pos. selection	gene families	Angiosperm gene	CAM-related	stomata func.	aquaporin reg.	drought res.	markers	anthocyanin	circadian	malate transferase	Assimilation	hydrates	flavonoid anthocyanin	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0011671	LG08	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	1868.000	1	1868	1868
Acc0011681	LG08	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	463.667	3	1391	208
Acc0011700	LG08	NA	NA	NA	yes	NA	NA	NA	NA										NA	NA				NA	201.500	6	1209	27
Acc0011705	LG08	NA	NA	NA	NA	NA	NA	NA	NA			yes							NA	NA			02_Aquaporins	333.000	3	999	190	
Acc0011709	LG08	NA	NA	NA	NA	NA	NA	NA	NA									ALB_L_malonyl-coenzyme a: anthocyanin 3-o-glucoside-6 -o-malonyltransferase	NA	NA				NA	526.000	3	1578	367
Acc0011710	LG08	yes	no	no	NA	NA	NA	NA	NA										NA	NA	yes			11_Glycolysis_night_C4_malate_p_atway	271.333	9	2442	21
Acc0011720	LG08	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	1083.000	2	2166	625
Acc0011724	LG08	yes	no	no	NA	NA	NA	NA	6785										NA	NA				NA	178.833	6	1073	51
Acc0011740	LG08	yes	no	no	NA	NA	NA	NA	6398										NA	NA				NA	233.400	5	1167	77
Acc0011789	LG08	yes	no	no	NA	NA	NA	NA	NA										NA	NA	yes			11_Glycolysis_night_C4_malate_p_atway	160.778	9	1447	63
Acc0011792	LG08	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	2186.000	1	2186	2186
Acc0011813	LG08	yes	no	no	NA	NA	NA	NA	NA										NA	PREDICTED: Virus vinifera putative tRNA pseudouridine synthase Pts10 (LOC100242087)_transcript variant X1_mRNA				NA	146.000	15	2190	55
Acc0011817	LG08	no	yes	no	yes	NA	NA	NA	NA										NA	NA				NA	290.250	4	1161	55
Acc0011837	LG08	no	no	yes	yes	NA	NA	NA	NA										NA	NA				NA	250.750	12	3009	60
Acc0011838	LG08	NA	NA	NA	yes	NA	NA	NA	NA										NA	NA				NA	286.400	5	1432	189
Acc0011860	LG08	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	429.750	4	1719	194
Acc0011881	LG03	yes	no	no	NA	yes	NA	NA	NA										NA	NA				NA	612.500	6	3675	161
Acc0011886	LG03	no	no	yes	NA	yes	NA	NA	NA		yes								NA	NA				01d_Auxin_reponse_factor	240.750	16	3852	56
Acc0011896	LG03	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	417.667	3	1253	259
Acc0011903	LG03	yes	no	no	NA	NA	NA	NA	NA							yes			NA	NA				00_#W_CircadianOscillator_Reveille1	331.571	7	2321	61
Acc0011912	LG03	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	369.333	3	1108	200
Acc0011916	LG03	no	yes	no	NA	NA	NA	NA	NA							yes			NA	NA				00_#W_HexoseTransporter_MSSP-like_transporter	620.500	6	3723	120
Acc0011917	LG03	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	283.200	5	1416	166
Acc0011930	LG03	no	no	yes	NA	NA	NA	NA	NA										NA	NA				NA	162.478	23	3737	31
Acc0011960	LG03	no	yes	no	yes	NA	NA	NA	NA										NA	NA				NA	159.455	11	1754	39
Acc0011969	LG03	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	665.000	2	1330	558
Acc0011974	LG03	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	780.000	3	2340	425
Acc0011976	LG03	no	yes	no	NA	NA	NA	NA	NA				yes						NA	NA				06_Drought_resistance_genes_LEA	440.400	5	2202	84
Acc0011977	LG03	yes	no	no	NA	NA	NA	NA	NA										NA	NA				NA	199.571	7	1397	112
Acc0011978	LG03	no	yes	no	NA	NA	NA	NA	NA				yes						NA	NA				06_Drought_resistance_genes_LEA	485.000	2	970	213

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_lunc	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	flavonoid_anthocyanin_Picalirnia_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0011994	LG03	no	no	yes	yes	NA	NA	NA									NA	NA			NA	248.625	8	1989	99
Ac0012009	LG03	no	yes	no	NA	NA	NA	NA			yes	yes					NA	NA			06_Drought_resistance_genes_LEA	617.000	1	617	617
Ac0012014	LG03	no	yes	no	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	381.250	4	1445	140
Ac0012041	LG03	yes	no	no	NA	yes	NA	NA									NA	NA			NA	411.000	4	1644	159
Ac0012049	LG03	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	594.000	4	2376	65
Ac0012057	LG03	no	yes	no	yes	NA	NA	NA									NA	NA			NA	608.000	5	3040	217
Ac0012085	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1152.000	2	2304	1055
Ac0012095	LG02	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	1322.000	1	1322	1322
Ac0012114	LG02	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	146.231	13	1901	70
Ac0012160	LG02	yes	no	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	244.000	3	732	80
Ac0012179	LG02	no	yes	no	NA	NA	NA	6649									NA	NA			NA	121.167	6	727	6
Ac0012217	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	443.857	7	3107	139
Ac0012218	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	442.000	5	2210	190
Ac0012234	LG02	no	no	yes	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_reponse_factor	265.867	15	3988	56
Ac0012256	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	302.250	8	2418	116
Ac0012270	LG01	no	yes	no	NA	NA	NA	NA					yes				NA	NA			00_#MB_XDH	335.875	16	5374	31
Ac0012284	LG01	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	102.250	4	409	71
Ac0012292	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	415.750	4	1663	119
Ac0012294	LG01	yes	no	no	NA	NA	NA	5280									NA	NA			NA	178.182	11	1960	23
Ac0012308	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	380.400	5	1902	149
Ac0012311	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	311.500	4	1246	138
Ac0012319	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1004.000	2	2008	957
Ac0012352	LG01	yes	no	no	NA	NA	NA	6559									NA	NA			NA	201.000	7	1407	76
Ac0012380	LG01	no	no	yes	NA	NA	NA	NA									NA	NA			NA	1083.333	3	3280	438
Ac0012389	LG01	yes	no	no	NA	NA	NA	7367									NA	NA			NA	148.500	4	594	89
Ac0012398	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	575.250	4	2301	133
Ac0012400	LG01	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	381.250	8	3050	76
Ac0012404	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1370.000	1	1370	1370
Ac0012412	LG01	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	136.625	8	1093	9
Ac0012435	LG01	no	yes	no	NA	yes	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	208.500	10	2085	94
Ac0012441	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	444.750	4	1779	144

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilia	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transfere	Assimilation_Carbohyd	flavonoid_anthocyanin_Picalirina_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc012460	LG13	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	698.333	3	2095	484
Acc012463	LG13	no	yes	no	NA	NA	yes	NA	NA										NA	NA			NA	181.529	17	3086	68
Acc012466	LG13	yes	no	no	yes	NA	NA	NA	NA										NA	NA			NA	221.333	6	1328	27
Acc012471	LG13	no	no	yes	NA	NA	NA	NA	NA										ALB_F_leucoanthocyanidin dioxygenase	NA	yes		11_Glycolysis_night_C4_malate_p	423.667	3	1271	330
Acc012481	LG13	no	yes	no	NA	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	283.857	7	1987	55
Acc012483	LG13	yes	no	no	NA	NA	NA	NA	5894										NA	NA			NA	175.833	6	1055	20
Acc012488	LG13	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	338.500	6	2031	126
Acc012491	LG13	yes	no	no	NA	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	225.667	6	1354	162
Acc012511	LG13	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	1514.000	1	1514	1514
Acc012530	LG13	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	1197.667	3	3593	615
Acc012547	LG13	yes	no	no	NA	NA	NA	NA	NA										ALB_F_chalcone isomerase	NA			NA	210.000	4	840	83
Acc012577	LG13	yes	no	no	NA	NA	NA	NA	NA			yes							NA	NA			06_Drought_resistance_genes_LEA	315.500	2	631	22
Acc012622	LG13	yes	no	no	NA	NA	NA	NA	NA				yes						NA	NA			NA	253.600	5	1268	135
Acc012641	LG13	yes	no	no	NA	NA	NA	NA	NA										NA	NA			06_Drought_resistance_genes_LEA	446.000	2	892	4
Acc012644	LG13	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	290.833	6	1745	163
Acc012651	LG05	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	578.500	2	1157	364
Acc012700	LG05	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	1937.000	1	1937	1937
Acc012704	LG05	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	1229.000	1	1229	1229
Acc012715	LG05	yes	no	no	yes	NA	NA	NA	NA										NA	NA			NA	303.750	8	2430	65
Acc012740	LG05	no	yes	no	NA	NA	NA	NA	6544										NA	NA			NA	201.176	17	3420	68
Acc012743	LG05	yes	no	no	NA	NA	yes	NA	NA										PREDICTED: Glycine max uridine-cytidine kinase C-like (LOC100781286)_mRNA	NA			NA	175.615	13	2283	3
Acc012745	LG05	no	yes	no	NA	NA	NA	NA	NA										ALB_L_flavonoid glucosyl-transferase	NA			NA	596.667	3	1790	257
Acc012760	LG05	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	2651.000	1	2651	2651
Acc012767	LG05	no	yes	no	NA	NA	NA	NA	NA										NA	PREDICTED: Solanum tuberosum ATPase family AAA domain containing protein 3C (LOC102578202)_mRNA			NA	267.125	8	2137	90
Acc012772	LG05	no	yes	no	NA	NA	NA	NA	NA										ALB_F_isoflavone reductase	NA	yes		10_Glucogenesis_specific_dayC3	193.250	12	2319	41
Acc012775	LG05	yes	no	no	NA	NA	NA	NA	6139										NA	NA			NA	292.500	6	1755	131
Acc012776	LG05	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	505.000	4	2020	202
Acc012793	LG05	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	689.000	2	1378	139
Acc012797	LG05	NA	NA	NA	NA	NA	NA	NA	NA										NA	10_Glucogenesis_specific_dayC3	yes			333.333	3	1000	183
Acc012808	LG03	yes	no	no	NA	NA	NA	NA	6387										NA	NA			NA	1088.000	1	1088	1088
Acc012817	LG03	yes	no	no	NA	NA	NA	NA	NA										NA	NA			NA	762.333	3	2287	261

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0012825	LG03	NA	NA	NA	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	348.000	3	1044	222
Ac0012837	LG03	no	yes	no	NA	NA	NA	NA									NA	Glycine max seed maturation protein PM37 (PM37)_mRNA			NA	236.500	8	1892	133
Ac0012850	LG03	yes	no	no	NA	NA	NA	5536									NA	NA			NA	150.167	6	901	57
Ac0012862	LG03	yes	no	no	NA	NA	NA	NA									NA	NA			NA	455.800	5	2279	182
Ac0012869	LG03	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	371.667	6	2230	44
Ac0012915	LG03	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	154.143	14	2158	59
Ac0012924	LG03	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	255.500	6	1533	127
Ac0012927	LG03	NA	NA	NA	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	329.000	1	329	329
Ac0012941	LG03	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Glycine max cullin-associated NEDD8-dissociated protein 1 (LOC100816200)_mRNA			NA	147.767	30	4433	50
Ac0012946	LG03	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	185.636	11	1822	55
Ac0012988	LG25	no	no	yes	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_reponse_factor	244.643	14	3425	56
Ac0012992	LG25	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	250.800	10	2508	85
Ac0013001	LG25	no	no	yes	NA	NA	NA	NA									NA	PREDICTED: Glycine max protein argonaute 1-like (LOC100778754)_mRNA			NA	156.000	22	3432	31
Ac0013003	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1838.000	1	1838	1838
Ac0013014	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1103.000	1	1103	1103
Ac0013015	LG25	yes	no	no	NA	NA	NA	5702									NA	NA			NA	169.800	10	1698	63
Ac0013023	LG25	no	yes	no	yes	NA	NA	NA									NA	NA			NA	556.200	5	2781	71
Ac0013026	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	777.500	2	1555	753
Ac0013034	LG25	no	no	yes	yes	NA	NA	NA									NA	NA			NA	283.125	8	2265	60
Ac0013037	LG25	yes	no	no	NA	NA	NA	6848									NA	NA			NA	161.667	6	970	62
Ac0013038	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	658.667	3	1976	197
Ac0013074	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1505.000	1	1505	1505
Ac0013080	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1340.000	1	1340	1340
Ac0013083	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1442.000	1	1442	1442
Ac0013096	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	687.500	2	1375	389
Ac0013100	LG25	yes	no	no	yes	NA	NA	NA									NA	NA			NA	115.455	11	1270	33
Ac0013121	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1325.000	2	2650	1148
Ac0013126	LG25	yes	no	no	NA	NA	NA	5960									NA	NA			NA	200.000	8	1600	69
Ac0013128	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	480.333	3	1441	302
Ac0013134	LG25	yes	no	no	NA	NA	NA	NA									NA	NA			NA	646.667	3	1940	422
Ac0013158	LG24	yes	no	no	NA	NA	NA	5123									NA	NA			NA	178.111	9	1603	60

gene	seqname	Single copy Broc_A nan_Tilla.	Low copy Broc_Ana n_Tilla.	High copy Broc_Ana	Diff_expressed CA an_Tilla.	pos_selection_CAM _shifts	gene_C3 families_CAM	Angiosperm_gene	CAM- related_Acomnosus- Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana nas	circadian_metabolis m_Walz2017	malate transferase Assimilation_Carbohy drates	flavonoid_anthocya nin_Picalirria_PS	pos_sel_portulugo	glucogenesis_glycol ysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0013178	LG24	no	yes	no	yes	NA	NA	NA									NA	NA			NA	204.333	3	613	86
Ac0013202	LG24	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_p athway	113.158	19	2150	44
Ac0013204	LG24	NA	NA	NA	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	197.000	1	197	197
Ac0013205	LG24	NA	NA	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	249.500	4	998	26
Ac0013210	LG24	no	yes	no	NA	yes	NA	NA									NA	NA			NA	276.000	12	3312	29
Ac0013213	LG24	yes	no	no	NA	NA	NA	NA									NA	NA			NA	662.500	2	1325	402
Ac0013214	LG24	yes	no	no	NA	yes	NA	NA									NA	NA			NA	2792.000	1	2792	2792
Ac0013224	LG24	yes	no	no	NA	NA	NA	NA									NA	NA			NA	469.400	5	2347	175
Ac0013227	LG24	NA	NA	NA	NA	NA	NA	NA			yes						NA	NA			01d_Auxin_reponse_factor	734.000	1	734	734
Ac0013253	LG24	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	719.333	3	2158	186
Ac0013267	LG24	no	yes	no	NA	NA	NA	NA									NA	NA			NA	266.267	15	3994	104
Ac0013270	LG24	yes	no	no	NA	NA	NA	NA									NA	NA			NA	558.333	3	1675	275
Ac0013283	LG24	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	155.333	9	1398	13
Ac0013293	LG24	no	yes	no	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	282.000	5	1410	61
Ac0013305	LG24	yes	no	no	NA	NA	NA	NA									NA	NA			NA	350.000	5	1750	114
Ac0013309	LG24	no	yes	no	NA	NA	NA	NA									NA	NA			NA	106.000	3	318	77
Ac0013351	LG15	yes	no	no	NA	NA	NA	NA									NA	NA			NA	293.125	8	2345	111
Ac0013353	LG15	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1916.000	1	1916	1916
Ac0013355	LG15	yes	no	no	NA	NA	NA	NA									NA	NA			NA	371.333	3	1114	149
Ac0013390	LG15	NA	NA	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	197.000	3	591	100
Ac0013412	LG15	no	no	yes	yes	NA	NA	NA						yes			NA	PREDICTED: Glycine max histidine kinase 2- like (LOC100781659)_transcript variant X3_ mRNA			00_#W_StomataMovement_Histidi ne_Kinase_HK5	312.455	11	3437	51
Ac0013414	LG15	yes	no	no	NA	NA	NA	NA									NA	NA			NA	709.250	4	2837	128
Ac0013421	LG15	no	yes	no	NA	NA	NA	NA						yes			NA	00_#W_StomataMovement_Respo nseRegulator2			NA	390.143	7	2731	76
Ac0013473	LG15	yes	no	no	NA	NA	NA	NA									ALB_F_leucoanthocyanidin reductase	NA			NA	213.200	5	1066	140
Ac0013474	LG15	yes	no	no	NA	NA	NA	NA									NA	NA			NA	813.333	3	2440	192
Ac0013489	LG20	yes	no	no	NA	NA	NA	NA									NA	NA			NA	510.667	3	1532	241
Ac0013510	LG20	yes	no	no	NA	NA	NA	NA									NA	NA			NA	408.500	4	1634	119
Ac0013517	LG20	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	133.909	11	1473	37
Ac0013524	LG20	no	yes	no	NA	NA	NA	5454									NA	NA			NA	110.083	24	2642	35
Ac0013533	LG20	yes	no	no	NA	NA	NA	NA									NA	NA			NA	972.000	3	2916	541

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff. expressed CA	pos_selection CAM	gene_selection CAM	Angiosperm_gene	CAM-related Miao2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism	malate/transferase	Assimilation_Carbohydrates	flavonoid_anthocyanin_Picalirina_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0013536	LG20	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	948.000	3	2844	871
Acc0013538	LG20	yes	no	no	NA	NA	NA	4806									NA	NA	NA			NA	198.333	6	1190	59
Acc0013550	LG20	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	404.000	4	1616	303
Acc0013584	LG20	NA	NA	NA	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	141.500	6	849	41
Acc0013587	LG20	no	yes	no	yes	NA	NA	NA									NA	NA	NA			NA	173.000	10	1730	56
Acc0013594	LG20	no	yes	no	NA	NA	NA	NA									NA	NA	NA	yes		11_Glycolysis_night_C4_malate_pathway	155.200	10	1552	68
Acc0013623	LG13	yes	no	no	NA	NA	NA	NA			yes						NA	NA	NA			02_Aquaporins	417.667	3	1253	224
Acc0013628	LG13	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	655.500	2	1311	263
Acc0013633	LG13	no	yes	no	NA	NA	NA	NA									NA		PREDICTED: Beta vulgaris subsp. vulgaris glucose-1-phosphate adenylyltransferase large subunit 3_chloroplast/amyloplastic (LOC104883901)_transcript variant X1_mRNA	yes		10_Glucogenesis_specific_dayC3	96.714	14	1354	55
Acc0013637	LG13	no	yes	no	NA	NA	NA	NA					yes				NA	NA	NA			00_#MB_Ph1YA	974.750	4	3899	290
Acc0013642	LG13	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	1142.000	2	2284	236
Acc0013643	LG13	yes	no	no	NA	NA	NA	NA				yes					NA	NA	NA			00_#W_CircadianOscillator_TF_P1_F3	302.800	10	3028	48
Acc0013654	LG13	no	yes	no	NA	NA	NA	6036									NA		PREDICTED: Solanum tuberosum guanine nucleotide-binding protein-like NSN1 (LOC102593978)_mRNA			NA	199.818	11	2198	41
Acc0013659	LG13	no	yes	no	NA	yes	NA	NA									NA	NA	NA			NA	682.000	3	2046	247
Acc0013678	LG13	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	1514.000	1	1514	1514
Acc0013680	LG13	yes	no	no	NA	NA	NA	6295									NA	NA	NA			NA	588.333	3	1765	93
Acc0013682	LG13	yes	no	no	NA	yes	NA	NA									NA		PREDICTED: Glycine max nuclear export mediator factor NEMF-like (LOC100803745)_transcript variant X3_mRNA			NA	263.786	14	3693	81
Acc0013691	LG13	no	yes	no	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	523.000	3	1569	388
Acc0013692	LG13	no	yes	no	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	873.500	2	1747	623
Acc0013695	LG13	no	yes	no	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	941.000	2	1882	771
Acc0013717	LG13	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	2141.000	1	2141	2141
Acc0013718	LG13	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	396.333	3	1189	219
Acc0013719	LG13	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	696.500	2	1393	254
Acc0013720	LG13	yes	no	no	NA	NA	NA	6198									NA	NA	NA			NA	109.909	11	1209	24
Acc0013735	LG13	no	yes	no	NA	NA	NA	4757									NA	NA	NA			NA	203.500	6	1221	103
Acc0013744	LG13	yes	no	no	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	390.000	3	1170	30
Acc0013769	LG24	no	yes	no	yes	NA	NA	NA									NA	NA	NA			NA	138.300	10	1383	45
Acc0013776	LG24	yes	no	no	NA	NA	NA	NA									NA		PREDICTED: Oryza sativa Japonica Group homeobox-leucine zipper protein ROC7 (LOC4344614)_mRNA			NA	288.375	8	2307	135
Acc0013825	LG24	no	no	yes	NA	NA	NA	NA									NA					NA	428.111	9	3853	101

gene	seqname	Single_copy_Broc_A	Low_copy_Broc_Ana	High_copy_Broc_Ana	High_copy_Broc_Ana_Tilla	Diff_expressed_CAM_MC3	pos_selection_CAM	gene_families_CAM_C3	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Wal2017	malate_transferrase	flavonoid_anthocyanin_Picalirina_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size	
Acc0013826	LG08	no	yes	no	yes	yes	NA	NA	NA									NA	NA			NA		452.400	5	2262	45
Acc0013898	LG08	yes	no	no	NA	NA	NA	NA	6496									NA	NA			NA		162.700	10	1627	48
Acc0013901	LG08	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA		228.833	6	1373	178
Acc0013935	LG03	no	yes	no	NA	NA	NA	NA	NA	yes								NA	Beta vulgaris subsp. vulgaris malate dehydrogenase, cytoplasmic (LOC104887939)_mRNA			00_#19_Malate_dehydrogenase	219.571	7	1537	111	
Acc0013938	LG03	NA	NA	NA	NA	NA	NA	NA	NA	yes								NA	00_#14_Phosphoenolpyruvate_carboxylase_kinase			00_#14_Phosphoenolpyruvate_carboxylase_kinase	434.000	3	1302	367	
Acc0013950	LG03	no	yes	no	NA	NA	NA	NA	NA		yes							NA	01d_Auxin_response_factor			01d_Auxin_response_factor	585.000	5	2925	78	
Acc0013951	LG03	no	yes	no	NA	NA	NA	NA	5936									NA	NA			NA		191.333	21	4018	45
Acc0013954	LG03	no	yes	no	NA	NA	NA	NA	NA			yes						NA	02_Aquaporins			02_Aquaporins	324.750	4	1299	140	
Acc0013958	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	10_Glucogenesis_specific_dayC3	yes		10_Glucogenesis_specific_dayC3	280.667	6	1684	101	
Acc0013959	LG03	no	yes	no	NA	NA	yes	NA	NA									NA	10_Glucogenesis_specific_dayC3	yes		10_Glucogenesis_specific_dayC3	450.000	4	1800	239	
Acc0013965	LG03	no	no	yes	yes	NA	NA	NA	NA									NA	1032.500	2		NA		1032.500	2	2065	446
Acc0013974	LG03	yes	no	no	NA	NA	NA	NA	NA								yes	NA	04_Oxygen_evolution_enhancer_OEE			04_Oxygen_evolution_enhancer_OEE	347.000	4	1388	45	
Acc0013975	LG03	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA		200.111	9	1801	61
Acc0013992	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA		673.333	3	2020	495
Acc0014001	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA		320.400	5	1602	179
Acc0014023	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	10_Glucogenesis_specific_dayC3	yes		10_Glucogenesis_specific_dayC3	649.000	3	1947	446	
Acc0014027	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA		240.100	10	2401	116
Acc0014051	LG03	no	no	yes	NA	NA	NA	NA	NA									NA	PREDICTED: Solanum tuberosum heat shock cognate protein 80-like (LOC102600008)_mRNA			NA		326.063	16	5217	17
Acc0014055	LG03	no	yes	no	NA	NA	NA	NA	NA								yes	NA	03_VATPase			03_VATPase	179.333	18	3228	44	
Acc0014058	LG03	NA	NA	NA	NA	NA	NA	NA	NA									NA	10_Glucogenesis_specific_dayC3	yes		10_Glucogenesis_specific_dayC3	222.500	2	445	65	
Acc0014061	LG03	no	yes	no	NA	NA	NA	NA	NA									ALB_F_anthocyanidin synthase	NA			NA		408.143	7	2857	224
Acc0014063	LG03	yes	no	no	NA	NA	NA	NA	NA			yes						NA	06_Drought_resistance_genes_LEA			06_Drought_resistance_genes_LEA	299.000	2	598	260	
Acc0014077	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA		411.333	3	1234	243
Acc0014086	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA		368.250	4	1473	204
Acc0014118	LG13	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA		1353.500	2	2707	791
Acc0014120	LG13	yes	no	no	NA	NA	NA	NA	NA			yes						NA	06_Drought_resistance_genes_LEA			06_Drought_resistance_genes_LEA	272.000	1	272	272	
Acc0014214	LG13	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA		1622.000	1	1622	1622
Acc0014228	LG13	yes	no	no	NA	NA	NA	NA	6570									NA	NA			NA		219.200	5	1096	127
Acc0014232	LG13	yes	no	no	NA	NA	NA	NA	NA						yes			ALB_L_chalcone isomerase				00_#Ming2019_Anthocyanin_biosynthesis	269.000	4	1076	158	
Acc0014241	LG13	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA		1601.000	1	1601	1601
Acc0014257	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA		1331.000	2	2662	736

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilia	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Acomnosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism	malate_transferrase	malate_transferrase	flavonoid_anthocyanin_Picatarria_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0014274	LG05	NA	NA	NA	NA	NA	NA	NA	NA										NA	NA	yes		10_Glucogenesis_specific_dayC3	210.800	5	1054	94
Acc0014278	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	843.333	3	2530	161
Acc0014285	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	350.750	4	1403	256
Acc0014286	LG05	yes	no	no	NA	NA	yes	NA	NA									NA	NA	NA			NA	2027.500	2	4055	1669
Acc0014301	LG05	no	yes	no	NA	NA	NA	NA	NA									ALB_L_flavonoid glucosyl-transferase	NA	NA			NA	790.500	2	1581	647
Acc0014304	LG05	no	yes	no	NA	NA	NA	NA	NA									ALB_L_flavonoid glucosyl-transferase	NA	NA			NA	1457.000	1	1457	1457
Acc0014314	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	359.000	4	1436	210
Acc0014315	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	1883.000	1	1883	1883
Acc0014317	LG05	yes	no	no	NA	NA	NA	NA	NA								yes		NA	NA			03_VATPase	222.000	4	888	77
Acc0014321	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	3179.000	1	3179	3179
Acc0014334	LG05	no	yes	no	NA	NA	NA	NA	NA									ALB_L_cyanidin 3-o-rutinoside 5-o-glucosyltransferase	NA	NA			NA	313.808	26	8159	50
Acc0014342	LG05	no	yes	no	NA	NA	NA	NA	NA									NA	NA	NA	yes		11_Glycolysis_night_C4_malate_P	163.667	9	1473	38
Acc0014347	LG05	yes	no	no	NA	NA	NA	NA	NA							yes		NA	NA	NA			00_#W_CircadianOscillator_GIGA	277.067	15	4156	54
Acc0014352	LG05	no	yes	no	NA	NA	NA	NA	NA							yes		NA	NA	NA			00_#W_StomatalMovement_Synta	746.000	2	1492	589
Acc0014364	LG05	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	534.000	3	1602	355
Acc0014393	LG07	no	no	yes	NA	NA	NA	NA	NA			yes						NA	NA	NA			06_Drought_resistance_genes_LEA	471.000	2	942	235
Acc0014402	LG07	no	yes	no	NA	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	200.500	10	2005	72
Acc0014404	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	1619.000	1	1619	1619
Acc0014406	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	254.500	6	1527	68
Acc0014415	LG07	no	no	yes	NA	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	432.500	4	1730	135
Acc0014417	LG07	no	yes	no	NA	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	125.846	13	1636	48
Acc0014433	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	549.200	5	2746	117
Acc0014452	LG07	no	yes	no	NA	NA	NA	NA	NA								yes	NA	NA	NA			04_Oxygen_evolution_enhancer_OEE	250.000	3	750	98
Acc0014474	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	1259.000	1	1259	1259
Acc0014488	LG07	yes	no	no	NA	NA	yes	NA	NA	yes								NA	NA	NA			00_#38_Pentatricopeptide	720.750	4	2883	376
Acc0014493	LG07	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	482.667	3	1448	445
Acc0014555	LG20	no	yes	no	yes	NA	NA	NA	NA									NA	NA	NA			03_VATPase	387.000	3	1161	251
Acc0014579	LG08	yes	no	no	NA	NA	NA	NA	6746								yes	NA	NA	NA			NA	246.000	4	984	71
Acc0014581	LG08	no	yes	no	yes	yes	NA	NA	NA									NA	NA	NA			NA	227.556	9	2048	98
Acc0014582	LG08	yes	no	no	NA	NA	NA	NA	NA									NA	NA	NA			NA	421.500	4	1686	142
Acc0014590	LG08	no	yes	no	NA	NA	NA	NA	NA									NA	NA	Populus trichocarpa hypothetical protein (POPTR_0002s25930g) mRNA_complete cds			NA	211.000	4	844	41

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilia	Diff_expressed CAM_MC3	pos_selection_CAM_shifts	gene_C3	Angiosperm_gene	CAM-related Mim2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate transferase Assimilation_Carbohyd	flavonoid_anthocyanin_Picirratia_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0014592	LG08	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	345.400	5	1727	125
Acc0014603	LG08	no	no	yes	NA	NA	NA	NA	NA				yes					NA	NA			00_#MB_RPB2	177.440	25	4436	44
Acc0014616	LG08	yes	no	no	NA	NA	NA	NA	6500									NA	NA			NA	227.800	10	2278	90
Acc0014628	LG08	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1697.000	1	1697	1697
Acc0014682	LG20	NA	NA	NA	NA	NA	NA	NA	NA								yes	NA	NA			04_Oxygen_evolution_enhancer_OEE	186.500	2	373	98
Acc0014690	LG20	no	yes	no	NA	NA	NA	NA	NA	yes								NA	PREDICTED: Beta vulgaris subsp. vulgaris malate dehydrogenase_chloroplastic-like (LOC104908004)_transcript variant X2_mRNA			00_#22_Malate_dehydrogenase	766.000	2	1532	11
Acc0014693	LG20	no	yes	no	NA	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	230.800	5	1154	61
Acc0014697	LG20	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	528.800	5	2644	160
Acc0014716	LG20	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1874.000	1	1874	1874
Acc0014745	LG20	yes	no	no	NA	NA	NA	NA	5551									NA	NA			NA	147.500	12	1770	46
Acc0014753	LG20	no	no	yes	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	320.000	5	1600	47
Acc0014755	LG20	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	782.500	2	1565	684
Acc0014770	LG20	yes	no	no	NA	NA	yes	NA	NA									NA	NA			NA	1319.000	1	1319	1319
Acc0014781	LG20	yes	no	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	287.667	3	863	215
Acc0014791	LG20	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	369.167	6	2215	222
Acc0014801	LG14	yes	no	no	NA	NA	NA	NA	5318									NA	NA	yes		10_Glucogenesis_specific_dayC3	174.375	8	1395	62
Acc0014810	LG14	no	yes	no	NA	NA	NA	NA	5296									NA	NA	yes		10_Glucogenesis_specific_dayC3	156.364	11	1720	34
Acc0014831	LG14	no	no	yes	NA	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_reponse_factor	287.647	17	4550	56
Acc0014837	LG14	NA	NA	NA	yes	NA	NA	NA	NA									NA	NA			NA	287.000	1	287	287
Acc0014842	LG14	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	568.000	3	1704	178
Acc0014855	LG14	yes	no	no	NA	NA	NA	NA	5304									NA	NA			NA	116.556	9	1049	45
Acc0014864	LG14	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	310.667	6	1864	122
Acc0014875	LG14	no	yes	no	NA	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	704.000	1	704	704
Acc0014884	LG14	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	275.400	5	1377	210
Acc0014888	LG14	yes	no	no	NA	NA	NA	NA	NA									ALB_F_anthocyanidin_reductase	NA	yes		10_Glucogenesis_specific_dayC3	175.500	6	1053	144
Acc0014895	LG14	yes	no	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	195.400	5	977	55
Acc0014899	LG14	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	153.000	13	1989	48
Acc0014931	LG04	yes	no	no	NA	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_pathway	143.250	8	1146	53
Acc0014937	LG04	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	792.333	3	2377	446
Acc0014942	LG04	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1172.000	1	1172	1172

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_families_CAM	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate transferase Assimilation_Carbohyd	flavonoid_anthocyanin_P5	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size			
Acc0014951	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	311.143	7	2178	150			
Acc0014959	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1385.000	1	1385	1385			
Acc0014967	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1132.500	2	2265	570			
Acc0015009	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	360.167	6	2161	215			
Acc0015014	LG04	yes	no	no	yes	NA	NA	NA									NA	NA			NA	1310.000	1	1310	1310			
Acc0015020	LG04	yes	no	no	NA	NA	NA	5528									NA	NA			NA	118.913	23	2735	38			
Acc0015031	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	903.500	2	1807	345			
Acc0015035	LG13	yes	no	no	NA	NA	NA	NA									NA	NA			NA	515.000	3	1545	423			
Acc0015042	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	730.333	3	2191	172			
Acc0015045	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	986.500	2	1973	954			
Acc0015047	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	371.667	3	1115	269			
Acc0015060	LG01	no	yes	no	NA	NA	NA	6732									NA	NA			NA	255.125	16	4082	41			
Acc0015072	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	656.000	2	1312	366			
Acc0015073	LG01	no	yes	no	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_reponse_factor	254.800	15	3822	14			
Acc0015090	LG01	no	yes	no	NA	NA	NA	NA									NA	NA		yes	10_Glucogenesis_specific_dayC3	252.300	10	2523	94			
Acc0015107	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	459.333	3	1378	130			
Acc0015108	LG01	yes	no	no	NA	NA	NA	5642									NA	NA			NA	101.300	10	1013	22			
Acc0015128	LG24	yes	no	no	NA	yes	NA	4848									NA	NA			NA	405.857	7	2841	136			
Acc0015141	LG24	yes	no	no	NA	NA	NA	NA									NA	NA			NA	384.333	3	1153	360			
Acc0015171	LG05	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Glycine max kinesin-4-like (LOC100777089)_transcript variant X4_mRNA										21
Acc0015180	LG05	yes	no	no	NA	NA	NA	NA									NA	NA			NA	204.524	21	4295	21			
Acc0015184	LG05	no	no	yes	NA	NA	NA	NA									NA	NA			11_Glycolysis_night_C4_malate_p_athway	123.929	14	1735	40			
Acc0015194	LG05	NA	NA	NA	NA	NA	NA	NA		yes							NA	NA			01b_Abscisic_acid_signal	315.500	2	631	17			
Acc0015197	LG05	yes	no	no	NA	NA	NA	NA									NA	NA			NA	272.400	5	1362	121			
Acc0015198	LG05	no	yes	no	NA	NA	yes	NA									NA	NA			NA	842.000	2	1684	378			
Acc0015199	LG05	no	yes	no	NA	NA	yes	NA									NA	NA			NA	907.000	3	2721	207			
Acc0015201	LG05	yes	no	no	NA	NA	NA	5430									NA	NA			NA	193.286	7	1353	68			
Acc0015211	LG05	yes	no	no	NA	NA	NA	NA									NA	NA			NA	576.000	2	1152	564			
Acc0015226	LG05	yes	no	no	NA	NA	NA	NA							yes		NA	NA			00_#W_ProtonPump_VHA-B_03_VATPase	128.533	15	1928	47			
Acc0015235	LG05	yes	no	no	NA	NA	NA	NA									NA	NA			NA	623.000	2	1246	566			
Acc0015240	LG05	no	yes	no	NA	NA	NA	NA									NA	NA		yes	11_Glycolysis_night_C4_malate_p_athway	313.000	4	1252	31			

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	an_Tilla	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	flavonoid_anthocyanin_Picalirnia_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0015263	LG20	yes	no	no	NA	NA	NA	NA	6050									NA	NA			NA	123.583	12	1483	56
Acc0015297	LG20	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	611.000	2	1222	235
Acc0015298	LG20	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	176.786	14	2475	44
Acc0015317	LG20	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	240.200	5	1201	111
Acc0015319	LG20	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	299.000	4	1196	159
Acc0015324	LG20	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	316.000	9	2844	38
Acc0015329	LG20	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	830.667	3	2492	549
Acc0015338	LG20	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	384.333	3	1153	136
Acc0015340	LG20	no	yes	no	yes	NA	NA	NA	NA									NA	NA			NA	564.500	2	1129	525
Acc0015349	LG20	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	345.000	5	1725	199
Acc0015352	LG20	yes	no	no	NA	NA	NA	NA	7194									NA	NA			NA	85.545	11	941	44
Acc0015377	LG23	no	yes	no	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	975.500	2	1951	887
Acc0015387	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	2051.000	1	2051	2051
Acc0015404	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	916.000	2	1832	852
Acc0015411	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	714.500	2	1429	456
Acc0015420	LG23	yes	no	no	NA	NA	NA	NA	5064									NA	NA			NA	226.000	3	678	105
Acc0015422	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1973.000	1	1973	1973
Acc0015434	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1448.000	1	1448	1448
Acc0015436	LG23	no	yes	no	NA	NA	yes	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	244.000	7	1708	85
Acc0015443	LG23	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	1337.000	1	1337	1337
Acc0015453	LG23	no	yes	no	yes	NA	NA	NA	NA									NA	NA			NA	152.667	9	1374	75
Acc0015495	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	218.571	7	1530	114
Acc0015513	LG03	yes	no	no	NA	NA	NA	NA	5090									NA	NA			NA	117.500	2	235	2
Acc0015521	LG03	yes	no	no	yes	NA	NA	NA	NA									NA	NA			NA	659.000	1	659	659
Acc0015530	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	396.600	5	1983	119
Acc0015541	LG03	no	yes	no	NA	NA	NA	NA	NA							yes		NA	NA			00_#W_StomatalMovement_Carbo	581.000	3	1743	384
Acc0015546	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	NA			0_n_dioxide_insensitive3_Cb13_11_				
Acc0015555	LG03	no	yes	no	yes	NA	yes	NA	NA									NA	NA			NA	1421.000	1	1421	1421
Acc0015577	LG03	no	no	yes	NA	NA	NA	NA	NA									NA	NA			NA	251.333	12	3016	55
Acc0015588	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	281.000	9	2529	101
Acc0015588	LG03	yes	no	no	NA	NA	NA	NA	NA									NA	NA			NA	580.000	2	1160	274

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_families_CAM	Angiosperm_gene	CAM-related_Ming2015	stomata_lunc	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	flavonoid_anthocyanin_Picalirnia_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc00155689	LG03	NA	NA	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	137.000	4	548	52
Acc00155592	LG03	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	356.600	5	1783	140
Acc00155593	LG03	yes	no	no	NA	NA	NA	NA									NA	NA			NA	544.714	7	3813	126
Acc00156568	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	199.000	9	1791	112
Acc00156685	LG19	no	yes	no	NA	NA	NA	4932									NA	NA			NA	195.444	9	1759	51
Acc00157111	LG19	no	yes	no	NA	NA	NA	NA									NA	NA			NA	395.000	4	1580	41
Acc0015727	LG09	no	yes	no	NA	yes	NA	5502									NA	NA			NA	127.682	22	2809	33
Acc0015741	LG09	yes	no	no	NA	NA	yes	NA									NA	NA			NA	947.000	1	947	947
Acc0015760	LG09	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	124.417	12	1483	40
Acc0015761	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2480.000	1	2480	2480
Acc0015775	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	727.500	2	1455	350
Acc0015779	LG09	no	yes	no	NA	NA	NA	NA							yes		NA	NA			00_#W_HexoseTransporter_MSPP-like transporter	484.286	7	3390	124
Acc0015799	LG09	yes	no	no	NA	NA	NA	5594									NA	NA			NA	382.667	3	1148	296
Acc0015842	LG01	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	114.385	13	1487	36
Acc0015885	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	451.667	3	1355	227
Acc0015877	LG22	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	625.667	3	1877	122
Acc0016004	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	381.500	4	1526	144
Acc0016014	LG22	yes	no	no	yes	NA	NA	NA									NA	NA			NA	720.500	2	1441	57
Acc0016019	LG22	yes	no	no	NA	yes	NA	NA									NA	NA			NA	234.500	6	1407	136
Acc0016029	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	608.250	4	2433	269
Acc0016047	LG21	yes	no	no	NA	NA	NA	7273									NA	NA			NA	108.500	10	1085	53
Acc0016057	LG21	yes	no	no	NA	NA	NA	7028									NA	NA			NA	220.250	4	881	69
Acc0016070	LG21	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	114.214	14	1599	36
Acc0016080	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	838.500	2	1677	194
Acc0016087	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1165.500	2	2331	124
Acc0016103	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	833.500	2	1667	696
Acc0016112	LG21	no	yes	no	NA	NA	NA	NA									NA	Populus trichocarpa adenosylhomocysteinase family protein (POPTR_0017s08610g) mRNA_complete cds			NA	508.000	4	2032	357
Acc0016139	LG21	yes	no	no	NA	yes	NA	NA									NA	NA			NA	2228.000	1	2228	2228
Acc0016151	LG21	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	280.000	5	1400	102
Acc0016200	LG17	yes	no	no	NA	NA	NA	NA						yes			ALB_F_chalcone synthase	NA			00_#Ming2019_Anthocyanin_biosynthesis	407.000	4	1628	155

gene	seqname	Single copy Broc_A	Low n Tillar	High copy Broc_Ana	Diff_Mc3	pos_selection_CAM	gene_families_CAM	Angiosperm_gene	CAM-related_Acomorus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferase_Assimilation_Carbohyd	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0016245	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2318.000	1	2318	2318
Acc0016255	LG03	no	yes	no	yes	NA	NA	NA									NA	NA			NA	416.000	3	1248	22
Acc0016258	LG03	no	yes	no	yes	NA	NA	NA									NA	NA			NA	156.444	9	1408	69
Acc0016261	LG03	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	602.000	1	602	602
Acc0016318	LG03	yes	no	no	NA	NA	yes	NA									NA	NA			NA	551.000	2	1102	462
Acc0016323	LG03	yes	no	no	NA	NA	NA	NA									NA	NA			NA	661.500	2	1323	203
Acc0016352	LG03	yes	no	no	NA	NA	NA	5660									NA	NA			NA	162.500	6	975	51
Acc0016357	LG03	no	yes	no	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	272.000	3	816	132
Acc0016360	LG03	yes	no	no	NA	yes	NA	7024									NA	NA			NA	151.750	4	607	34
Acc0016390	LG10	yes	no	no	yes	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_P	175.500	8	1404	38
Acc0016393	LG10	no	yes	no	NA	NA	NA	NA									ALB_F_isoflavone_reductase	NA			NA	209.286	7	1465	46
Acc0016395	LG10	yes	no	no	NA	NA	NA	5639									NA	NA			NA	242.286	7	1696	91
Acc0016429	LG10	yes	no	no	NA	NA	NA	NA	yes								NA	NA			00_#12_10_Glucogenesis_specific_dayC3	192.947	19	3666	61
Acc0016437	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1568.000	1	1568	1568
Acc0016438	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	730.333	3	2191	167
Acc0016511	LG11	yes	no	no	NA	NA	NA	NA									NA	NA			NA	384.000	3	1152	227
Acc0016554	LG11	no	no	yes	NA	NA	NA	NA							yes		NA	NA			00_#W_StomatalMovement_Phospholipase_Dat	493.900	10	4939	74
Acc0016562	LG09	yes	no	no	NA	NA	NA	5116									NA	NA			NA	241.875	8	1935	86
Acc0016569	LG09	no	yes	no	NA	NA	NA	NA	yes								NA	PREDICTED: Beta vulgaris subsp. vulgaris NAD-dependent malic enzyme 59 kDa isoform_mitochondrial (LOC10488560)_mRNA			00_#10_Glucogenesis_specific_dayC3	112.300	20	2246	22
Acc0016575	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	929.000	2	1858	707
Acc0016607	LG08	yes	no	no	NA	NA	NA	NA									NA	NA			NA	369.667	3	1109	200
Acc0016636	LG08	yes	no	no	NA	NA	NA	NA									NA	NA			NA	218.167	6	1309	120
Acc0016649	LG08	yes	no	no	NA	yes	NA	NA							yes		NA	NA			00_#W_CircadianOscillator_CCA1/LHY	352.000	8	2816	61
Acc0016654	LG08	NA	NA	NA	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_response_factor	614.000	1	614	614
Acc0016662	LG08	yes	no	no	NA	yes	NA	NA									NA	NA			NA	1628.000	1	1628	1628
Acc0016690	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1406.000	1	1406	1406
Acc0016697	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	742.000	4	2968	160
Acc0016708	LG17	yes	no	no	NA	NA	NA	NA									NA	NA			NA	743.000	2	1486	272
Acc0016710	LG17	no	no	yes	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	564.000	2	1128	356
Acc0016712	LG17	no	no	yes	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	749.000	2	1498	359

gene	seqname	Single_copy_Broc_A nan_Tilia.	Low_copy_Broc_Ana n_Tilia.	High_copy_Broc_An n_Tilia.	Diff_expressed_CA MC3	pos_selection_CAM _shifts	gene_C3 families_CAM	Angiosperm_gene	CAM- related_Acomosus- Ming2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana nas	circadian_metabolis m_Walz2017	malate_transferase Assimilation_Carbohy drates	flavonoid_anthocya nin_Picalirnia_PS	pos_sel_portulugo	glucogenesis_glycol ysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0016722	LG17	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	402.000	3	1206	145
Acc0016727	LG17	NA	NA	NA	NA	NA	NA	NA	yes							NA	NA	NA			00_#18_Malate_dehydrogenase	249.500	4	998	182
Acc0016732	LG17	no	yes	no	NA	NA	NA	6462								NA	NA	NA			NA	203.111	9	1828	33
Acc0016748	LG17	no	yes	no	NA	NA	NA	NA		yes						NA	NA	NA			01d_Auxin_reponse_factor	514.667	3	1544	195
Acc0016764	LG17	yes	no	no	NA	NA	NA	NA				yes				NA	NA	NA			06_Drought_resistance_genes_LEA	355.200	5	1776	55
Acc0016768	LG17	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	1547.000	1	1547	1547
Acc0016769	LG17	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	367.500	4	1470	266
Acc0016773	LG17	no	no	yes	NA	NA	NA	NA								NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	371.143	7	2598	32
Acc0016774	LG17	no	no	yes	NA	NA	NA	NA								NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	229.857	7	1609	112
Acc0016784	LG17	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	2372.000	1	2372	2372
Acc0016803	LG04	no	yes	no	NA	NA	NA	5131								NA	NA	NA			NA	272.143	7	1905	90
Acc0016807	LG04	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	633.333	3	1900	216
Acc0016809	LG04	no	yes	no	yes	NA	NA	NA								NA	NA	NA			NA	323.167	6	1939	123
Acc0016819	LG04	yes	no	no	NA	NA	NA	NA								NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	190.286	7	1332	36
Acc0016848	LG04	no	no	yes	NA	NA	NA	NA								NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	347.143	7	2430	37
Acc0016862	LG04	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	411.000	4	1644	224
Acc0016876	LG04	yes	no	no	NA	NA	NA	6685								NA	NA	NA			NA	152.636	11	1679	49
Acc0016888	LG02	no	yes	no	NA	NA	NA	5921								NA	NA	NA			NA	150.630	27	4067	60
Acc0016897	LG02	no	no	yes	NA	NA	NA	NA								NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	437.800	5	2189	267
Acc0016898	LG02	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	253.800	5	1269	111
Acc0016907	LG02	no	yes	no	NA	NA	NA	5596								NA	NA	NA			NA	425.333	12	5104	45
Acc0016943	LG02	no	no	yes	yes	NA	NA	NA								NA	NA	NA			NA	957.333	3	2872	214
Acc0016948	LG02	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	653.000	2	1306	410
Acc0016978	LG08	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	596.667	3	1790	172
Acc0016979	LG08	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	373.833	6	2243	144
Acc0016986	LG08	yes	no	no	yes	NA	NA	NA								NA	NA	NA			NA	517.667	3	1553	195
Acc0016993	LG08	yes	no	no	NA	yes	NA	NA								NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	170.800	15	2562	57
Acc0017006	LG08	yes	no	no	NA	NA	NA	6946								NA	NA	NA			NA	160.600	5	803	26
Acc0017008	LG08	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	1583.000	1	1583	1583
Acc0017030	LG08	yes	no	no	NA	NA	NA	NA								NA	NA	NA			NA	569.000	3	1707	524
Acc0017042	LG08	yes	no	no	yes	NA	NA	NA								NA	NA	NA			NA	534.750	4	2139	373

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_families_CAM	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate transferase Assimilation_Carboxhydrates	flavonoid anthocyanin_Picalirina_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0017062	LG14	yes	no	no	NA	yes	NA	NA									NA	NA			NA	863.500	2	1727	427
Acc0017076	LG14	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Beta vulgaris subsp. vulgaris glucose-1-phosphate adenyltransferase small subunit, chloroplast/amyloplastic (LOC104903071), mRNA	yes		10_Glucogenesis_specific_dayC3	202.778	9	1825	98
Acc0017100	LG14	yes	no	no	NA	NA	NA	NA									NA	NA			NA	235.800	5	1179	154
Acc0017111	LG14	yes	no	no	NA	NA	NA	5460									NA	NA			NA	90.000	9	810	38
Acc0017115	LG14	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	122.250	12	1467	48
Acc0017128	LG14	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	396.500	2	793	311
Acc0017133	LG14	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_p athway	142.316	19	2704	44
Acc0017138	LG14	yes	no	no	NA	NA	NA	NA									NA	NA			NA	680.000	4	2720	226
Acc0017169	LG18	no	yes	no	NA	NA	NA	NA					yes				STA_L_Flavonoid 3- hydroxylase	NA			00_#Ming2019_Anthocyanin_biosy nthesis	572.000	3	1716	194
Acc0017171	LG18	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1937.000	1	1937	1937
Acc0017183	LG18	NA	NA	NA	NA	NA	NA	NA		yes							NA	NA			01b_Abscisic_acid_signal	726.500	2	1453	57
Acc0017195	LG18	yes	no	no	NA	NA	NA	NA									NA	NA			NA	2033.000	1	2033	2033
Acc0017212	LG18	yes	no	no	NA	NA	NA	NA									NA	NA			NA	257.000	6	1542	183
Acc0017240	LG03	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1401.000	2	2802	234
Acc0017241	LG03	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	1268.000	1	1268	1268
Acc0017259	LG03	yes	no	no	NA	NA	NA	NA									NA	NA			NA	306.400	5	1532	125
Acc0017266	LG03	yes	no	no	NA	NA	NA	NA									NA	NA			NA	442.000	3	1326	144
Acc0017276	LG03	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1808.000	1	1808	1808
Acc0017282	LG03	yes	no	no	NA	NA	NA	NA				yes					NA	00_#MB_phosphothiolokinasase-like protein (PRK) gene				406.500	4	1626	244
Acc0017286	LG03	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Solanum tuberosum ATPase family AAA domain-containing protein 3C (LOC102578202), mRNA			NA	272.286	7	1906	90
Acc0017324	LG01	yes	no	no	yes	NA	NA	NA									NA	NA			NA	600.667	3	1802	152
Acc0017325	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	347.250	4	1389	152
Acc0017331	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	486.000	3	1458	152
Acc0017335	LG01	yes	no	no	NA	NA	NA	6779									NA	NA			NA	216.200	5	1081	80
Acc0017336	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1583.000	1	1583	1583
Acc0017339	LG01	yes	no	no	NA	yes	NA	NA									NA	NA			NA	1281.500	2	2563	1054
Acc0017365	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	969.000	2	1938	271
Acc0017433	LG09	no	no	yes	NA	NA	NA	NA									NA	PREDICTED: Glycine max protein argonaute 1- like (LOC100718754), mRNA			NA	159.591	22	3511	20
Acc0017479	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	354.222	9	3188	124
Acc0017482	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1802.000	1	1802	1802

gene	seqname	Single copy Broc_A nan_Tilia	Low copy Broc_Ana n_Tilia	High copy Broc_Ana n_Tilia	Diff_expressed_CAM3	pos_selection_CAM-shifts	gene_C3_families_CAM	Angiosperm_gene	CAM-related Mim2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Anas	circadian_metabolism_Wal2017	malate_transferrase_Assimilation_Carbohydrites	flavonoid_anthocyanin_PSc	pos_sel_portulungo	glucogenesis_ghycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0017488	LG22	yes	no	no	NA	NA	NA	NA									NA	PREDICTED: Virus vinniera cell division control protein 48 homolog C (LOC100261441)_mRNA			NA	254.818	11	2803	64
Acc0017525	LG22	no	yes	no	NA	NA	NA	NA	yes								NA				00_#23_Lactate_dehydrogenase	471.667	3	1415	326
Acc0017526	LG22	no	yes	no	NA	NA	NA	NA	yes								NA				00_#24_hypotheical_protein	243.500	2	487	23
Acc0017527	LG22	no	yes	no	NA	NA	NA	NA	yes								NA				00_#25_hypotheical_protein	197.000	5	985	15
Acc0017528	LG22	no	yes	no	NA	NA	NA	NA	yes								NA				00_#26_Lactate_dehydrogenase	351.000	4	1404	110
Acc0017531	LG22	yes	no	no	NA	NA	NA	NA									NA				NA	459.000	3	1377	275
Acc0017533	LG22	no	yes	no	NA	NA	NA	NA							yes		NA				00_#W_VacuolarAcidInvertase_AI	268.875	8	2151	8
Acc0017540	LG22	yes	no	no	NA	NA	NA	NA	54.28								NA				NA	612.500	2	1225	43
Acc0017546	LG22	yes	no	no	NA	NA	NA	NA									NA				NA	319.750	4	1279	161
Acc0017554	LG09	yes	no	no	NA	NA	NA	NA	73.71								NA				NA	335.000	2	670	240
Acc0017556	LG09	yes	no	no	NA	NA	NA	NA		yes							NA				01b_Absciscic_acid_signal	312.600	5	1563	12
Acc0017624	LG14	no	yes	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	435.250	4	1741	304
Acc0017637	LG14	yes	no	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	156.000	8	1248	46
Acc0017688	LG14	no	no	yes	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	373.286	7	2613	32
Acc0017701	LG05	yes	no	no	NA	NA	NA	NA		yes							NA				01d_Auxin_reponse_factor	342.200	5	1711	40
Acc0017705	LG05	yes	no	no	NA	NA	NA	NA									NA				NA	319.750	4	1279	116
Acc0017761	LG05	yes	no	no	NA	NA	NA	NA									NA				NA	463.400	5	2317	118
Acc0017767	LG21	NA	NA	NA	NA	NA	NA	NA							yes		NA				00_#W_CircadianOscillator_casein_kinase2_subunit_beta	294.200	5	1471	106
Acc0017802	LG21	yes	no	no	NA	NA	NA	NA									NA				NA	645.500	2	1291	400
Acc0017829	LG21	no	yes	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	191.750	8	1534	65
Acc0017858	LG23	no	no	yes	NA	NA	NA	NA							yes		NA				00_#W_StomatalMovement_ATP-binding_cassette_C4	350.769	13	4560	63
Acc0017860	LG23	no	yes	no	NA	NA	NA	NA									NA				NA	138.304	23	3181	31
Acc0017874	LG04	yes	no	no	NA	NA	NA	NA									NA				NA	1457.000	1	1457	1457
Acc0017939	LG09	NA	NA	NA	NA	NA	NA	NA									ALB_F_flavonol_synthase				NA	338.000	3	1014	194
Acc0017945	LG09	no	yes	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	174.154	13	2264	26
Acc0017965	LG19	no	yes	no	NA	NA	NA	NA		yes							NA				01d_Auxin_reponse_factor	275.750	4	1103	95
Acc0017967	LG19	yes	no	no	NA	NA	NA	NA									NA				NA	1117.500	2	2235	968
Acc0017975	LG19	yes	no	no	NA	NA	NA	NA									NA				NA	2432.000	1	2432	2432
Acc0017983	LG19	no	yes	no	NA	NA	NA	NA									NA				NA	309.000	9	2781	92
Acc0017991	LG01	yes	no	no	NA	NA	NA	NA									NA		yes		10_Glucogenesis_specific_dayC3	224.875	8	1799	36
Acc0018002	LG01	yes	no	no	NA	yes	NA	NA									NA				NA	181.909	11	2001	62

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohyd	flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0018009	LG01	no	yes	no	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	1012.000	2	2024	260
Acc0018038	LG01	yes	no	no	NA	NA	NA	NA									NA	PREDICTED: Beta vulgaris subsp. vulgaris adenylate kinase 2, chloroplastic (LOC104898882)_transcript variant X1_mRNA			NA	183.857	7	1287	74
Acc0018061	LG11	no	yes	no	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_reponse_factor	256.333	12	3076	56
Acc0018084	LG11	NA	NA	NA	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_p_athway	348.500	4	1394	122
Acc0018093	LG11	no	yes	no	NA	NA	NA	NA	yes								NA	PREDICTED: Beta vulgaris subsp. vulgaris phosphoenolpyruvate carboxylase 1 (LOC104903486)_mRNA			00_#10_Glucogenesis_specific_dayC3	343.100	10	3431	84
Acc0018101	LG11	NA	NA	NA	NA	NA	NA	NA		yes							NA	NA			01b_Abscicic_acid_signal	617.000	1	617	617
Acc0018117	LG11	no	yes	no	NA	NA	NA	NA									NA	NA			01d_Auxin_reponse_factor	131.857	7	923	4
Acc0018141	LG12	yes	no	no	NA	NA	NA	NA			yes						NA	NA			00_#Ming2019_Anthocyanin_biosynthesis	1400.000	1	1400	1400
Acc0018176	LG12	yes	no	no	NA	NA	NA	NA									NA	NA			NA	620.667	3	1862	115
Acc0018265	LG01	no	yes	no	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	371.500	2	743	339
Acc0018268	LG01	no	yes	no	NA	NA	NA	NA									NA	NA			06_Drought_resistance_genes_LEA	614.000	1	614	614
Acc0018269	LG01	no	yes	no	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	168.000	3	504	4
Acc0018270	LG01	no	yes	no	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	584.000	1	584	584
Acc0018272	LG01	no	yes	no	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	677.000	1	677	677
Acc0018273	LG01	no	yes	no	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	677.000	1	677	677
Acc0018321	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	388.333	3	1165	142
Acc0018350	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	901.000	2	1802	656
Acc0018356	LG07	yes	no	no	NA	NA	NA	5840									NA	NA	yes		10_Glucogenesis_specific_dayC3	161.500	8	1292	33
Acc0018440	LG01	yes	no	no	yes	NA	NA	NA									NA	NA			NA	217.200	10	2172	43
Acc0018451	LG01	no	no	yes	yes	NA	NA	NA									NA	NA			NA	262.000	10	2620	6
Acc0018452	LG01	yes	no	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	393.667	3	1181	71
Acc0018478	LG21	no	yes	no	NA	NA	NA	6036									NA	PREDICTED: Solanum tuberosum guanine nucleotide-binding protein-like NSN1 (LOC102593978)_mRNA			NA	223.000	10	2230	66
Acc0018512	LG06	yes	no	no	NA	NA	NA	NA									NA	NA			NA	400.000	3	1200	229
Acc0018514	LG06	yes	no	no	NA	NA	NA	NA									NA	NA			NA	600.000	2	1200	538
Acc0018517	LG06	yes	no	no	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	329.500	2	659	170
Acc0018526	LG06	no	yes	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_right_C4_malate_p_athway	123.737	19	2351	44
Acc0018527	LG06	yes	no	no	NA	NA	NA	NA									NA	NA			NA	297.000	4	1188	155
Acc0018533	LG06	yes	no	no	yes	NA	NA	NA									NA	NA			NA	377.000	1	377	377
Acc0018556	LG06	yes	no	no	NA	NA	NA	NA									NA	NA			NA	590.333	3	1771	153
Acc0018557	LG06	yes	no	no	NA	NA	NA	NA									NA	NA			NA	676.500	2	1353	170

gene	seqname	Single copy Broc_A nan_Tilla.	Low copy Broc_Ana n_Tilla.	High copy Broc_Ana Tilla.	Diff expressed CA MC3	pos_selection CAM shifts	gene_family CAM C3	Angiosperm_gene	CAM-related Acomosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Wal2017	malateTransferase Assmilation_Carbohyd	flavonoid_anthocyanin_Pitcarnia_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size	
Acc0018567	LG01	no	yes	no	NA	NA	NA	NA							yes		NA	NA				00_#NW_ProteinPump_VHA-C_03_VATPase	161.273	11	1774	67
Acc0018571	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	NA	678.667	3	2036	167
Acc0018578	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	NA	341.600	5	1708	125
Acc0018579	LG01	yes	no	no	NA	NA	NA	7279									NA	NA			NA	NA	166.700	10	1667	50
Acc0018598	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	NA	437.250	4	1749	303
Acc0018609	LG22	yes	no	no	NA	NA	NA	NA						yes			ALB_F_flavanone 3-hydroxylase	NA			00_#Ming2019_Anthocyanin_biosynthesis	572.000	2	1144	440	
Acc0018625	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	NA	749.000	3	2247	432
Acc0018626	LG22	yes	no	no	NA	NA	NA	4744									NA	NA			NA	NA	182.364	11	2006	49
Acc0018637	LG22	no	no	yes	yes	NA	NA	NA									NA	NA		yes		10_Glucogenesis_specific_dayC3	414.500	4	1658	258
Acc0018645	LG22	no	no	yes	NA	NA	NA	NA									NA		PREDICTED: Solanum tuberosum heat shock cognate protein 80-like (LOC102600008)_mRNA		NA	NA	529.250	4	2117	122
Acc0018648	LG22	no	no	yes	NA	NA	NA	NA									NA		PREDICTED: Solanum tuberosum heat shock cognate protein 80-like (LOC102600008)_mRNA		NA	NA	230.000	9	2070	26
Acc0018665	LG14	yes	no	no	NA	NA	NA	6652									NA	NA			NA	NA	205.300	10	2053	87
Acc0018673	LG14	yes	no	no	NA	NA	NA	NA									NA	NA			NA	NA	793.667	3	2381	137
Acc0018674	LG14	yes	no	no	NA	NA	NA	NA									NA	NA			NA	NA	690.500	2	1381	117
Acc0018695	LG14	yes	no	no	NA	NA	NA	5347									NA	NA			NA	NA	304.750	4	1219	72
Acc0018780	LG13	NA	NA	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	175.333	3	526	62	
Acc0018851	LG08	yes	no	no	NA	NA	NA	NA									NA	NA			NA	NA	1637.000	1	1637	1637
Acc0018853	LG08	no	yes	no	NA	NA	NA	NA		yes							NA	NA			01b_Absciscic_acid_signal		680.000	1	680	680
Acc0018892	LG20	yes	no	no	NA	NA	NA	NA									NA	NA			NA	NA	1598.000	1	1598	1598
Acc0018949	LG12	yes	no	no	NA	NA	NA	6531									NA	NA			NA	NA	87.667	9	789	6
Acc0018974	LG12	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	126.800	10	1268	44	
Acc0018978	LG02	no	no	yes	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_response_factor	250.643	14	3509	56	
Acc0018982	LG02	yes	no	no	NA	NA	NA	NA							yes		NA	NA			09_Circadian_clock	211.600	5	1058	57	
Acc0019042	LG20	no	yes	no	NA	yes	NA	NA									NA	NA			NA	NA	291.083	12	3493	45
Acc0019058	LG20	NA	NA	NA	NA	NA	NA	NA				yes					NA	NA			00_#Versieux_Leafy	365.000	1	365	365	
Acc0019059	LG20	NA	NA	NA	NA	NA	NA	NA					yes				NA	NA			00_#Versieux_Leafy	124.250	4	497	11	
Acc0019060	LG20	yes	no	no	NA	NA	NA	NA									NA	NA			NA	NA	499.444	9	4495	125
Acc0019065	LG20	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	141.125	16	2258	59	
Acc0019068	LG20	no	yes	no	NA	NA	NA	NA									ALB_F_dihydroflavonol 4-reductase	NA			NA	NA	165.875	8	1327	39
Acc0019096	LG06	no	yes	no	NA	NA	NA	NA									NA	NA		yes	10_Glucogenesis_specific_dayC3	318.500	8	2548	102	

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff. expressed CA	pos_selection_CAM	gene_families_CAM	Angiosperm_gene	CAM-related_Acomorus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Wal2017	malate_transferrase	Flavonoid_anthocyanin_Picalirra_PS	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size	
Acc0019116	LG06	yes	no	no	NA	NA	NA	NA									NA	NA	yes			10_Glucogenesis_specific_dayC3	150.143	7	1051	39
Acc0019211	LG13	no	yes	no	NA	NA	NA	6041									NA	NA				NA	165.267	15	2479	55
Acc0019275	LG13	no	yes	no	yes	NA	NA	NA						yes			STA_F_Flavonoid 3-hydroxylase	NA				00_#Ming2019_Anthocyanin_biosynthesis	578.333	3	1735	329
Acc0019302	LG03	yes	no	no	NA	NA	NA	NA									NA	NA				NA	895.000	2	1790	464
Acc0019314	LG03	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Glycine max cleavage and polyadenylation specificity factor subunit 1-like (LOC100801803)_transcript variant X1_mRNA				NA	187.769	26	4882	51
Acc0019315	LG03	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Glycine max cleavage and polyadenylation specificity factor subunit 1-like (LOC100801803)_transcript variant X1_mRNA				NA	245.500	12	2946	65
Acc0019356	LG05	no	yes	no	NA	NA	NA	5257									NA	NA				NA	252.167	6	1513	129
Acc0019360	LG05	yes	no	no	NA	NA	NA	NA									NA	NA				NA	1805.000	1	1805	1805
Acc0019362	LG05	yes	no	no	NA	yes	NA	NA									NA	NA				NA	431.333	3	1294	143
Acc0019382	LG05	yes	no	no	NA	NA	NA	6265									NA	NA				NA	137.000	8	1096	41
Acc0019394	LG06	yes	no	no	yes	NA	NA	NA									NA	NA				NA	149.600	30	4488	54
Acc0019493	LG03	yes	no	no	NA	NA	NA	NA									NA	NA				NA	2135.000	1	2135	2135
Acc0019534	LG07	yes	no	no	NA	NA	NA	NA						yes			NA	PREDICTED: Glycine max protein XAP5 CIRCULIAN TIMEKEEPER-like (LOC100814574)_transcript variant X3_mRNA			09_Circadian_clock	115.167	12	1382	53	
Acc0019543	LG07	no	yes	no	NA	NA	NA	NA			yes						NA	NA				02_Aquaporins	319.000	4	1276	140
Acc0019586	LG10	yes	no	no	NA	NA	NA	NA									NA	NA				NA	348.000	5	1740	140
Acc0019600	LG09	no	yes	no	yes	NA	NA	NA									NA	NA				NA	263.500	6	1581	94
Acc0019622	LG09	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Glycine max T-complex protein 1 subunit alpha (LOC100805255)_mRNA				NA	126.706	17	2154	50
Acc0019691	LG20	yes	no	no	NA	NA	NA	NA									NA	NA				NA	481.667	3	1445	343
Acc0019702	LG20	yes	no	no	NA	NA	NA	NA									NA	NA	yes			10_Glucogenesis_specific_dayC3	168.778	9	1519	62
Acc0019777	LG05	no	no	yes	NA	NA	NA	5578									NA	NA				NA	343.500	4	1374	22
Acc0019822	LG07	yes	no	no	NA	NA	NA	NA									NA	NA				NA	393.400	5	1967	135
Acc0019932	LG22	no	yes	no	yes	NA	NA	NA									NA	NA				NA	250.625	8	2005	55
Acc0019983	LG12	no	no	yes	NA	NA	NA	NA									NA	Oryza sativa Japonica Group cDNA clone:J033097.124_full insert sequence				NA	198.727	11	2186	74
Acc0020038	LG03	yes	no	no	NA	NA	NA	NA									NA	NA				NA	651.333	3	1954	348
Acc0020099	LG10	yes	no	no	yes	NA	NA	NA									NA	NA				NA	244.667	3	734	102
Acc0020108	LG10	no	yes	no	NA	NA	NA	NA		yes							NA	NA				01d_Auxin_reponse_factor	226.571	14	3172	56
Acc0020123	LG21	yes	no	no	yes	NA	NA	NA									NA	NA				NA	242.667	3	728	134
Acc0020124	LG21	no	yes	no	yes	NA	NA	NA									NA	NA				NA	637.600	5	3188	81
Acc0020155	LG09	yes	no	no	NA	NA	NA	NA			yes						NA	NA				06_Drought_resistance_genes_LEA	307.000	3	921	38

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	flavonoid_anthocyanin_P5	pos_sel_portuligo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0020162	LG09	no	yes	no	yes	NA	NA	NA									NA	NA			NA	1605.500	2	3211	712
Ac0020163	LG09	yes	no	no	NA	NA	NA	4989									NA	NA			NA	272.857	7	1910	89
Ac0020178	LG18	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1455.667	3	4367	1033
Ac0020317	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1778.000	1	1778	1778
Ac0020323	LG10	yes	no	no	NA	NA	NA	NA									NA	NA			NA	488.000	5	2440	177
Ac0020359	LG11	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1772.000	1	1772	1772
Ac0020409	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	144.429	14	2022	50
Ac0020413	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	180.000	10	1800	42
Ac0020426	LG11	NA	NA	NA	yes	NA	NA	NA									NA	NA			NA	160.333	6	962	18
Ac0020475	LG19	no	yes	no	NA	NA	NA	NA									NA	NA			NA	202.737	19	3852	47
Ac0020506	LG02	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	119.923	13	1559	51
Ac0020510	LG02	no	no	yes	yes	NA	NA	NA									NA	NA			NA	149.875	24	3597	10
Ac0020532	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	568.667	3	1706	237
Ac0020543	LG09	no	yes	no	NA	NA	yes	NA									NA	NA			NA	181.182	11	1993	83
Ac0020616	LG02	yes	no	no	NA	yes	NA	NA									NA	NA			NA	1715.000	1	1715	1715
Ac0020630	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	951.000	3	2853	189
Ac0020728	LG15	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1796.000	1	1796	1796
Ac0020733	LG15	yes	no	no	NA	NA	NA	NA									NA	NA			NA	372.500	4	1490	170
Ac0020738	LG15	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	123.583	12	1483	13
Ac0020770	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	864.000	2	1728	224
Ac0020771	LG22	no	yes	no	NA	NA	NA	NA									NA	NA			NA	224.538	13	2919	53
Ac0020775	LG22	NA	NA	NA	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	108.500	2	217	37
Ac0020809	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	287.667	6	1726	143
Ac0020843	LG09	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	218.400	5	1092	123
Ac0020881	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	3314.000	1	3314	3314
Ac0020962	LG24	no	yes	no	NA	yes	NA	NA									NA	NA			NA	113.588	17	1931	47
Ac0020989	LG13	yes	no	no	NA	NA	NA	5933									NA	NA	yes		10_Glucogenesis_specific_dayC3	306.333	3	919	207
Ac0021008	LG15	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	180.200	5	901	8
Ac0021148	LG15	yes	no	no	yes	NA	NA	NA									NA	NA			NA	523.000	2	1046	508
Ac0021151	LG15	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1232.000	1	1232	1232
Ac0021230	LG10	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	212.167	6	1273	154

gene	seqname	Single copy Broc_A nan_Tilia.	Low copy Broc_Ana n_Tilia.	High copy Broc_Ana	Diff_expressed_CA an_Tilia.	pos_selection_CAM _shifts	gene_C3 families_CAM	Angiosperm_gene	CAM- related_Acomosus- Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Ana nas	circadian_metabolis m_Walz2017	malate_transferase Assimilation_Carbohy drates	flavonoid_anthocya nin_Picalirra_PS	pos_sel_portulugo	glucogenesis_glycol ysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0021241	LG10	no	yes	no	yes	NA	NA	NA									NA	NA			NA	622.667	3	1868	349
Ac0021278	LG12	NA	NA	NA	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	171.667	3	515	53
Ac0021294	LG14	yes	no	no	NA	NA	NA	NA									NA	NA			NA	327.750	4	1311	127
Ac0021318	LG06	yes	no	no	NA	NA	NA	NA									NA	NA			NA	965.000	2	1930	459
Ac0021330	LG06	no	yes	no	NA	NA	NA	NA		yes							NA	NA			01b_Abscisic_acid_signal	244.300	10	2443	27
Ac0021372	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	277.250	4	1109	151
Ac0021378	LG16	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1142.000	1	1142	1142
Ac0021382	LG16	no	yes	no	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_response_factor	217.917	12	2615	56
Ac0021395	LG17	no	yes	no	NA	NA	NA	6373									NA	NA			NA	161.923	13	2105	59
Ac0021408	LG17	yes	no	no	yes	NA	NA	NA									NA	NA			NA	193.000	7	1351	28
Ac0021447	LG18	yes	no	no	NA	NA	NA	NA									NA	NA			NA	739.500	2	1479	166
Ac0021534	LG23	yes	no	no	NA	NA	NA	NA									NA	NA			NA	538.600	5	2693	137
Ac0021544	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	564.500	4	2258	269
Ac0021597	LG14	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	376.800	5	1884	58
Ac0021598	LG14	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	264.667	6	1588	85
Ac0021600	LG14	yes	no	no	NA	NA	NA	5849									NA	NA			NA	116.286	7	814	2
Ac0021616	LG14	yes	no	no	yes	NA	NA	NA									NA	NA			NA	223.250	4	893	6
Ac0021664	LG07	yes	no	no	NA	NA	yes	NA									NA	NA			NA	1067.000	1	1067	1067
Ac0021721	LG23	no	no	yes	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_P athway	177.800	5	889	116
Ac0021839	LG10	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Glycine max leucine-tRNA ligase_chloroplast/mitochondrial (LOC100781232)_transcript variant X1_ mRNA			NA	153.800	20	3076	56
Ac0021905	LG13	no	yes	no	yes	NA	NA	NA									NA	NA			NA	722.500	2	1445	673
Ac0021924	LG13	no	yes	no	NA	yes	NA	NA							yes		NA	00_#W_ProtonPump_VHA- H_03_VATPase			00_#W_ProtonPump_VHA- H_03_VATPase	154.417	12	1853	33
Ac0021999	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	798.333	3	2395	307
Ac0022001	LG04	yes	no	no	NA	NA	NA	5162									NA	NA			NA	495.750	4	1983	284
Ac0022068	LG05	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	161.500	6	969	5
Ac0022074	LG05	NA	NA	NA	NA	NA	NA	NA			yes						NA	06_Drought_resistance_genes_LEA			06_Drought_resistance_genes_LEA	597.500	2	1195	480
Ac0022095	LG04	yes	no	no	yes	NA	NA	NA									NA	NA			NA	196.375	8	1571	74
Ac0022109	LG04	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	750.000	2	1500	410
Ac0022112	LG04	yes	no	no	NA	NA	NA	6221									NA	NA			NA	212.000	5	1060	62
Ac0022148	LG08	no	yes	no	yes	NA	NA	NA									NA	NA			NA	194.364	11	2138	24

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed CAM	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Ming2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Wal2017	malate_transferrase	Assimilation_Carbohyd	flavonoid_anthocyanin_P5	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0022172	LG08	NA	NA	NA	NA	NA	NA	NA										ALB_F_anthocyanin_5-aromaticacyltransferase	NA			NA	428.000	3	1284	105
Ac0022185	LG08	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	1292.000	1	1292	1292
Ac0022207	LG17	no	yes	no	NA	NA	NA	NA										NA	Populus trichocarpa adenosylhomocysteinease family protein (POPTR_0017s08610g) mRNA_complete cds			NA	397.750	4	1591	45
Ac0022226	LG16	no	yes	no	NA	NA	NA	NA							yes			NA	00_#W_ProtonPump_VHA-G_03_VATPase			NA	200.500	4	802	81
Ac0022281	LG16	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	375.333	3	1126	115
Ac0022288	LG16	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	573.667	6	3442	142
Ac0022306	LG08	yes	no	no	NA	NA	NA	6284									NA	NA	NA			NA	413.750	4	1655	206
Ac0022383	LG02	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	512.333	3	1537	187
Ac0022384	LG02	no	yes	no	NA	NA	NA	NA										NA	10_Glucogenesis_specific_dayC3	yes		NA	156.917	12	1883	29
Ac0022402	LG02	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	763.333	3	2290	124
Ac0022525	LG22	yes	no	no	NA	NA	NA	NA	yes								NA	NA	00_#15_calciumdependent_proK			NA	529.600	5	2648	210
Ac0022558	LG18	yes	no	no	NA	NA	NA	5299									NA	NA	NA			NA	141.500	10	1415	28
Ac0022561	LG18	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	171.000	8	1368	30
Ac0022563	LG18	NA	NA	NA	NA	NA	NA	6449									NA	NA	NA			NA	130.167	6	781	22
Ac0022583	LG09	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	611.333	3	1834	311
Ac0022694	LG01	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	188.778	9	1699	30
Ac0022696	LG01	NA	NA	NA	NA	NA	NA	NA									NA	NA	NA			NA	572.000	1	572	572
Ac0022713	LG18	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	655.000	2	1310	606
Ac0022846	LG03	yes	no	no	NA	yes	NA	NA									NA	NA	10_Glucogenesis_specific_dayC3	yes		NA	998.000	3	2994	272
Ac0022847	LG03	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	2762.000	1	2762	2762
Ac0022851	LG03	no	yes	no	NA	NA	NA	5940									NA	NA	NA			NA	723.500	2	1447	370
Ac0022874	LG01	no	yes	no	NA	NA	NA	NA									NA	NA	10_Glucogenesis_specific_dayC3	yes		NA	139.000	9	1251	41
Ac0022875	LG01	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	1018.000	2	2036	673
Ac0022927	LG24	yes	no	no	NA	NA	NA	5328									NA	NA	NA			NA	402.667	3	1208	222
Ac0022932	LG24	yes	no	no	NA	NA	NA	6274									NA	NA	NA			NA	231.167	6	1387	51
Ac0022947	LG08	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	370.250	4	1481	114
Ac0022956	LG08	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	290.429	7	2033	126
Ac0023000	LG06	NA	NA	NA	NA	NA	NA	NA									NA	NA	10_Glucogenesis_specific_dayC3	yes		NA	291.000	2	582	182
Ac0023030	LG20	no	yes	no	NA	yes	NA	NA							yes		NA	NA	00_#W_VacuolarAcidInvertase_AI			NA	288.000	7	2016	93
Ac0023036	LG20	no	yes	no	NA	NA	NA	NA							yes		NA	NA	00_#W_VacuolarAcidInvertase_AI			NA	380.200	5	1901	96

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CAM3	pos_selection_CAM_shifts	gene_C3	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Anas	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohyd	flavonoid_anthocyanin_PScitral	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0023062	LG07	NA	NA	NA	NA	NA	NA	NA			yes						NA	NA			02_Aquaporins	173.000	3	519	96
Ac0023087	LG09	no	yes	no	NA	NA	NA	NA						yes			NA	NA			00_Ming2019_Anthocyanin_biosynthesis	666.500	2	1333	457
Ac0023089	LG09	no	yes	no	NA	NA	NA	NA						yes			NA	NA			00_Ming2019_Anthocyanin_biosynthesis	644.000	2	1288	585
Ac0023092	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	331.750	4	1327	113
Ac0023132	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1150.000	2	2300	437
Ac0023139	LG19	no	yes	no	NA	NA	NA	NA									NA	PREDICTED: Beta vulgaris subsp. vulgaris glucose-1-phosphate adenylyltransferase small subunit, chloroplast/amyloplastic (LOC104903071), mRNA	yes		10_Glucogenesis_specific_dayC3	220.667	9	1986	98
Ac0023175	LG11	yes	no	no	NA	NA	NA	NA									NA	NA			NA	454.000	4	1816	161
Ac0023190	LG10	no	yes	no	yes	NA	NA	NA									NA	NA			NA	511.500	2	1023	485
Ac0023264	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	352.250	4	1409	137
Ac0023275	LG02	yes	no	no	NA	NA	NA	NA									NA	NA			NA	296.250	4	1185	152
Ac0023302	LG16	yes	no	no	NA	NA	NA	7013									NA	NA			NA	158.300	10	1583	49
Ac0023369	LG18	no	yes	no	yes	NA	NA	NA									NA	NA			NA	311.714	7	2182	101
Ac0023410	LG24	yes	no	no	NA	NA	NA	4802									NA	NA			NA	130.400	10	1304	37
Ac0023418	LG24	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1280.000	1	1280	1280
Ac0023513	LG21	yes	no	no	NA	NA	NA	NA									NA	NA			NA	730.000	3	2190	406
Ac0023525	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	387.167	6	2323	158
Ac0023573	LG06	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	179.000	2	358	99
Ac0023597	LG09	yes	no	no	NA	NA	NA	6977									NA	NA			NA	134.375	8	1075	39
Ac0023600	LG09	no	yes	no	yes	NA	NA	NA									NA	NA			NA	206.286	7	1444	61
Ac0023607	LG09	yes	no	no	NA	NA	NA	6978									NA	NA			NA	180.125	8	1441	72
Ac0023620	LG19	yes	no	no	NA	NA	NA	NA									NA	NA			NA	736.000	2	1472	339
Ac0023631	LG24	no	yes	no	NA	NA	NA	NA							yes		NA	NA			03_VATPase	212.500	6	1275	107
Ac0023654	LG14	NA	NA	NA	NA	NA	NA	NA			yes						NA	NA			06_Drought_resistance_genes_LEA	728.000	1	728	728
Ac0023656	LG14	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1928.000	1	1928	1928
Ac0023689	LG20	yes	no	no	NA	NA	NA	NA									NA	NA			NA	562.000	2	1124	304
Ac0023810	LG20	yes	no	no	NA	NA	NA	NA									NA	NA			NA	283.200	5	1416	209
Ac0023812	LG20	no	yes	no	NA	NA	NA	NA				yes					NA	NA			06_Drought_resistance_genes_LEA	150.000	3	450	34
Ac0023961	LG22	no	yes	no	yes	NA	NA	NA									NA	NA			NA	91.400	5	457	35
Ac0024014	LG14	no	yes	no	NA	NA	NA	NA									NA	Populus trichocarpa Argonaute 4 family protein (POPTR_0001s22710g) mRNA_complete cds			NA	139.773	22	3075	31
Ac0024020	LG14	yes	no	no	NA	yes	NA	NA									NA	NA			NA	1730.000	1	1730	1730

gene	seqname	Single copy Broc_A	Low copy Broc_Ana	High copy Broc_Ana	Diff_expressed_CA	pos_selection_CAM	gene_C3	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism_Walz2017	malate_transferrase	hydrates	flavonoid_anthocyanin_Picalirina_PS	pos_sel_portulungo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Acc0024092	LG20	no	yes	no	NA	NA	NA	NA									NA	NA	NA	yes		11_Glycolysis_might_C4_malate_p athway	155.429	21	3264	71
Acc0024146	LG20	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	870.000	3	2610	181
Acc0024224	LG21	no	yes	no	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	304.800	5	1524	5
Acc0024279	LG05	no	yes	no	yes	NA	NA	NA									NA	NA	NA			NA	2252.000	1	2252	2252
Acc0024340	LG01	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	589.500	2	1179	164
Acc0024368	LG07	yes	no	no	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	142.778	9	1285	39
Acc0024370	LG11	no	no	yes	NA	NA	NA	NA									NA	NA	NA	yes		10_Glucogenesis_specific_dayC3	187.688	16	3003	95
Acc0024440	LG03	no	yes	no	NA	NA	NA	NA			yes						NA	NA	02_Aquaporins				259.714	7	1818	13
Acc0024489	LG01	yes	no	no	yes	NA	NA	NA									NA	NA	NA			NA	1064.000	1	1064	1064
Acc0024553	LG01	no	yes	no	NA	NA	NA	NA									NA	NA	10_Glucogenesis_specific_dayC3	yes			313.333	3	940	146
Acc0024634	LG01	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	741.750	4	2967	191
Acc0024663	LG19	no	yes	no	NA	NA	NA	NA						yes			NA	NA	00_#Ming2019_Anthocyanin_biosy nthesis				1388.000	1	1388	1388
Acc0024665	LG19	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	1499.000	1	1499	1499
Acc0024749	LG24	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	2072.000	1	2072	2072
Acc0024760	LG15	yes	no	no	NA	NA	NA	NA			yes						NA	NA	06_Drought_resistance_genes_LEA				523.500	2	1047	220
Acc0024818	LG04	yes	no	no	NA	NA	NA	NA	yes								NA	NA	00_#10_Glucogenesis_specific_da yC3	yes			160.263	19	3045	77
Acc0024860	LG08	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	1223.000	2	2446	201
Acc0024898	LG01	yes	no	no	NA	NA	NA	6130									NA	NA	NA			NA	136.500	10	1365	60
Acc0025098	LG10	yes	no	no	NA	NA	NA	NA									ALB_F_chalcone isomerase	NA				NA	222.600	5	1113	101
Acc0025121	LG06	no	yes	no	NA	NA	NA	NA									NA		PREDICTED: Glycane max adaggio protein 1-like (LOC100805523)_mRNA			NA	761.000	3	2283	86
Acc0025167	LG17	yes	no	no	NA	NA	NA	5656									NA	NA	NA			NA	158.000	8	1264	65
Acc0025277	LG10	no	no	yes	yes	NA	NA	NA									NA	NA	NA			NA	138.522	46	6372	15
Acc0025306	LG05	yes	no	no	NA	NA	NA	6782									NA	NA	NA			NA	214.000	5	1070	77
Acc0025351	LG06	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	452.333	3	1357	193
Acc0025372	LG20	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	1047.500	2	2095	348
Acc0025401	LG24	yes	no	no	NA	NA	NA	NA		yes							NA	NA	01d_Auxin_reponse_factor				211.786	14	2965	56
Acc0025404	LG24	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	312.143	7	2185	125
Acc0025405	LG24	yes	no	no	yes	NA	NA	NA									NA	NA	NA			NA	1205.000	1	1205	1205
Acc0025417	LG19	no	yes	no	NA	NA	NA	6176									NA	NA	NA			NA	169.231	13	2200	48
Acc0025467	LG10	yes	no	no	NA	NA	NA	NA									NA	NA	NA			NA	1637.000	1	1637	1637
Acc0025560	LG01	yes	no	no	NA	NA	NA	NA		yes							NA	NA	01b_Absciscic_acid_signal				148.522	23	3416	14

gene	seqname	Single copy Broc. A	Low copy Broc. Ana	High copy Broc. Ana	Diff. expressed CA	pos. selection CAM	gene_C3	Angiosperm_gene	CAM-related Mim2015	stomata_lunc	aquaporin_reg	draught_res	markers	anthocyanin_Sl_Ana	circadian_metabolism	malate transferase	flavonoid_anthocyanin_Ps	pos_sel_portulugo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0025623	LG02	yes	no	no	NA	NA	NA	6679									NA	NA			NA	140.375	8	1123	70
Ac0025769	LG23	no	yes	no	NA	NA	NA	5670									NA	NA			NA	162.158	19	3081	44
Ac0025773	LG23	no	no	yes	yes	NA	NA	NA									NA	NA			NA	281.625	8	2093	99
Ac0025955	LG02	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	222.000	8	1776	62
Ac0026063	LG13	no	yes	no	NA	NA	NA	NA		yes					yes		NA	NA			00_#W_SomatolMovement_NHX1/NHX2_01D_Abscasic_acid_signal	187.357	14	2623	46
Ac0026064	LG13	yes	no	no	NA	NA	NA	NA									NA	NA			NA	351.800	5	1759	145
Ac0026097	LG07	yes	no	no	NA	NA	NA	NA									NA	NA			NA	843.000	2	1686	467
Ac0026485	LG23	yes	no	no	NA	yes	NA	5857									NA	NA			NA	602.000	2	1204	509
Ac0026631	LG02	yes	no	no	NA	NA	NA	NA									NA	NA		yes	09_Circadian_clock	120.714	7	845	12
Ac0026647	LG24	yes	no	no	NA	NA	NA	6825									NA	NA			NA	279.750	4	1119	173
Ac0026648	LG24	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1749.000	2	3498	810
Ac0026652	LG17	yes	no	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_pathway	98.222	9	884	7
Ac0026809	LG05	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	159.444	9	1435	64
Ac0026869	LG08	no	yes	no	NA	NA	NA	5843									NA	NA			NA	226.250	4	905	85
Ac0027040	LG02	no	yes	no	NA	NA	NA	NA		yes							NA	NA			01d_Auxin_response_factor	139.500	4	558	70
Ac0027113	LG09	yes	no	no	NA	NA	NA	NA									NA	NA			NA	361.250	4	1445	124
Ac0027210	LG24	yes	no	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	204.889	9	1844	46
Ac0027296	LG22	no	yes	no	yes	NA	NA	NA									NA	NA			NA	162.333	3	487	69
Ac0027353	LG05	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	911.000	3	2733	173
Ac0027484	LG21	yes	no	no	NA	yes	NA	NA									NA	NA			NA	511.667	3	1535	139
Ac0027613	LG24	yes	no	no	NA	NA	NA	NA									NA	NA			NA	1343.000	1	1343	1343
Ac0027695	LG09	yes	no	no	NA	NA	NA	NA								yes	NA	NA			04_Oxygen_evolution_enhancer_OEE	326.667	3	980	295
Ac0027726	LG22	yes	no	no	NA	NA	NA	NA									NA	NA			NA	651.000	4	2604	124
Ac0027767	LG23	yes	no	no	NA	yes	NA	NA									NA	NA			NA	133.300	10	1333	35
Ac0027851	LG04	yes	no	no	NA	NA	NA	NA									NA	NA			NA	289.400	5	1447	115
Ac0027955	LG04	yes	no	no	NA	NA	NA	NA									NA	NA	yes		11_Glycolysis_night_C4_malate_pathway	667.333	6	4004	111
Ac0028383	LG05	no	no	yes	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	241.000	3	723	140
Ac0028390	LG04	yes	no	no	NA	NA	NA	6432									NA	NA			NA	223.200	5	1116	38
Ac0028555	LG21	no	yes	no	NA	NA	NA	NA									NA	NA	yes		10_Glucogenesis_specific_dayC3	206.000	5	1030	102
Ac0028694	LG01	yes	no	no	NA	NA	NA	NA									NA	NA			NA	567.750	4	2271	137

gene	seqname	Single copy Broc_A nan_Tilla.	Low copy Broc_Ana n_Tilla.	High copy Broc_Ana	an_Tilla.	Diff_expressed CAM3	pos_selection_CAM-shifts	gene_families_CAM-C3	Angiosperm_gene	CAM-related_Acomosus_Ming2015	stomata_func	aquaporin_reg	drought_res	markers	anthocyanin_Sl_Anas	circadian_metabolism_Walz2017	malate_transferrase_Assimilation_Carbohydrates	flavonoid_anthocyanin_Picalirria_PS	pos_sel_portuligo	glucogenesis_glycolysis	circadian_clock	type	average_exon_size	nr_exons_per_gene	total_exonic_size	min_exon_size
Ac0028598	LG01	yes	no	no	NA	NA	NA	NA	NA	NA								NA	NA	NA		NA	646.667	3	1940	203
Ac0029197	LG17	no	yes	no	NA	NA	NA	NA	NA	NA	yes							NA	NA	NA		01d_Auxin_reponse_factor	214.214	14	2999	56
Ac0029846	LG03	no	yes	no	NA	NA	NA	NA	NA	NA								NA	NA	NA	yes	10_Glucogenesis_specific_dayC3	227.000	3	681	83
Ac0029912	LG13	yes	no	no	NA	NA	NA	NA	NA	NA								NA	NA	NA		NA	771.500	2	1543	500
Ac0030202	LG12	yes	no	no	yes	NA	NA	NA	NA	NA								NA	NA	NA		NA	485.000	2	970	395
Ac0030327	LG07	yes	no	no	NA	NA	NA	NA	NA	NA								NA	NA	NA		NA	2255.000	1	2255	2255
Ac0030333	LG21	no	yes	no	NA	NA	NA	NA	NA	NA								NA	NA	NA	yes	10_Glucogenesis_specific_dayC3	132.500	10	1325	59
Ac0030647	LG11	yes	no	no	NA	NA	NA	NA	NA	NA								NA	NA	NA		NA	2840.000	1	2840	2840
Ac0030651	LG07	yes	no	no	NA	NA	NA	NA	NA	NA								NA	NA	NA		NA	1140.333	3	3421	295
Ac0031044	LG01	NA	NA	NA	NA	NA	NA	NA	NA	NA	yes							NA	NA	NA		01d_Auxin_reponse_factor	482.000	2	964	32
Ac0031251	LG04	no	yes	no	NA	NA	NA	yes	NA	NA								NA	NA	NA		NA	170.000	6	1020	15
Ac0031276	LG20	no	yes	no	NA	NA	NA	NA	NA	NA								NA	NA	NA	yes	10_Glucogenesis_specific_dayC3	237.500	6	1425	38
Ac0031754	LG20	yes	no	no	yes	NA	NA	NA	NA	NA								NA	NA	NA		NA	402.500	2	805	204

Legend for table S1.

gene	gene identifier	
seqname	sequence location	
Single_copy_Broc_Anan_Tilla.	present in single copy category	output of clustering of transcriptome assemblies
Low_copy_Broc_Anan_Tilla.	present in low copy category (≤ 5 copies)	output of clustering of transcriptome assemblies
High_copy_Broc_Anan_Tilla.	present in high copy category (> 5 copies)	output of clustering of transcriptome assemblies
Diff_expressed_CAMC3	Significant for differential Expression analysis genes between CAM and C3 species	de la Harpe et al., 2020
pos_selection_CAM_shifts	genes that underwent adaptive protein evolution during C3/CAM transitions	de la Harpe et al., 2021
gene_families_CAMC3	belonging to gene families with putative association to C3 vs CAM phenotype in <i>Tillandsia</i> spp	de la Harpe et al., 2021
Angiosperm_gene	orthologous to Angiosperm353 gene	citation - table 2
CAM-related_Acomosus_Ming2015	Putative CAM-related genes from pineapple genome assembly	citation - table 2
stomata_func	putatively associated with stomatal function	citation - table 2
aquaporin_reg	putatively associated with aquaporin regulation	citation - table 2
draught_res	putatively associated with drought-resistance	citation - table 2
markers	previously used marker in bromeliad phylogenies (phyc. prk, malate synthase, NADH1, PGIC, RBP2, XDH, LEAFY, AGT1, PEPC, PHYA, PHYAE/B)	citation - table 2
anthocyanin_SI_Ananas	genes regulating anthocyanin biosynthesis and self incompatibility in Ananas	citation - table 2
circadian_metabolism_Wal2017	genes coding for circadian expression or key roles in metabolism in Ananas	citation - table 2
malateTransferase_Assimilation_Carbohydrates	putatively associated with malate transferase in the vacuole, assimilation of inorganic carbon and compartmentalization of carbohydrates	citation - table 2
flavonoid_anthocyanin_Pitcairnia_PS	genes orthologous to transcripts encoding for key enzymes in the flavonoid and anthocyanin biosynthesis pathways in Pitcairnia transcriptome	citation - table 2
pos_sel_portullugo	under positive selection within the portullugo clade (Caryophyllales)	citation - table 2
glucogenesis_glycolysis	putatively associated with glucogenesis and starch synthesis	citation - table 2
circadian_clock	putatively associated with circadian clock	citation - table 2
type	defined as candidate genes, as follows:	
	1. Stomate opening/ closing genes (SOCG): 1.1 CO ₂ - sensing HIGH LEAF TEMP (HLT)	
	1.2 Abscissic acid signalling 1.3 K ⁺ /Ca ²⁺ /Cl ²⁺ channels + Na ⁺ /H ⁺ anti transporter 1.4 Auxin response factor (ARF4)	
	2. Aquaporins	
	3. V-ATPase	
	4. Oxygen evolving enhancer (OEE)	
	5. Chloroplasts activity	
	6. Drought resistance genes (LEA)	
	7. Mitochondrial pyruvate transporters	
	09. Circadian clock	
	10. Glucogenesis specific (day: C3 pathway)	
	11. Glycolysis night C4 malate pathway	

Table S2 Categories of pathways and traits used to choose genes of interest for the Bromeliad1776 bait set, including literature source and number of genes in each category.

category	number_genes	citation
Differentially expressed between CAM/C3	190	Harpe, Marylaure De La et al. 2018. "Genomic Footprints of Repeated Evolution of CAM Photosynthesis in Tillandsioid Bromeliads." IN PRESS: 495812.
showing positive selection [dN/dS] along CAM shifts	22	Harpe, Marylaure De La et al. 2018. "Genomic Footprints of Repeated Evolution of CAM Photosynthesis in Tillandsioid Bromeliads." IN PRESS: 495812.
belong in gene families putatively associated to CAM/C3 correlated trait syndrome	57	Harpe, Marylaure De La et al. 2018. "Genomic Footprints of Repeated Evolution of CAM Photosynthesis in Tillandsioid Bromeliads." IN PRESS: 495812.
Matching Angiosperm-353 set	281	Johnson, Matthew G et al. 2019. "A Universal Probe Set for Targeted Sequencing of 353 Nuclear Genes from Any Flowering Plant Designed Using K-Medoids Clustering" ed. Susanne Renner. Systematic Biology 68(4): 594–606
Purative CAM-related genes from pineapple genome assembly	31	Ming, Ray et al. 2015. "The Pineapple Genome and the Evolution of CAM Photosynthesis." Nature Genetics 47(12): 1435–42.
putatively associated with stomatal function	32	Christin, Pascal-Antoine et al. 2014. "Shared Origins of a Key Enzyme during the Evolution of C4 and CAM Metabolism." Journal of Experimental Botany 65(13): 3609–21. Winter, Klaus, and Joseph A. M. Holtum. 2014. "Facultative Crassulacean Acid Metabolism (CAM) Plants: Powerful Tools for Unravelling the Functional Elements of CAM Photosynthesis." Journal of Experimental Botany 65(13): 3425–41.
putatively associated with aquaporin regulation	21	Vera-Estrella, Rosario, Bronwyn J. Barkla, Julio C. Amezcua-Romero, and Omar Pantoja. 2012. "Day/Night Regulation of Aquaporins during the CAM Cycle in Mesembryanthemum Crystallinum." Plant, Cell & Environment 35(3): 485–501.
putatively associated with malate transferase in the vacuole, assimilation of inorganic carbon and compartmentalization of carbohydrates	32	Cosentino, Cristian et al. 2013. "Proteomic Analysis of Mesembryanthemum Crystallinum Leaf Microsomal Fractions Finds an Imbalance in V-ATPase Stoichiometry during the Salt-Induced Transition from C3 to CAM." Biochemical Journal 450(2): 407–15.
putatively associated with drought-resistance	52	Xiao, Benze, Yuenmin Huang, Ning Tang, and Lizhong Xiong. 2007. "Over-Expression of a LEA Gene in Rice Improves Drought Resistance under the Field Conditions." Theoretical and Applied Genetics 115(1): 35–46.
putatively associated with circadian clock	8	McClung, C. Robertson. 2006. "Plant Circadian Rhythms." The Plant Cell 18(4): 792–803.
putatively associated with glucogenesis and starch synthesis	288	Cushman, John C. et al. 2008. "Large-Scale mRNA Expression Profiling in the Common Ice Plant, Mesembryanthemum Crystallinum, Performing C3 Photosynthesis and Crassulacean Acid Metabolism (CAM)." Journal of Experimental Botany 59(7): 1875–94. Antony, Edna et al. 2008. "Cloning, Localization and Expression Analysis of Vacuolar Sugar Transporters in the CAM Plant Ananas Comosus (Pineapple)." Journal of Experimental Botany 59(7): 1895–1908.
under positive selection within the portulugo clade (Caryophyllales), lineage contains multiple evolutionary origins of all known photosynthesis types	86	Wada, Hajime, and Norio Murata. 2009. Lipids in Photosynthesis: Essential and Regulatory Functions. Springer Science & Business Media.
transcripts encode for key enzymes in the flavonoid and anthocyanin biosynthesis pathways in Placatinia transcriptome	119	Goolsby, Eric W. et al. 2018. "Molecular Evolution of Key Metabolic Genes during Transitions to C4 and CAM Photosynthesis." American Journal of Botany 105(3): 602–13. Palma-Silva, C., M. Ferro, M. Bacci, and A. C. Turchetto-Zolet. 2016. "De Novo Assembly and Characterization of Leaf and Floral Transcriptomes of the Hybridizing Bromeliad Species (Placatinia Spp.) Adapted to Neotropical Inseabergs." Molecular Ecology Resources 16(4): 1012–22.
circadian expression or key roles in metabolism	41	Wai, Ching M. et al. 2017. "Temporal and Spatial Transcriptomic and MicroRNA Dynamics of CAM Photosynthesis in Pineapple." The Plant Journal 92(1): 19–30.
anthocyanin biosynthesis and self incompatibility in Ananas	21	Chen, Li-Yu et al. 2019. "The Bracteatus Pineapple Genome and Domestication of Clonally Propagated Crops." Nature Genetics 51(10): 1549–58. Machado, Talita Mota et al. 2020. "Systematics of Vriesea (Bromeliaceae): Phylogenetic Relationships Based on Nuclear Gene and Partial Plastome Sequences." Botanical Journal of the Linnean Society 192(4): 656–74.
marker (phyc, prk, malate synthase, NADH1, PGIC, RBP2, XDH, LEAFY, AGT1, PEPC, PHYA, PHYAE/B)	13	Schulte, Katharina, Michael H. J. Barfuss, and Georg Zizka. 2009. "Phylogeny of Bromelioideae (Bromeliaceae) Inferred from Nuclear and Plastid DNA Loci Reveals the Evolution of the Tank Habit within the Subfamily." Molecular Phylogenetics and Evolution 51(2): 327–39. Sass, Chodon, and Chelsea D. Specht. 2010. "Phylogenetic Estimation of the Core Bromelioids with an Emphasis on the Genus Aechmea (Bromeliaceae)." Molecular Phylogenetics and Evolution 55(2): 559–71. Versieux, Leonardo M. et al. 2012. "Molecular Phylogenetics of the Brazilian Giant Bromeliads (Alcantarea, Bromeliaceae): Implications for Morphological Evolution and Biogeography." Molecular Phylogenetics and Evolution 64(1): 177–89. Heller, Sascha et al. 2017. "Barcoding of Bromeliaceae (Poales)." Genome 60(11): 943–45. Barfuss, Michael H. J. et al. 2016. "Taxonomic Revision of Bromeliaceae Subfam. Tillandsioideae Based on a Multi-Locus DNA Sequence Phylogeny and Morphology." Phytotaxa 279(1): 1–97.
all	1091	private correspondence

Table S3 List of accessions used in this study. For samples of *Tillandsia* subgenus *Tillandsia* locality codes are also indicated

sample code	DNA Bank ID	species	subfamily	tribe	locality	country	source	locality code	N_Angio353	N_Brom1776	herbarium acc.	sample_type	collector	in pop-gen analysis
		<i>Tillandsia fasciculata</i>	Tillandsioideae	Tillandsieae	Poza Azul Pinoltepec, Veracruz	Mexico	field collection	fas_VER	4	4	unknown	silicagel dried	G. Yardeni	yes
	MHJB-B1837	<i>Tillandsia fasciculata</i>	Tillandsioideae	Tillandsieae	Depto. Escuintla, Sta. Teresa Esquintla	Guatemala	HBV: WU-0025390 (WU-B97/94)	ES_GUA	1	1	WU: 0006333, 0011189, 0013562	fresh	W. Welz 3132	yes
	MHJB-B1838	<i>Tillandsia fasciculata</i>	Tillandsioideae	Tillandsieae	no data	Honduras	HBV: WU-0025334 (WU-B99B53-1)	HOND	1	1	WU: 0013708	fresh	Kamm s.n.	yes
	MHJB-B1839	<i>Tillandsia fasciculata</i>	Tillandsioideae	Tillandsieae	Depto. Sololá, Lago de Atitlán	Guatemala	HBV: WU-0026950 (WU-B103/94)	SOL_GUA	1	1	WU: 0006321	fresh	G. Noller 9106	yes
	MHJB-B1840	<i>Tillandsia fasciculata</i>	Tillandsioideae	Tillandsieae	Prov. Puntarenas, SW declivities of Cordillera de Tilaran, along the road from Sta. Elena to Rancho Grande.	Costa Rica	HBV: WU-0024657 (WU-B179/91)	PUN_CR	1	1	WU: 0013642	fresh	W. & S. Till 7116	yes
	MHJB-B1841	<i>Tillandsia fasciculata</i>	Tillandsioideae	Tillandsieae	Prov. Heredia, Barva north of Heredia	Costa Rica	HBV: WU-0024655 (WU-B99/91)	HER_CR	1	1	WU: 0013632, 0013754	fresh	W. & S. Till 7004	yes
P1, P13, P14, P15	MHJB-B360	<i>Tillandsia fasciculata</i>	Tillandsioideae	Tillandsieae	Esldo. Chiapas, Sumidero Cañon bei Tuxtla Gutierrez	Mexico	HBV: WU-0024656 (WU-B108/94)		1	1	WU: 0002124	silicagel dried	K.-D. & R. Ehlers EM890701	no
		<i>Tillandsia gymmoborhya</i>	Tillandsioideae	Tillandsieae	San Felipe Usila, Oaxaca	Mexico	field collection (Snergia)	MH124	5	7	unknown	silicagel dried	M de la Harpe	yes
		<i>Tillandsia gymmoborhya</i>	Tillandsioideae	Tillandsieae	Coatepec, Veracruz	Mexico	field collection (Snergia)	MH027	4	4	unknown	silicagel dried	M de la Harpe	yes
		<i>Tillandsia leiboldiana</i>	Tillandsioideae	Tillandsieae	Mexico, Veracruz, Cascada De Texelo	Mexico	field collection	lei_VER	2	2	unknown	silicagel dried	G. Yardeni	yes
	MHJB-B1842	<i>Tillandsia leiboldiana</i>	Tillandsioideae	Tillandsieae	Prov. Alajuela, 2 km N San Ration	Costa Rica	HBV: WU-0024715 (WU-B82/91)	ALA_CR	1	1	WU: 0003058	fresh	W. & S. Till 7112	yes
	MHJB-B323	<i>Tillandsia leiboldiana</i>	Tillandsioideae	Tillandsieae	Prov. Cartago, Turrialba, Centro Agronomico Tropical de Investigacion y Enseñanza (CATIE)	Costa Rica	HBV: WU-0000663 (WU-B84/91)	CAR_CR	1	1	WU: 0001725, 0003008	fresh	W. & S. Till 7043	yes
		<i>Tillandsia punctulata</i>	Tillandsioideae	Tillandsieae	San Andrés Tuxtla, Veracruz	Mexico	field collection (Snergia)	MH016	6	6	unknown	silicagel dried	M de la Harpe	yes
		<i>Tillandsia punctulata</i>	Tillandsioideae	Tillandsieae	San Felipe Usila, Oaxaca	Mexico	field collection (Snergia)	MH098	6	6	unknown	silicagel dried	M de la Harpe	yes
		<i>Tillandsia foliosa</i>	Tillandsioideae	Tillandsieae	Mexico, Veracruz, Cascada De Texelo	Mexico	field collection	fol_VER	4	4	unknown	silicagel dried	G. Yardeni	yes
B1832	MHJB-B1832	<i>Cipriopsis subandina</i>	Tillandsioideae	Cipuropsidinae	Depto. Amazonas, Bagua	Peru	HBV: WU-0027605 (WU-B5/95)				WU: HBV: WU-0027605 (photo)	fresh	no data	
B2	MHJB-B2	<i>Catopsis nutans</i>	Tillandsioideae	Catopsideae	Prov. San José	Costa Rica	HBV: WU-0023734 (WU-B5/95)				WU: 0007683, 0009236	fresh	E. Trauner s.n.	
B4	MHJB-B4	<i>Vriesea saundersii</i>	Tillandsioideae	Vrieseae	no data	Brazil	HBV: WU-0000137 (WU-BRO000135, WU-B144/80)				WU: 0004316, 0005675	fresh	no data (old holdings)	
B13	MHJB-B13	<i>Glomeropitcairnia penduliflora</i>	Tillandsioideae	Glomeropitcairnieae	Parish St. Joseph, road Layou - Pont Cassé	Dominica	HBV: WU-0023999 (WU-B89/96)				WU: 0015907 (photo)	fresh	W. & S. Till 12012	
B18	MHJB-B18	<i>Werauhia tarmaensis</i>	Tillandsioideae	Cipuropsidinae	Depto. Junin below Tarma	Peru	HBV: WU-0025584 (WU-B97B277-1, HEID-103077)				WU: 0008378	fresh	W. Rauh 21147a	
B22	MHJB-B22	<i>Guzmania monostachia</i>	Tillandsioideae	Tillandsieae	Trelawny Parish	Jamaica	HBV: WU-0024036 (WU-B204/91)				WU: 0001763, 0008630, 0010416	fresh	H. & I. Seethaler s.n.	
B29	MHJB-B29	<i>Catopsis juncifolia</i>	Tillandsioideae	Catopsideae	Prov. Puntarenas Bosque Esquinas near La Gamba	Costa Rica	HBV: WU-B98B1-1				WU: 0008124	fresh	W. & S. Till 15050	
B30	MHJB-B30	<i>Glomeropitcairnia erectiflora</i>	Tillandsioideae	Glomeropitcairnieae	Esldo. Sucre Peninsula Paria, Cerro la Cerbatana	Venezuela	HBV: WU-0023998 (WU-B206/96)				WU: 0008343, 0011768	fresh	F.-J. Hase s.n.	

sample code	DNA Bank ID	species	subfamily	tribe	locality	country	source	locality code	N_Angio353	N_Brom1776	herbarium acc.	sample_type	collector	in pop-gen analysis
B57	MHJB-B57	<i>Racinaea ropalocarpa</i>	Tillandsioideae	Tillandsieae	no data	Ecuador	HBV: WU-0024525 (WU-B256/96)				WU: 0003462-0003464, 0007342 WU: 0008802, 0008948, 0009245-0009247, 0012735-0012736, 0015103 WU: J. Uteley & K. Burt-Uteley 8940	fresh	H. Pirnsler s.n.	
B78	MHJB-B78	<i>Puya laxa</i>	Puyoideae		Depto. Santa Cruz	Bolivia	HBV: WU-0032184 (WU-B229/95)					fresh	K. Sasse s.n.	
B79	MHJB-B79	<i>Hechtia carlsoniae</i>	Hechtioideae		Estado. Guerrero	Mexico	HBV: WU-0024096 (WU-B178/96)					fresh	J. Uteley & K. Burt-Uteley 8940	
B90	MHJB-B90	<i>Tillandsia biflora</i>	Tillandsioideae	Tillandsieae	Estado. Lara, S Cubiro	Venezuela	HBV: WU-0024591 (WU-B266/96)				WU: 0000452, 0007348, 0013882, 0014775	fresh	F. G. Gruber	
B118	MHJB-B118	<i>Aechmea nudicaulis</i>	Bromelioideae		Prov. La Altagracia Parque Nacional del Este, S of Boca de Yuma	Dominican Republic	field collection				WU: 0011130-0011131	silicagel dried	W. Till 18094 et al.; M.H.J. Barfuss 020221-43 et al.	
B1586	MHJB-B1586	<i>Vriesea drepanocarpa</i>	Tillandsioideae	Vrieseae	Estado. Minas Gerais, near Bahia border, Santa Maria do Salto, Talismã	Brazil	Leme: 8175				RB: E.M.C. Leme 8175 & L. Kollmann	silicagel dried	E.M.C. Leme 8175 & L. Kollmann	
B1579	MHJB-B1579	<i>Waltilia hatschbachii</i>	Tillandsioideae	Vrieseae	Estado. Minas Gerais, Gouveia	Brazil	Leme: 8417				RB: E.M.C. Leme 8417 et al.	silicagel dried	E.M.C. Leme 8417 et al.	
B1604	MHJB-B1604	<i>Jagrantia monstrum</i>	Tillandsioideae	Cipuropsideinae	no data	Costa Rica	BGBM: B-021-14-74-63				B: Horich s.n.	silicagel dried	Horich s.n.	
B92	MHJB-B92	<i>Tillandsia rauhii</i>	Tillandsioideae	Tillandsieae	Depto. Lambayeque, Vally of Rio Santa (type locality)	Peru	HBV: WU-0024778 (HEID-104552)				WU: 0012008-0012018	fresh	W. Rauh 69417	
B133	MHJB-B133	<i>Cornelia sp.</i>	Lindmanioidaeae		no data	Venezuela or Colombia	field collection				WU: A. Weissenhofer s.n. (photo)	silicagel dried	A. Weissenhofer s.n.	
B148	MHJB-B148	<i>Zizkaea tuerckheimii</i>	Tillandsioideae	Cipuropsideinae	no data	Dominican Republic	BGBM: B-289-33-00-63				B: B100162215-B100162217, GH 42669	silicagel dried	G. Palm s.n.	
B150	MHJB-B150	<i>Brocchinia micrantha</i>	Brocchinioideae		Estado. Bolivar La Escalera, Ciudad Guayana - Santa Elena de Uairén, immediately NW Cerro Venado	Venezuela	HBV: WU-0020269 (WU-BRO140825, WU-B12/6)				HEID: 602934-602935	fresh	W. Rauh 69323	
B173	MHJB-B173	<i>Eduandrea seloana</i>	Bromelioideae		Estado. Minas Gerais near Palácios	Brazil	HBV: WU-0023621 (WU-B00895-2)				WU: E.M.C. Leme 1830	fresh	E.M.C. Leme 1830	
B174	MHJB-B174	<i>Deuterocohnia longipetala</i>	Pitcairnioideae		Prov. Tucumán. Depto. Trancas, 2 km N bifurcation to Gonzalo, road to San Pedro de Colalao	Argentina	HBV: WU-0025179 (WU-B23/93)				WU: W. Till 10045	fresh	W. Till 10045	
B175	MHJB-B175	<i>Deuterocohnia lotteae</i>	Pitcairnioideae		Depto. Tarija, near pass Cumbre del Condor, between Tarija and Narvaez	Bolivia	HBV: WU-B177/95				WU: W. Till 62	fresh	W. Till 62, E. Haugg & H. Hronadnik	
B381	MHJB-B381	<i>Puya mirabilis</i>	Puyoideae		no data	Bolivia or Argentina	HBV: WU-0024486 (WU-B70/96)				WU: HBV: WU-B70/96	fresh	no data	
B452	MHJB-B452	<i>Pitcairnia paraguayensis</i>	Pitcairnioideae		Depto. Amambay, ca. 11 km W Cerro Corá, main road to Yley Yau	Paraguay	HBV: WU-0024355 (WU-B70/91)				WU: W. & S. Till 6025	fresh	W. & S. Till 6025	
B543	MHJB-B543	<i>Grelgia vulcanica</i>	Bromelioideae		Prov. Carchi, Páramo del Ángel	Ecuador	field collection				QCNE: W. Till 21027	silicagel dried	W. Till 21027, M.H.J. Barfuss 060204-10 & al.	
B558	MHJB-B558	<i>Puya glomerifera</i>	Puyoideae		Prov. Pichincha, ca. 13 km above Pifo, road to Papallacta pass	Ecuador	field collection				QCNE: W. Till 21077	silicagel dried	W. Till 21077, M.H.J. Barfuss 060208-21b & al.	
B559	MHJB-B559	<i>Puya clava-herculis</i>	Puyoideae		Prov. Napo, path Papallacta pass - Termas de Papallacta	Ecuador	field collection				QCNE: W. Till 21081	silicagel dried	W. Till 21081, M.H.J. Barfuss 060208-22a & al.	
B590	MHJB-B590	<i>Pitcairnia cosangaensis</i>	Pitcairnioideae		Prov. Napo, Cordillera de los Guacamayos, road Cosanga – Cocolfillo	Ecuador	field collection				QCNE: W. Till 21123	silicagel dried	W. Till 21123, M.H.J. Barfuss 060209-48 & al.	

sample code	DNA Bank ID	species	subfamily	tribe	locality	country	source	locality code	N_Angio353	N_Brom1776	herbarium acc.	sample_type	collector	in pop-gen analysis
-------------	-------------	---------	-----------	-------	----------	---------	--------	---------------	------------	------------	----------------	-------------	-----------	---------------------

B594	MHJB-B594	<i>Aechmea zebrina</i>	Bromelioideae		Prov. Napo, Estación Biológica Jatun Sacha	Ecuador	field collection				QCNE: W. Till 21151	silicagel dried	W. Till 21151, M.H.J. Barfuss 060212-51 & al.	
B596	MHJB-B596	<i>Aechmea illandsioides</i>	Bromelioideae		Prov. Napo, Estación Biológica Jatun Sacha	Ecuador	field collection				QCNE: W. Till 21152	silicagel dried	W. Till 21152, M.H.J. Barfuss 060212-52b & al.	

legend

WU = Institutional Code (Wien Universität)
HBV = Hortus Botanicus Vindobonensis
(Botanical Garden of the University of Vienna)

N_Angio353 number of samples included in the Angiosperms353 data-set
N_Brom1776 number of samples for Bromeliad1776 data-set

localities description limited: for full locality collection data, please contact the authors

Table S4 Number of reads, numbers and percentage of read mapping to target in all samples for both bait sets.

code	species	bait_kit	Reads surviving_QC			Mapping success		Reads mapped HQ	Reads target	Percent on_target	Matrix recovered length_ATGC	Matrix recovered relative to total target
P1	<i>Tillandsia fasciculata</i>	Bromeliad1776	2669424	2153711	0.80680738616271	1926038	1763939	0.915838	1077237	0.42459		
P2	<i>Tillandsia punctulata</i>	Bromeliad1776	3114626	2509191	0.8056155056819	2234791	2034506	0.910379	1295465	0.46836		
P3	<i>Tillandsia punctulata</i>	Bromeliad1776	3031356	2416439	0.79714787705568	2145604	1968306	0.917367	1349306	0.56325		
P4	<i>Tillandsia punctulata</i>	Bromeliad1776	3077884	2521473	0.81922288169405	2239837	2077502	0.927524	1069002	0.58665		
P5	<i>Tillandsia punctulata</i>	Bromeliad1776	3260914	2704548	0.82938341826862	2411667	2268828	0.940772	1220583	0.46478		
P6	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	2270494	1874042	0.82538954077835	1662206	1566088	0.942174	1344222	0.53069		
P7	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	1630016	1347428	0.82663483057835	1187509	1117364	0.940931	1327973	0.58444		
P8	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	3565312	2912157	0.81680284923171	2606042	2397680	0.920047	751547	0.57738		
P9	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	4391420	3584750	0.8163077091237	3181017	2905918	0.913519	1334783	0.32676		
P10	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	3239560	2655668	0.81976194297991	2366246	2175200	0.919262	1337797	0.58034		
P11	<i>Tillandsia leiboldiana</i>	Bromeliad1776	2807756	2339016	0.83305529397854	2070531	1902692	0.918939	1354655	0.58165		
P12	<i>Tillandsia leiboldiana</i>	Bromeliad1776	3892048	3197777	0.82161807870818	2818092	2551399	0.905364	1294642	0.58898		
P13	<i>Tillandsia fasciculata</i>	Bromeliad1776	2899466	2210876	0.76251144176204	1941225	1740558	0.896629	1296043	0.56289		
P14	<i>Tillandsia fasciculata</i>	Bromeliad1776	2742836	2156627	0.78627632129664	1900077	1694722	0.891923	1118106	0.56350		
P15	<i>Tillandsia fasciculata</i>	Bromeliad1776	2827008	2153745	0.76184609311328	1885606	1681074	0.891530	1361624	0.48613		
P16	<i>Tillandsia punctulata</i>	Bromeliad1776	3334248	2686609	0.80576159901723	2384865	2184198	0.915858	1350965	0.59201		
P17	<i>Tillandsia punctulata</i>	Bromeliad1776	4672066	3716587	0.79549111677789	3257097	3011345	0.924549	1357563	0.58738		
P18	<i>Tillandsia punctulata</i>	Bromeliad1776	6074946	5032673	0.8284309029249	4432691	4021146	0.907157	1311902	0.59024		
P19	<i>Tillandsia punctulata</i>	Bromeliad1776	6256346	5154496	0.82388282233751	4558707	4149881	0.910320	1366049	0.57039		
P20	<i>Tillandsia punctulata</i>	Bromeliad1776	6280070	5176383	0.82425562135454	4575244	4130192	0.902726	1274360	0.59393		
P21	<i>Tillandsia punctulata</i>	Bromeliad1776	5965530	4997135	0.83766823735695	4416874	3989756	0.903299	1179268	0.55407		
P22	<i>Tillandsia punctulata</i>	Bromeliad1776	5465612	4602836	0.84214466742242	4085739	3765890	0.921716	1301415	0.51273		
P23	<i>Tillandsia punctulata</i>	Bromeliad1776	6044558	5053954	0.83611638766639	4470080	4025319	0.900503	1326100	0.56583		
P24	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	5748662	4803579	0.83559948384511	4217722	3895223	0.923537	1323329	0.57657		
P25	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	3891812	3127374	0.80357787066796	2748616	2532437	0.921350	1320943	0.57536		
P28	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	3760206	3101024	0.82469524276064	2732992	2544375	0.930985	1337747	0.57432		
P29	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	3146722	2517081	0.79990574318291	2229365	2057525	0.922920	1275798	0.58163		
P30	<i>Tillandsia foliosa</i>	Bromeliad1776	3217278	2630344	0.81756814300785	2346838	2182993	0.930185	1327392	0.55469		

code	species	bait_kit	Reads surviving_QC	Reads mapped	Mapping success	Reads mapped HQ	Reads target	Percent on target	Matrix length_ATGC	Matrix recovered relative to total target
P31	<i>Tillandsia foliosa</i>	Bromeliad1776	4473752	3551061	0.7937545487546	3141529	2845417	0.905743	1360496	0.57713
P32	<i>Tillandsia foliosa</i>	Bromeliad1776	1877560	1403495	0.74751006625621	1248078	1152773	0.923639	1339738	0.59152
P33	<i>Tillandsia foliosa</i>	Bromeliad1776	3537992	2838784	0.80237151468969	2521387	2272335	0.901224	1356962	0.58249
P34	<i>Tillandsia fasciculata</i>	Bromeliad1776	4779290	3791890	0.79340027493624	3325897	2957134	0.889124	1365932	0.58998
P35	<i>Cipuropopsis subandina</i>	Bromeliad1776	2423866	1964203	0.81035956608162	1733864	1558150	0.898658	1080812	0.59388
P36	<i>Tillandsia fasciculata</i>	Bromeliad1776	5381438	4278950	0.79513133850097	3748232	3310614	0.883247	1342997	0.46992
P37	<i>Tillandsia fasciculata</i>	Bromeliad1776	3994304	3158848	0.79083815353063	2767346	2469330	0.892310	1345898	0.58391
P38	<i>Tillandsia fasciculata</i>	Bromeliad1776	3922952	3110013	0.79277365616505	2727773	2426229	0.889454	1326968	0.58517
P39	<i>Tillandsia fasciculata</i>	Bromeliad1776	2931168	2305883	0.78667718806974	2016208	1828287	0.906795	1344054	0.57694
P40	<i>Tillandsia fasciculata</i>	Bromeliad1776	3996100	3167174	0.79256625209579	2780397	2498905	0.898758	1203965	0.58437
P41	<i>Tillandsia telboidiana</i>	Bromeliad1776	1728360	1386100	0.80197412576084	1226244	1108880	0.904290	1305328	0.52346
P42	<i>Catopsis nutans</i>	Bromeliad1776	3204928	2655716	0.82863515186613	2389550	2154929	0.901814	1330651	0.56753
P43	<i>Vriesea saundersii</i>	Bromeliad1776	3137568	2613581	0.83299581076809	2337550	2114382	0.904529	1352150	0.57854
P44	<i>Glomeropitcairnia penduliflora</i>	Bromeliad1776	2027710	1722591	0.84952532659996	1540635	1395022	0.905485	1305873	0.58789
P46	<i>Werauhia tarmaensis</i>	Bromeliad1776	1929428	1595926	0.82714980812966	1430354	1303779	0.911508	1336429	0.56777
P47	<i>Guzmania monostachia</i>	Bromeliad1776	2084610	1560393	0.74852994085225	1338932	1023982	0.764775	1263840	0.58106
P48	<i>Catopsis juncifolia</i>	Bromeliad1776	3231016	2684660	0.83090272533469	2396516	2133997	0.890458	1338541	0.54950
P49	<i>Glomeropitcairnia erectiflora</i>	Bromeliad1776	2051312	1741594	0.84901467938568	1560251	1414385	0.906511	1280979	0.58197
P50	<i>Racinaea ropalocarpa</i>	Bromeliad1776	3326574	2687245	0.80781158032258	2383630	2145572	0.900128	1320084	0.55695
P51	<i>Puya laxa</i>	Bromeliad1776	3031142	2807947	0.92636603629919	2501369	2305674	0.921765	1358608	0.57395
P52	<i>Hechtia carlsoniae</i>	Bromeliad1776	2952558	2443967	0.82774563615685	2195838	1973375	0.898689	1356860	0.59070
P53	<i>Tillandsia biflora</i>	Bromeliad1776	3098240	2400809	0.7748944562074	2114823	1827226	0.864009	1062287	0.58994
P54	<i>Aechmea nudicaulis</i>	Bromeliad1776	1678686	1532026	0.9126340482973	1196876	976173	0.815601	1346781	0.46186
P55	<i>Waltilia hatschbachii</i>	Bromeliad1776	2362680	2022955	0.85621201347622	1817506	1660250	0.913477	1365281	0.58556
P56	<i>Vriesea saundersii</i>	Bromeliad1776	3352122	2817674	0.84056427540525	2542284	2321864	0.913298	1188150	0.59360
P57	<i>Jagranthia monstrum</i>	Bromeliad1776	3236846	2708990	0.83692273280842	2423985	2203100	0.908875	1267191	0.51659
P58	<i>Tillandsia rauhii</i>	Bromeliad1776	5585238	4603005	0.824137666432156	4073484	3553616	0.872378	1358623	0.55095
P59	<i>connellia</i> sp.	Bromeliad1776	2696872	2295998	0.85135594125342	2047593	1846274	0.901680	1357153	0.59071
P60	<i>zizkaea tuerckheimii</i>	Bromeliad1776	1489336	1235610	0.82963817432735	1108414	998171	0.900540	1360100	0.59007
P61	<i>Brocchinia micrantha</i>	Bromeliad1776	2032974	1256465	0.61804282789647	1103624	968073	0.877176	1324318	0.59135

code	species	bait_kit	Reads surviving_QC	Reads mapped	Mapping success	Reads mapped HQ	Reads target	Percent on target	Matrix recovered length_ATGC	Matrix recovered relative to total target
P62	<i>Eduandrea selloana</i>	Bromeliad1776	4084300	3916783	0.95898513821218	3468313	3205162	0.924127	1224849	0.57579
P63	<i>Deuterocohnia longipetala</i>	Bromeliad1776	3688036	3296876	0.89881233444819	2949054	2676037	0.907422	1312276	0.53254
P64	<i>Deuterocohnia lotteae</i>	Bromeliad1776	3090086	2755169	0.89161563788192	2478817	2268130	0.915005	1309692	0.57055
P65	<i>Tillandsia leiboldiana</i>	Bromeliad1776	1431140	1008156	0.70444261218329	853246	653122	0.765456	1295884	0.56943
P66	<i>Puya mirabilis</i>	Bromeliad1776	2660328	2453613	0.922229717538589	2196826	2032682	0.925281	1305524	0.56343
P67	<i>Pitcairnia paraguayensis</i>	Bromeliad1776	1581380	1307309	0.82668871491988	1158854	1062553	0.916900	1340233	0.56762
P68	<i>Greigia vulcanica</i>	Bromeliad1776	1367928	1243369	0.90894330695768	1087891	980259	0.901064	1309890	0.58271
P69	<i>Puya glomerifera</i>	Bromeliad1776	3156522	2877174	0.91150132962799	2536062	2295847	0.905280	1341597	0.56952
P70	<i>Puya clava-herculis</i>	Bromeliad1776	1279850	1167196	0.91197874750947	1029551	940500	0.913505	807111	0.58330
P71	<i>Pitcairnia cosangaensis</i>	Bromeliad1776	1355932	1108289	0.81736326010449	979029	884110	0.903048	1183427	0.35092
P72	<i>Aechmea zebrina</i>	Bromeliad1776	2944188	2759139	0.93714769573139	2440682	2245036	0.919840	1259534	0.51453
P73	<i>Aechmea tillandsioides</i>	Bromeliad1776	4566162	4282734	0.93792861488489	3819633	3509240	0.918737	1325715	0.54762
P104	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	3523594	2582743	0.73298541205372	2205905	1998581	0.906014	1274725	0.57640
P105	<i>Tillandsia gymnobotrya</i>	Bromeliad1776	2837254	2166356	0.76353967603887	1867737	1710295	0.915704	1311369	0.55423
P1	<i>Tillandsia fasciculata</i>	Angiosperm353	5386732	1392546	0.2585140675274	543799	30625	0.056317		
P2	<i>Tillandsia punctulata</i>	Angiosperm353	6637144	1611675	0.24282658324123	634184	34229	0.053973		
P3	<i>Tillandsia punctulata</i>	Angiosperm353	6984846	1594717	0.22831097493059	575012	30625	0.053260		
P4	<i>Tillandsia punctulata</i>	Angiosperm353	5967368	1654093	0.27718970909788	677663	37341	0.055103		
P5	<i>Tillandsia punctulata</i>	Angiosperm353	4837242	1438587	0.29739818681803	595852	36481	0.061225		
P6	<i>Tillandsia gymnobotrya</i>	Angiosperm353	4192080	1302899	0.31080012786016	548506	37044	0.067536		
P7	<i>Tillandsia gymnobotrya</i>	Angiosperm353	3646304	1172888	0.321664896838	466786	26947	0.057729		
P8	<i>Tillandsia gymnobotrya</i>	Angiosperm353	6463272	1469809	0.22740942977489	599049	36876	0.061558		
P9	<i>Tillandsia gymnobotrya</i>	Angiosperm353	8810522	1953403	0.22171251601211	818492	48661	0.059452		
P10	<i>Tillandsia gymnobotrya</i>	Angiosperm353	6484156	1529174	0.23583238898015	665540	43349	0.065134		
P11	<i>Tillandsia leiboldiana</i>	Angiosperm353	6661442	1683405	0.2527087978849	655895	39760	0.060620		
P12	<i>Tillandsia leiboldiana</i>	Angiosperm353	12646348	3123699	0.24700403626407	1194227	61429	0.051438		
P13	<i>Tillandsia fasciculata</i>	Angiosperm353	4389156	1779661	0.4054677026745	1225200	216318	0.176557		
P14	<i>Tillandsia fasciculata</i>	Angiosperm353	876302	406872	0.46430568456993	273169	44598	0.163262		
P15	<i>Tillandsia fasciculata</i>	Angiosperm353	2595260	1034267	0.39852153541456	701653	119604	0.170460		
P16	<i>Tillandsia punctulata</i>	Angiosperm353	1281520	540493	0.42175931706099	372371	88323	0.237191		

code	species	bait_kit	Reads surviving_QC	Reads mapped	Mapping success	Reads mapped HQ	Reads target	Percent on target	Matrix recovered length_ATGC	Matrix recovered relative to total target
P17	<i>Tillandsia punctulata</i>	Angiosperm353	4631308	1872568	0.40432810773976	1232773	309219	0.250832		
P18	<i>Tillandsia punctulata</i>	Angiosperm353	1941586	867951	0.4470319625296	591947	119728	0.202261		
P19	<i>Tillandsia punctulata</i>	Angiosperm353	7613792	3306543	0.43428333739614	2308348	483102	0.209285		
P20	<i>Tillandsia punctulata</i>	Angiosperm353	7225456	3104148	0.42961274693251	2128897	409928	0.192554		
P21	<i>Tillandsia punctulata</i>	Angiosperm353	7059160	3311952	0.46917083619014	2252053	430470	0.191146		
P22	<i>Tillandsia punctulata</i>	Angiosperm353	4849036	2106478	0.43441170574935	1425622	313714	0.220054		
P23	<i>Tillandsia punctulata</i>	Angiosperm353	9264762	4297243	0.46382659371066	2960833	554806	0.187382		
P24	<i>Tillandsia gymnobotrya</i>	Angiosperm353	4845072	2209450	0.45602005501673	1505300	371053	0.246498		
P25	<i>Tillandsia gymnobotrya</i>	Angiosperm353	797228	347047	0.43531712383409	242349	66784	0.275570		
P26	<i>Tillandsia gymnobotrya</i>	Angiosperm353	950808	407605	0.42869327982095	275539	82807	0.300527		
P27	<i>Tillandsia gymnobotrya</i>	Angiosperm353	78072	29964	0.38379956962804	19232	4338	0.225562		
P28	<i>Tillandsia gymnobotrya</i>	Angiosperm353	265692	113304	0.42644866988844	78065	22347	0.286261		
P29	<i>Tillandsia gymnobotrya</i>	Angiosperm353	1257654	523333	0.4161184236682	365852	97300	0.265955		
P30	<i>Tillandsia leiboldiana</i>	Angiosperm353	2927578	1146704	0.39169033241813	763637	159776	0.209230		
P31	<i>Tillandsia leiboldiana</i>	Angiosperm353	3872700	1456270	0.37603480775686	981304	157373	0.160371		
P32	<i>Tillandsia leiboldiana</i>	Angiosperm353	2496866	685286	0.27445846112687	453351	85929	0.189542		
P33	<i>Tillandsia leiboldiana</i>	Angiosperm353	2365022	953264	0.40306770930672	667191	103932	0.155775		
P34	<i>Tillandsia fasciculata</i>	Angiosperm353	4415566	2115091	0.47900790068589	1478490	250100	0.169159		
P35	<i>Cipuropopsis subandina</i>	Angiosperm353	6089062	1793811	0.29459562080334	820576	59671	0.072718		
P36	<i>Tillandsia fasciculata</i>	Angiosperm353	5072400	2469932	0.48693557290435	1707563	272328	0.159483		
P37	<i>Tillandsia fasciculata</i>	Angiosperm353	4535562	2175739	0.47970659424345	1473579	259843	0.176335		
P38	<i>Tillandsia fasciculata</i>	Angiosperm353	4228848	2066579	0.48868604404793	1410533	230395	0.163339		
P39	<i>Tillandsia fasciculata</i>	Angiosperm353	3619738	1672241	0.46197846363466	1080756	218295	0.201984		
P40	<i>Tillandsia fasciculata</i>	Angiosperm353	4028558	1906005	0.47312338558859	1282614	234571	0.182885		
P41	<i>Tillandsia leiboldiana</i>	Angiosperm353	3082500	1245404	0.40402400648824	820664	137487	0.167531		
P42	<i>Catopsis nutans</i>	Angiosperm353	5098068	1727232	0.33880128707581	1096933	98491	0.089788		
P43	<i>Vriesea saundersii</i>	Angiosperm353	3709342	1782848	0.48063726666347	1293946	202194	0.156262		
P44	<i>Glomeropitcairnia penduliflora</i>	Angiosperm353	5660056	1884113	0.33287886197592	923369	65905	0.071375		
P46	<i>Werauhia tarmaensis</i>	Angiosperm353	4054118	1214269	0.29951496231733	700484	59068	0.084325		
P47	<i>Guzmania monostachia</i>	Angiosperm353	4062640	1842660	0.45356221570211	1315051	78222	0.059482		

code	species	bait_kit	Reads surviving_QC	Reads mapped	Mapping success	Reads mapped HQ	Reads target	Percent on target	Matrix recovered length_ATGC	Matrix recovered relative to total target
P48	<i>Catopsis juncifolia</i>	Angiosperm353	7649860	2140250	0.27977636192035	1377948	108358	0.078637		
P49	<i>Glomeropitcairnia erectiflora</i>	Angiosperm353	5687844	1846802	0.32469280099806	988144	75743	0.076652		
P50	<i>Racinaea ropalocarpa</i>	Angiosperm353	4153564	1808566	0.43542509517128	1247365	185554	0.148757		
P51	<i>Puya laxa</i>	Angiosperm353	4582336	3029801	0.66119136615037	1917582	279360	0.145683		
P52	<i>Hechtia carlsoniae</i>	Angiosperm353	4794524	2407049	0.5020412871017	1644830	246326	0.149758		
P53	<i>Tillandsia biflora</i>	Angiosperm353	3482194	1550197	0.44517824107445	1068708	131170	0.122737		
P54	<i>Aechmea nudicaulis</i>	Angiosperm353	1246890	963273	0.77254048071602	312517	14308	0.045783		
P55	<i>Waltillia hatschbachii</i>	Angiosperm353	6858722	2616032	0.3814168295493	1478586	135344	0.091536		
P56	<i>Vriesea drepanocarpa</i>	Angiosperm353	3801376	1915638	0.50393278644365	1400381	246268	0.175858		
P57	<i>Jagrantia monstrum</i>	Angiosperm353	3837150	1761575	0.45908421615001	1213960	184215	0.151747		
P58	<i>Tillandsia rauhii</i>	Angiosperm353	4267808	2354788	0.55175584281205	1643313	202852	0.123441		
P59	<i>connellia</i> sp.	Angiosperm353	4712632	2477695	0.52575609553218	1508049	233522	0.154850		
P60	<i>zizkaea tuerckheimii</i>	Angiosperm353	3918008	1120431	0.28596955391617	603743	50248	0.083228		
P61	<i>Brocchinia micrantha</i>	Angiosperm353	5612748	1684404	0.30010326492478	1131608	211401	0.186815		
P62	<i>Eduandrea seloana</i>	Angiosperm353	5337672	4158719	0.77912599350428	1768773	134230	0.075889		
P63	<i>Deuterocohnia longipetala</i>	Angiosperm353	6302958	3942026	0.62542476088211	2487901	345820	0.139001		
P64	<i>Deuterocohnia lotteae</i>	Angiosperm353	4098292	2375274	0.57957656506662	1573923	247851	0.157473		
P65	<i>Tillandsia leiboldiana</i>	Angiosperm353	2831380	1059110	0.37406141174975	742650	49798	0.067055		
P66	<i>Puya mirabilis</i>	Angiosperm353	3637370	2357148	0.64803635593849	1461202	219439	0.150177		
P67	<i>Pitcairnia paraguayensis</i>	Angiosperm353	3808146	1706160	0.44802904090337	991449	156761	0.158113		
P68	<i>Greigia vulcanica</i>	Angiosperm353	2542430	1518183	0.59713856428692	758237	75851	0.100036		
P69	<i>Puya glomerifera</i>	Angiosperm353	2669070	1566406	0.58687333041097	895673	116479	0.130046		
P70	<i>Puya clava-herculis</i>	Angiosperm353	2243284	1349521	0.60158276883355	788198	111467	0.141420		
P71	<i>Pitcairnia cosangaensis</i>	Angiosperm353	1794638	812511	0.4527436730973	450907	59534	0.132032		
P72	<i>Aechmea zebrina</i>	Angiosperm353	2134172	1349628	0.63238951687118	579454	43930	0.075813		
P73	<i>Aechmea tillandsioides</i>	Angiosperm353	5111106	3445851	0.67418891331935	1804092	147917	0.081990		

Table S5 Averaged levels of nucleotide diversity at synonymous (π_S) and non-synonymous (π_N) for 5 *Tillandsia* subgenus *Tillandsia* species.

species	π_S	π_N	π_N/π_S
<i>T. fasciculata</i>	0.008134	0.001061	0.130317
<i>T. foliosa</i>	0.004188	0.000585	0.139684
<i>T. leiboldiana</i>	0.005715	0.000716	0.125284
<i>T. punctulata</i>	0.007411	0.001039	0.140215
<i>T. gymnotrya</i>	0.006449	0.000926	0.143588

Chapter 2

Rapid radiation coupled with pervasive hybridization in the Neotropical *Tillandsia* subgenus *Tillandsia*

Gil Yardeni¹, Michael H. J. Barfuss¹, Walter Till¹, Matthew R. Thornton¹, Clara Groot
Crego¹, Christian Lexer¹, Thibault Leroy^{1,2,*}, Ovidiu Paun^{1,*}

¹*Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria.*

²*IRHS-UMR1345, Université d'Angers, INRAE, Institut Agro, SFR 4207 QuaSaV, 49071
Beaucouzé, France.*

**shared last authorship*

Status: Prepared for submission to *Systematic Biology*

Contribution: sample collection, laboratory work, data analysis, interpretation of the results,
original draft preparation.

Abstract

The recent rapid radiation of *Tillandsia* subgenus *Tillandsia* (family Bromeliaceae) provides an attractive system for studying the drivers and limitations of species diversification. This species-rich Neotropical monocot lineage consists of predominantly epiphytic plants with tremendous phenotypic diversity. So far, phylogenetic resolution has been lacking for this genus, which is thought to have diversified within the last 3 million years during its expansion from South America into Central America. We used whole-genome resequencing data to explore the evolutionary history of 32 species of *Tillandsia*, including inferring phylogenomic relationships with a tree-based and a network approach. Our results indicate that co-occurrence does not predict relatedness between lineages and reveal significant deviations from a tree-like structure, coupled with pervasive gene tree discordance. Focusing on hybridization, Patterson's D (ABBA-BABA) and related statistics were used to describe the rates and timing of introgression events, whereas topology weighting uncovered high heterogeneity of the phylogenetic signal along the genome. High rates of hybridization within and between clades suggest that the expansion of the subgenus into Central America proceeded in several migration events, followed by episodes of diversification and gene flow. Network analysis supported a reticulated history during colonization of different ecological niches. This work provides an example of how hybridization and introgression, rather than being a complicating issue for phylogenomic inference, should be investigated as a key driver of genomic processes during rapid species diversification.

Keywords: *Tillandsia*, gene tree discordance, hybridization, phylogenomics, bromeliad, Neotropical diversity, species network

Rapid radiations are characterised by accelerated diversification, usually following dispersal into novel habitats. As a diversity of ecological spaces become available, a lineage can opportunistically adapt to quickly occupy them in a process termed adaptive radiation (Hughes et al. 2015; Linder 2008; Naciri and Linder 2020; Rundell and Price 2009). While not all rapid radiations are adaptive (i.e., not necessarily associated with an increase in species diversification or ecological occupancy; Givnish 2015; Schluter 2000; Stroud and Losos 2016), they are generally associated with climatic fluctuations, habitat heterogeneity, landscape fragmentation and/or orogenesis (Boschman and Condamine 2022; Hughes and Atchison 2015; Paun et al. 2016; Qian and Ricklefs 2000; Richardson et al. 2001; Valente et al. 2010). Both ubiquitous among angiosperms in general and in the Neotropics in particular, rapid radiations provide ample opportunities to investigate the mechanisms driving plant diversification and Neotropical biodiversity in particular (Drummond et al. 2012; Lagomarsino et al. 2016; Linder 2008; Pérez-Escobar et al. 2017; Richardson et al. 2001; Soltis et al. 2019).

Phylogenomic studies of recent, rapid radiations must however overcome significant challenges. Quick diversification coincides with limited phylogenomic signal due to low sequence divergence, increased probability of short internal branches, longer terminal branches, and incomplete reproductive barriers, impeding phylogenomic resolution (Giarla and Esselstyn 2015; Straub et al. 2014; Whitfield and Lockhart 2007). Moreover, frequent morphological innovation is often associated with phylogenetic conflict, stemming from complex population-level processes like changes in population size, incomplete lineage sorting (ILS) and reticulation (Oliver 2013; Parins-Fukuchi et al. 2021). Researchers traditionally viewed the problem as analytical noise when aiming to construct bifurcating species trees, yet an increasing amount of research indicates that such episodes of phylogenomic conflict represent microevolutionary processes which are fundamental to rapid species diversification. Hence, gene-tree conflict and non-bifurcating relationships in rapid radiations should be investigated as defining features of diversification processes during rapid radiations (Filiault et al. 2018; Malinsky et al. 2018; Parins-Fukuchi et al. 2021).

Hybridization particularly influences phylogenomic signals, augmenting rates of discordance, but has emerged as a key driver of species' diversification (e.g., Abbott et al. 2013; Seehausen 2004). Historically viewed as an uncommon event during speciation and mostly accounted for secondary contact in allopatric processes, ancestral and recent gene-

flow is now regarded as a fundamental evolutionary process (Barth et al. 2020; Linck and Battey 2019; Mallet et al. 2016; Taylor and Larson 2019). Not only does speciation frequently persist in the face of hybridization, ecological diversification may be prompted through adaptive introgression, a process in which advantageous alleles are incorporated from one gene pool to another. Gene flow between recently diverged species may increase their genetic variability and enrich the genomic substrate for natural selection to act on. In addition, introgression can introduce advantageous alleles upon which selection has already acted, thus catalyzing adaptive radiations (Abbott et al. 2013; Edelman et al. 2019; Harrison and Larson 2014; Meier et al. 2019; Suarez-Gonzalez et al. 2018).

The rapidly radiating and highly diverse Neotropical genus *Tillandsia* belongs to the Bromeliaceae family. It consists of approximately 600 species of predominantly epiphytic plants distributed from southeastern United States to central Argentina. *Tillandsia* exhibit tremendous morphological diversity of adaptive traits which allowed them to occupy disparate habitats, from tropical rainforests to deserts and from lowlands to highlands (Benzing 2000; Givnish et al. 2011). Adaptations such as adventitious roots, trichomes modified for water and nutrient absorption, a water-impounding leaf rosette tank habit and Crassulacean acid metabolism (CAM) photosynthesis appear to have undergone correlated and contingent evolution, resulting in adaptive syndromes (Givnish et al. 2014).

Subgenus *Tillandsia* is one of seven subgenera in the genus and exhibits a similar morphological range as the entire genus (Barfuss et al. 2016). The phylogenetic relationships within this exceptionally young radiation (likely < 3 Mya) had been poorly resolved and its evolution remains elusive, as previous work employing conserved plastid and nuclear markers encountered issues like inconsistent relationships, “fuzzy” clade and species boundaries (Barfuss et al. 2005; 2016; Chew et al. 2010; Granados Mendoza et al. 2017; Pinzón et al. 2016). Barfuss et al. (2016) and Granados Mendoza et al. (2017) identified a monophyletic clade ‘K’ within the subgenus, with members distinctly distributed in North and Central America. Rivera (2019) expanded upon the naming scheme to describe subclades K.1 and K.2 (see Supporting Table S1). Using an ancestral area reconstruction strategy, Granados Mendoza et al. (2017) further showed that clade K’s ancestral area was likely Central and North American. Previous inferences estimated an Andean ancestor for *Tillandsia* subgenus *Tillandsia*, with continental islands formation playing an important role

in the evolution of the central American groups (Givnish et al. 2014; Till 2000; Winkler, 1990).

Here, we used whole-genome data to explore phylogenomic relationships and patterns of gene flow among 32 representative species of the recently and rapidly radiating *Tillandsia* subgenus *Tillandsia*. With sampling focused on Mexican groups within clade K, we inferred relationships between species and used a tree- and a network-based approach to shed light on the group's evolutionary history. We identify deviations from a tree structure and prevalent gene tree discordance, suggesting that phylogenomic inference in this group is complicated by the signature of microevolutionary processes. Finally, we used several variations of the D-statistic (Patterson's D) to characterise patterns of hybridization and gene flow that may have contributed to adaptive trait shift in this Neotropical rapid radiation.

Materials and Methods

Plant Material collection

We sampled a total of 69 individuals representing 33 species of *Tillandsia* subgenus *Tillandsia* and 2 species of *Tillandsia* subgenus *Allardtia* as outgroups. Within our ingroup sampling, 25 *Tillandsia* species represent the Central-American radiation (henceforth: K clades) and eight species represent South American *Tillandsia* (henceforth: SA subclades; Supporting Table S2). We sampled with the intention to represent a variety of morphological and physiological syndromes of the Mexican species (Supporting Table S1). One typical example of this physiological variation within species of the clade K is photosynthetic syndrome. The large K.2.3 clade includes species exhibiting CAM photosynthetic syndrome and varying distributions, endemic or widespread, extending into the Antilles. Clade K.2.2 is a small group of widespread C3 plants, while the other two clades, K.1 and K.2.1, include widespread or endemic species with intermediate CAM-C3 syndromes (K.1) or various CAM-like photosynthetic syndromes (K.2.1). Species within all clades are mainly epiphytic and ornithophilous. Finally, the South American sampling included epiphytic, saxicolous and tank forming xerophytic species, predominantly endemic to Peru (with the exception of *Tillandsia adpressiflora*; Supporting Tables S1-2). Voucher information for the samples is provided in Supporting Table S1. Leaves were cut lengthwise and immediately dried in

powdered silica gel. Fresh samples from the University of Vienna botanical collection were also included.

DNA library preparation

DNA extractions were performed as previously described (Yardeni et al. 2022) using a modified CTAB protocol (Doyle and Doyle, 1987) or the QIAGEN DNeasy® Plant Mini Kit (Qiagen, USA), and purified using the Nucleospin® gDNA cleanup kit from Macherey-Nagel (Hudlow et al. 2011) following the supplier's instructions. The extracts were subsequently diluted in water and quantified with a Qubit® 3.0 Fluorometer (Life Technologies, Carlsbad, CA, USA). Prior to library preparation, DNA extracts containing a maximum of 600ng of DNA were sheared at 4°C to a target average length of 400 bp using a Bioruptor® Pico sonication device (Diagenode, Denville, NJ, USA). Libraries were then prepared using a modified KAPA protocol with the KAPA LTP Library Preparation Kit (Roche, Basel, Switzerland). Sample cleaning, end repair reaction and A-tailing steps were carried out using the KAPA LTP library preparation kit following the supplier's instructions and adaptor ligation and adaptor fill-in reactions were added based on Meyer and Kircher (2010). Libraries were amplified using 8 cycles of PCR with additional cycles for samples of low amount or degraded DNA. Samples were either double indexed with a set of 60 dual-index primers, as recommended by Kircher et al. (2012) and described in Loiseau et al. (2019) or with Illumina TrueSeq PCR-free single index (Supporting Table 3). Finally, the libraries underwent a final size selection using AMPure Beads (Agencourt). Libraries were then pooled and sequenced at Vienna BioCenter Core Facilities on Illumina HiSeqV4 PE125 or on NovaSeqS1 PE150.

Data processing

The raw sequence data was demultiplexed using deML v.1.1.3 (Renaud et al. 2015) and bamtools 2.5.1 (Barnett et al. 2011). Statistics were then collected with samtools v.1.9 (Li et al. 2009). The data in BAM format was converted to fastq using bedtools v.2.29.2 (Quinlan and Hall 2010) and quality checked using fastqc v.0.11.9 (Andrews 2010). Reads were trimmed for adapter content and quality using trimalore v.0.6.5 (Krueger 2019), a wrapper tool around fastqc and cutadapt, using the following settings: --fastqc --retain unpaired. Sequence quality and adapter removal were confirmed with FastQC reports.

Quality- and adapter-trimmed reads were aligned to a *Tillandsia fasciculata* reference genome (Clara Groot Crego et al. in prep.) using bowtie2 v.2.3.5.1 (Langmead and Salzberg 2012) with the --very-sensitive-local option to increase accuracy. Low-quality mapped reads were removed, alignments were sorted by position using samtools v.1.15.1 (Li et al. 2009) and PCR duplicates were marked using MarkDuplicates from PicardTools v.2.25.2 (Picard Toolkit 2018). When a single accession was sequenced twice on different lanes, the reads were merged after confirming relatedness with inference of KING kinship coefficient, indicating a duplicate or 1st-degree relationship (kinship > 0.177; Manichaikul et al. 2010). Samples with an average coverage below 2x were removed from subsequent analysis, as well as accessions for which a high kinship coefficients was inferred: for this step, a kinship coefficient was estimated using KING and when a 1st degree relationship among accessions was indicated, only one sample was retained (see Supporting Tables S2-3).

Next, variants were called using GATK HaplotypeCaller v.4.1.9.0 followed by joint calling with GenotypeGVCFs (Poplin et al. 2018). The resulting variant call file (VCF) was filtered to exclude indels and any SNPs 3bp or closer from indels using bcftools v.1.15 (Li 2011) and GATK SelectVariants (DePristo et al. 2011). Regions reported to contain transportable elements were excluded using bedtools intersect (Quinlan and Hall 2010). We used the transition/transversion ratio as a guideline to define the filtering parameter values leading to a set of SNPs exhibiting the highest ratio using SnpSift (Cingolani et al. 2012). The following parameters were finally used: MQ < 20, genotype depth (DP) < 4, QD < 4, FS > 40, SOR > 3, MAF < 0.045 and missing rate > 0.2. Summary statistics were generated using bcftools stats (Li 2011). The full data processing pipeline and description is available at <https://github.com/giyany/TillandsiaPhylo>.

Phylogenetic tree inference

We inferred phylogenetic relationships for all samples using a genome-wide concatenated matrix and a coalescent-based method using genomic windows. To infer a genome-wide concatenated phylogeny we first converted the VCF containing only variant sites to a FASTA file using vcf2phylip with the --resolve-IUPAC option (Ortiz 2019), followed by using the raxml_ascbias.py script to remove any variant sites exclusively resulting from assignment of heterozygotes (https://github.com/btmartin721/raxml_ascbias). We used IQ-TREE 1.6.12 (Nguyen et al. 2015) to construct a maximum-likelihood tree,

allowing IQ-TREE to estimate the best-fit model and assess branch-support using ultra-fast bootstrap estimation with 1,000 replicates. After removal of species that could not be delimited (see Supplementary note regarding *T. zoquensis*), we inferred maximum-likelihood trees on a whole-genome and chromosome-by-chromosome basis.

Next, we constructed a species tree using a summary method with ASTRAL-III v.5.7.8 (hereafter: ASTRAL, Zhang et al. 2018). The VCF file was separated into non-overlapping windows of 100kb and a maximum-likelihood tree was inferred for each genomic window as described above, excluding genomic windows with fewer than 40 SNPs. Loci with insufficient signal may reduce the accuracy of species tree estimation (Mirarab 2019), hence, across all gene trees, nodes with a bootstrap support below 10 were collapsed using Newick utilities v. 1.1.0 (Junier and Zdobnov 2010). A species tree was then generated in ASTRAL, measuring branch support as posterior probability. To explore gene tree discordance we calculated overall quartet support for the main, alternative and second alternative topologies, calculated the number of concordant and discordant bipartitions on each node using phyparts (Smith et al. 2015) with a cutoff of ten for informative bootstrap values (≥ 10) and visualised gene tree discordance using the script `phypartspiecharts.py` (available from <https://github.com/mossmatters/MJPythonNotebooks>). A claudogram of all gene trees was visualised using `ggdensitree` from the R package `ggtree` v.3.4.0 (Yu 2020). Trees were modified using R packages `phytool` v.1.0-3 (Revell 2012) and `treeio` v.1.20.0 (Wang et al. 2020), with branch lengths forced into ultrametric structure for visualisation purposes only.

To explore the sources of inconsistency between phylogenies inferred with different methods and the clade relationships that differed from our expectations based on previous research (e.g., Barfuss et al. 2016), we also generated maximum-likelihood trees for subsets of the data, excluding different clades: a species tree was constructed (i) containing exclusively K clades, (ii) excluding clade K.2.1, (iii) excluding the species *T. mima* and *T. marnier-lapostollei*, which shared a branch showing inconsistent placement between inference methods.

Investigating monophyly and deviations from tree structure

To further explore the relationships within and among clades and deviations from tree structure by employing several variations of the D-statistic, we performed D-statistic tests.

Admixture between groups was inferred by exploiting asymmetries in the frequencies of nonconcordant gene trees in a three-population tree ordered (((P1,P2),P3),O) (Durand et al. 2011; Green et al. 2010). The statistic is based on the frequencies of different patterns of allele sharing where the outgroup O carries the ancestral allele A and the derived allele is denoted as B. Following this, ABBA sites are those where P2 and P3 share a derived allele while P1 has the ancestral allele, and BABA represents sites at which P1 and P3 share the derived allele while P2 has the ancestral allele. Under a null hypothesis assuming no gene flow, ABBA and BABA patterns are assumed to be equally likely due to incomplete lineage sorting, while excess gene flow between any two groups should result in significant deviations from that expectation. The deviation from the expected ratio is termed the D statistic, calculated as the differences in the count of ABBA and BABA patterns across the genome, divided by their sum: $D = [\Sigma(\text{ABBA}) - \Sigma(\text{BABA})] / [\Sigma(\text{ABBA}) + \Sigma(\text{BABA})]$. D would equal 0 under the null hypothesis, a positive D indicates an excess of ABBA patterns and accordingly, $D < 0$ indicates an excess of BABA patterns (Martin et al. 2015). Specifically, the D-statistic we implemented uses allele frequency estimates which allows to include several individuals per taxon and does not require the implicit assumption that the outgroup is fixed for the ancestral allele (Malinsky et al. 2021).

We first examined the consistency of assigning individuals to a clade by testing whether individuals assigned to the same clade always share more derived alleles with each other than with any individual from another clade by implementing a variation on the D-statistic (henceforth referred to as D_G). Following Malinsky et al., (2018) we calculated $D(A,G_1;G_2,O)$ for all possible combinations of individuals, where G_1 and G_2 are assigned to the same clade and A is assigned to a different clade. *Tillandsia complanata* was used as the outgroup for all comparisons. For this comparison and all related D-statistics hereafter, significance was assessed using a jackknife procedure with 200kb window size, and family-wise error rate (FWER) was calculated and corrected for multiple comparisons following the Holm–Bonferroni method with the p.adjust function in R (R Core Team 2020).

To further characterise the relationship among clades, we examined the consistency of monophyly of several clades within our set of window-based trees using the check_monophyly function in the ETE Toolkit v.3 (Huerta-Cepas et al. 2016). We next performed a principal components analysis (PCA) with SNPRelate v.1.20.1 (Zheng et al. 2012) on a set of biallelic SNPs, pruned to be at least 10kb apart and filtered to allow a

maximum of 10% missing data. We then used $D_{\min}$, another variation of the D-statistic (Malinsky et al. 2018), to infer if allele-sharing among clades is consistent with a tree-like structure. Assuming no known topology, $D_{\min}$ provides the minimum absolute value of D across all possible arrangements of a trio. This provides a ‘lower bound’ estimate of the amount of gene flow within each trio, so that a significantly positive $D_{\min}$ score signifies the sharing of derived alleles between the species in the trio is inconsistent with a single species tree relating them, even in the presence of incomplete lineage sorting.

Finally, to quantify and visualise the relationship among clades along the genome we used topology weighting by iterative sampling of subtrees with *Twisst* v.0.2 (Martin and Van Belleghem 2017). This method focuses on the relationships among taxa or groups and quantifies the contribution of each individual taxon topology to the full tree, enabling to locate genomic regions that are associated with certain topologies. Since large genomic windows may lead to reduced accuracy due to overestimation of the most prevalent topology and small genomic windows may lead to underestimation of the most prevalent topology, we chose windows of 50 SNPs following the strategy of Martin and Van Belleghem (2017). We focused on the three largest clades and reduced the number of possible topologies by including only the Peruvian South American, K.2.1 and K.2.3 clades. To produce subtrees, the VCF file was partitioned using the `biostar497922` script from Jvarkit (Lindenbaum 2015) and maximum-likelihood trees were inferred for each window as described above. Visualisation of the summary of weighted topologies was produced using the `plot_twisst.R` script from *Twisst* and visualisation of topologies along genomic coordinates was performed using a modified version of `plot_twisst.R` to avoid averaging the obtained signal over regions with high rates of missing data.

Characterising hybridization and introgression

To assess the rates of hybridization among all species in our data-set, we used the original implementation of the D-statistic and estimates of admixture fraction f (henceforth: f_4 -ratio) in *Dsuite* v.0.5r45 (Malinsky et al. 2021), obtaining a genome-wide signal of introgression between species. We additionally computed D-statistic for each reference chromosome, using a jackknife procedure with windows of 150 SNPs to obtain a sufficient number of blocks for computing p-values. Given a certain level of uncertainty regarding the true relationship among species, we specified no input tree to focus on estimates where

Dsuite orders each trio so that the BBAA pattern is more common than the discordant patterns, instead using a priori knowledge of taxon relationships.

Gene flow events between groups are not independent. Instead, groups that are involved in hybridization may share branches on a phylogenomic tree and a single gene-flow event can present multiple correlated instances of significantly elevated D-statistic. This is especially expected when introgression events involve ancestral lineages, hence affecting internal branches of a phylogenetic tree. With Dsuite, we estimated the f -branch metric (Malinsky et al. 2021; henceforth: $f_b(C)$), another estimator of allele sharing that was developed to create a less correlated summary. Using all f -scores on a given tree, $f_b(C)$ captures excess allele sharing involving branches and by utilising multiple f_4 -ratio calculations, provides additional information about the timing of introgression events. We next focused on hypotheses regarding hybridization events between Mexican species by inferring a species network under a maximum pseudo-likelihood approach using PhyloNet v.3.8.0 (Cao et al. 2019; Than et al. 2008; Y. Yu and Nakhleh 2015). For this analysis, we produced a reduced 16-taxon set excluding the branch containing *T. marnier-lapostollei* and *T. mima* and sampled 1,000 trees at random, inferred as previously described from non-overlapping windows of 100kb.

To investigate the signature of introgression on specific loci and to locate highly admixed loci in regions of interest (see results), we used the Dinvestigate function implemented in Dsuite, which calculates several statistics suitable for analysis in sliding genomic windows. Our analysis was performed on windows of 50 SNPs with a step size of ten SNPs. The D-statistic itself shows large variance when applied to genomic windows (Martin et al. 2015), hence we used the statistic f_{dM} , which was designed to investigate introgression in small genomic windows (Malinsky et al. 2015). While the different statistics offer different advantages, their limitations and benefits across different evolutionary scenarios have yet to be thoroughly understood (Malinsky et al. 2015; Martin et al. 2015). We chose to focus on results using the f_{dM} statistic due to its ability to better account for allele sharing between P1 and P3 - nonetheless, we found in our analysis that the different statistics were correlated and offered consistent patterns. We further used *Twisst* as previously described to explore and visualise topologies along the genome. To investigate loci on chromosome 18, where we recognized high values of f_{dM} involved in hybridization with *T. achyrostachys*, we chose to investigate patterns of hybridization between *T. punctulata*, *T.*

butzii and *T. achyrostachys* (P1, P2 and P3). Shared variation is quantified as positive values of f_{dM} when shared between P2 and P3 and as negative values when shared between P3 and P1. Finally, to gain insight into the possible functional basis of genes that were involved in ancient hybridization, we identified regions which exhibited D-statistic values exceeding the 95% quantile of the D-statistic distribution.

Results

Read mapping and variant calling

After removing samples with low coverage, high inbreeding coefficient and species with unclear circumscription, we retained for our analyses 64 accessions. The average number of reads retained per accession ranged between 22.6M and 254.5M. The average mapping rates were 89.8%, ranging between 69.4% and 97.5%, resulting in an average sequence coverage of 11.3x. Mapping rates were slightly higher for clade K.2.3 and slightly lower for the SA clade (97.5% and 92.9% on average), however the differences in mapping rates between clades were not significant (Kruskal–Wallis test, $P = 0.22$), suggesting limited biases towards the used reference genome. After variant calling and filtering we retained 2,162,143 high-quality SNPs.

Phylogenomic inference recovers several topologies

IQ-TREE determined TVMe+ASC+R2 as the best-fit substitution model - a transversion model with equal base frequencies, two substitution rates and ascertainment bias correction, which was used for maximum likelihood estimations hereafter. Coalescence-based analyses were performed on 3,785 genomic windows of 100kb.

The concatenated maximum-likelihood (ML) tree and the species tree (Supporting figure S1 and Figure 1, respectively) recovered different back-bone phylogenies for *Tillandsia* and both contradicted previously inferred phylogenies (Barfuss et al. 2016; Granados Mendoza et al. 2017; Rivera 2019). Within species clustering in both trees agreed with previously obtained assignments (Barfuss et al. 2016; Rivera 2019), however neither the South American species nor the Mexican clades formed monophyletic groups. Instead, the South American species were separated into a clade consisting of the endemic Peruvian species together with the widespread species *T. adpressiflora*, and a second clade containing

T. marnier-lapostollei and *T. mimia*, with clade K.2.1 placed as sister to the first South American clade (Supporting figure S1; Figure 1). The species tree recovered an overall similar phylogeny to the concatenated tree, but placed the branch containing *T. marnier-lapostollei* and *T. mimia* as sister to the other South American and the K.2.1 clades. Several other differences between the trees inferred with different methods concerned relationships among species within clade K.2.3: for example, in the ML tree (Supporting figure S1), a clade consisting of *T. bulbosa* and *T. ionantha* was a sister to a clade containing *T. caput-medusae* and *T. concolor*; however in the species tree the latter was sister to a different clade, containing *T. festucoides*, *T. schiedeana* and *T. juncea*. These differences were coupled with overall poor resolution of the relationships between currently recognized taxa within clade K.2.3, reflected in low bootstrap values in the concatenated tree and high levels of gene tree incongruence in the species tree (see below), suggesting substantial allele sharing and high rates of gene flow within the clade.

Gene tree discordance was widespread within the data-set, especially on internal nodes with inconsistent placement in trees inferred with different methods. Such was the case of the node representing the common ancestor of K.2.1 and all South American clades, as well as the majority of internal nodes within clade K.2.3 (Figure 1). The relationships between South American species and clade K.2.1 were characterised by many alternative topologies and often more than half of the gene tree topologies were discordant with a main topology. High levels of discordance were also found within the intermediate CAM clade K.1.

Maximum likelihood trees constructed separately for concatenated matrices of each reference chromosome retrieved many different topologies: solely considering relationships between clades, we recognized nine different topologies among the 25 different trees (see Supporting file 1). The signal from eight chromosomes recovered within-clade relationships similar to those in the concatenated whole-genome tree while several other trees presented different placements of basal branches, and while the South American species formed a monophyly in six trees, in four of them a monophyletic South American clade was placed as a sister clade to K.2.1. While each chromosome tree likely contains high amounts of gene tree discordance, the abundance of different topologies imply that several evolutionary histories can be traced along the genome in *Tillandsia* and that several genomic processes muddy the relationship between gene trees and the true species tree. Maximum-likelihood trees

constructed for subsets of the data recovered similar topologies in the first and third subset (Mexican *Tillandsia* only and excluding the branch containing *T. marnier-lapostollei* and *T. mima*), however when excluding clade K.2.1, the South American species were recovered as monophyletic (Supporting file 1).

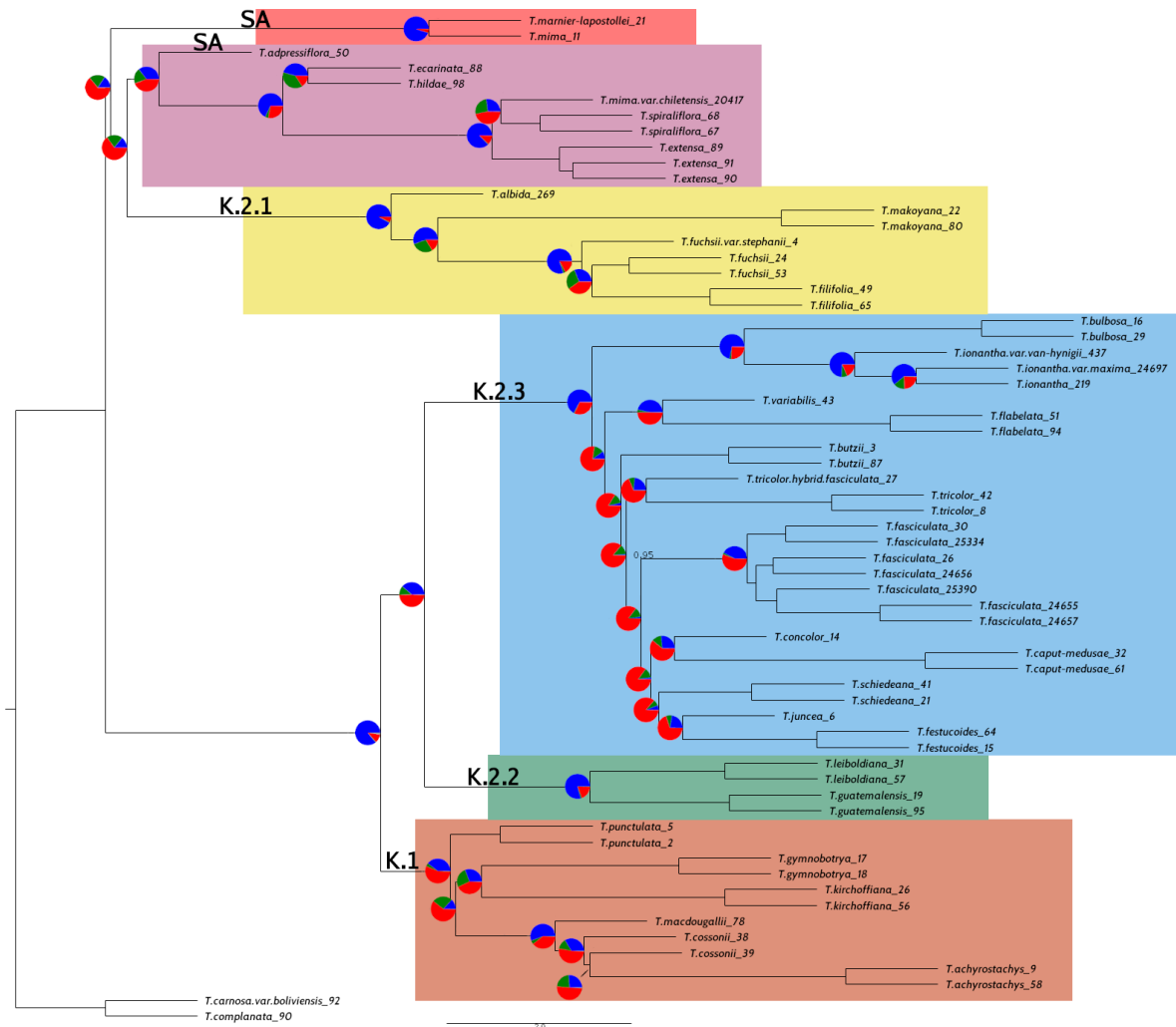


Figure 1. Coalescent-based species trees generated on 3,785 genomic windows with ASTRAL-III for 64 individuals representing 32 species of *Tillandsia* subgenus *Tillandsia*, plus two outgroups. Branch lengths are given in coalescent units. Node values represent local posterior probabilities for the main topology and are equal to one unless noted otherwise. Pie charts at the nodes show levels of gene tree discordance: the percentages of concordant gene trees (blue), the top alternative bipartition (green), other conflicting topologies (red) and uninformative gene trees (grey).

Lack of monophyly and deviations from tree-like structure

Patterns of allele sharing as tested with D_G fully supported clade assignment for all species, indicating that individual species assigned to the same clade always shared more alleles with other species within the clade than with species from other clades. Patterns of clade monophyly across the 3,785 genomic windows generally indicated consistent clade monophyly but varying relationships among clades: each of the Mexican clades were monophyletic in the majority of the windows tested with the exception of clade K.1, which was monophyletic in fewer than 40% of the window. Clades K.2.1, K.2.2, K.2.3 and K.1 were monophyletic in 3,485, 3,007, 2,537 and 1,501 trees, respectively. All the South American species in our data-set formed a monophyletic clade in a basal position in 646 trees, while the Peruvian South American group was monophyletic in 1,354 trees. Applying the same strategy to investigate relationships among clades, we recovered a group clustering all South American species and clade K.2.1 (similar to the topology of the species tree) in 587 trees, while a topology similar to that obtained in the ML tree was found in 542 trees. Finally, all K clades were together monophyletic in just 622 trees. After distance-pruning and allowing for maximum 10% missing data we retained 16,204 SNPs for the PCA analysis. The analysis provided evidence for consistent interspecific genetic structure, separating the South American group, clade K.2.1 and the remainder of the K clades. Clades K.1, K.2.2 and K.2.3 clustered closely together (Supporting figure S2).

Patterns of allele sharing indicated widespread deviations from tree-like structure, driven mostly by allele sharing between Mexican taxa. We computed the D_{min} statistic for a total of 7,141 trios (i.e., plus the outgroup) and found 3,650 comparisons with a significantly elevated value ($P < 0.05$) with more than 95% of those being highly significant ($P < 0.01$), indicating that patterns of allele sharing in 51.5% of the trios tested did not conform with a tree-like structure. The rate of elevated D_{min} values was highest for comparisons involving accessions from clade K.2.1 with 58.7% of the trios showing significant D_{min} values, followed by K.2.2 with 56.1%, K.1 with 57.9% and K.2.3 with 51.4%, respectively (Supporting figure S2). Fewer significantly elevated values included South American groups, with 42% of comparisons and 10.4% of comparisons including the South Peruvian group and the group of *T. marnier-lapostollei* and *T. mima*, respectively. In general, the majority of deviations from tree-like structure were found in trios containing species from different clades, rather than within a clade (Supporting figure S2).

We considered three distinct tree topologies for topology weighting: the most frequent topology was congruent with the one recovered in phylogenomic inference, which placed the Peruvian South American clade as sister to K.2.1 in 42% of the windows. The second topology, placing the Mexican clades as sister clades, was recovered from 35% of the genome, while a third topology placing K.2.3 as sister to the Peruvian South American clade appeared in 23% of the genome (Figure 2). The three topologies were broadly equally distributed along the genome – however, in few chromosomes the first topology frequented the centromeric regions, characterized by low recombination rates and reduced genic density (Gill et al. 2008; Yan and Jiang 2007). Topology weighing is a descriptive method which does not explicitly test for introgression or ILS, hence, it does not allow to discern the reason for a specific phylogenetic signal. Regardless, these findings suggest that the common topology represents the phylogeny’s backbone (Figure 2).

Correlated and widespread gene-flow events

We performed a total of 7,141 tests with the D-statistic measure, of which 4,331 were significantly elevated values with D values ranging between 0.021 and 0.581 (after Holm–Bonferroni correction; Figure 3). All species were involved in potential hybridizations. A prominent signal revealed introgression between species in clade K.2.1 and all other Mexican clades. Other notable signals for hybridization were indicated within clade K.2.3 and within the South American clade. *f*₄-ratio scores ranged between 0.0017 and 0.357 and indicated that in most hybridization events, the proportion of the genome involved was smaller than 10% (Supporting figure S3). A few exceptions where larger parts of the genome were admixed were found in events within clades: for example, *T. caput-medusae* was involved in two hybridization events with *T. butzii* and *T. fasciculata* involving 31.7% of the genome (see Supporting figure S3). D-statistic signal affected all chromosomes and widespread signals inferred in the whole-genome analysis generally were not located to a specific chromosome – for example, a significant D-statistic signal between clade K.2.1 and the other Mexican clades appeared on all 25 chromosomes (Supporting file 2). Notably, a significantly elevated D-statistic was found on chromosome 18 involving *T. achyrostachys* and the majority of species in the CAM clade K.2.3 (Supporting file 2). *Tillandsia achyrostachys* is a member of clade K.1, a group of generally intermediate C3-CAM species which exhibit small to intermediate tank structures and intermediary values of stable carbon isotope ratios ($\delta^{13}\text{C}$:

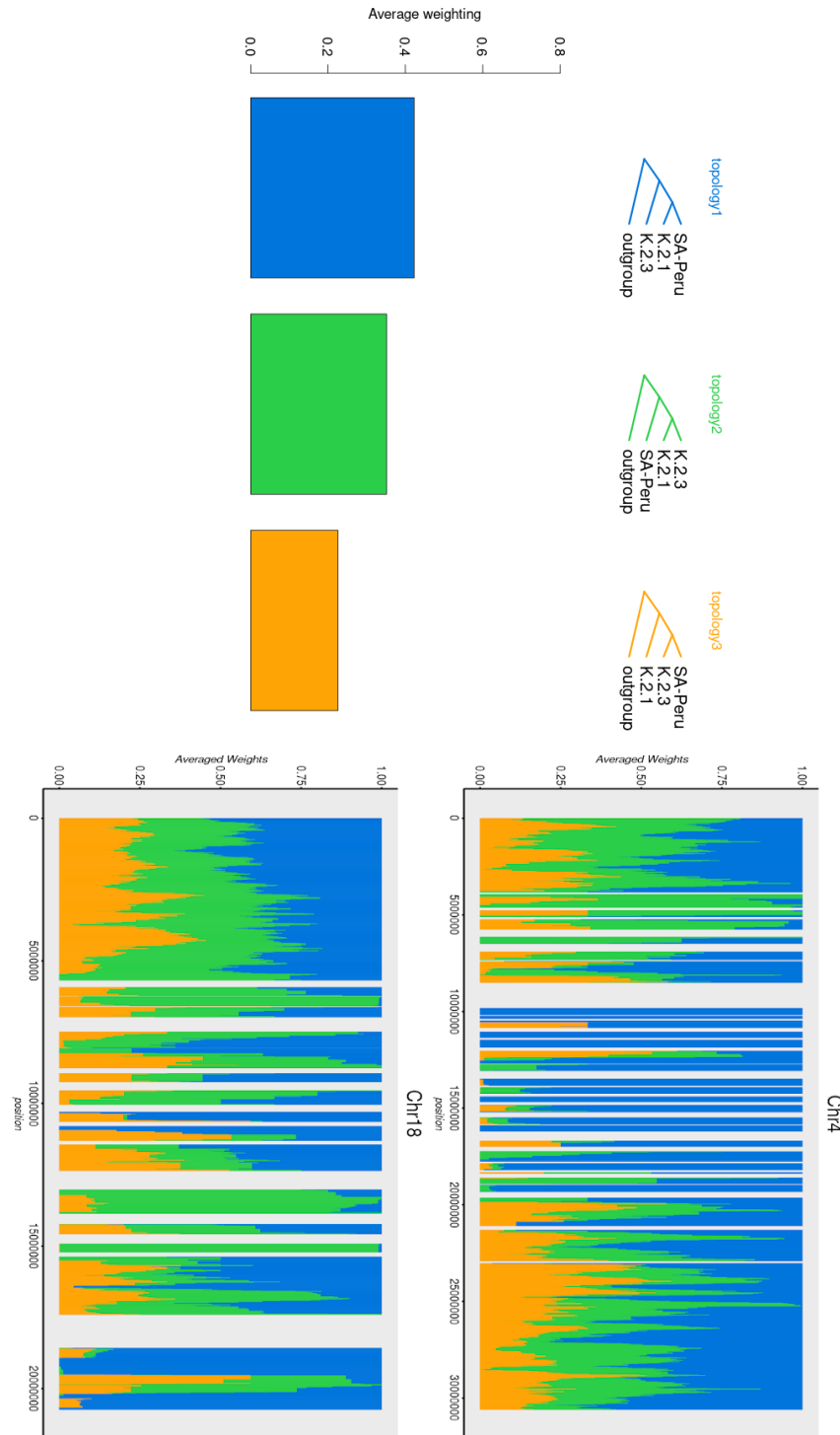


Figure 2. Variation along the genome in relationships between the major *Tillandsia* clades. Visualization was produced using topology weighting by iterative sampling of subtrees in genomic windows of 50 SNPs. Colours represent the frequency of each topology along chromosomal position in each chromosome: grey lines indicate regions excluded due to high levels of missing data, including highly repetitive regions. Two chromosome are presented: chromosome 4, to depict typical distribution of the main topologies, and chromosome 18, to contrast.

see Supporting Table S3, but see also Messerschmid et al. 2021). Within our sampling, *T. achyrostachys* exhibits the strongest CAM phenotype and $\delta^{13}\text{C}$ values (-14.7 ; Crayn et al. 2015).

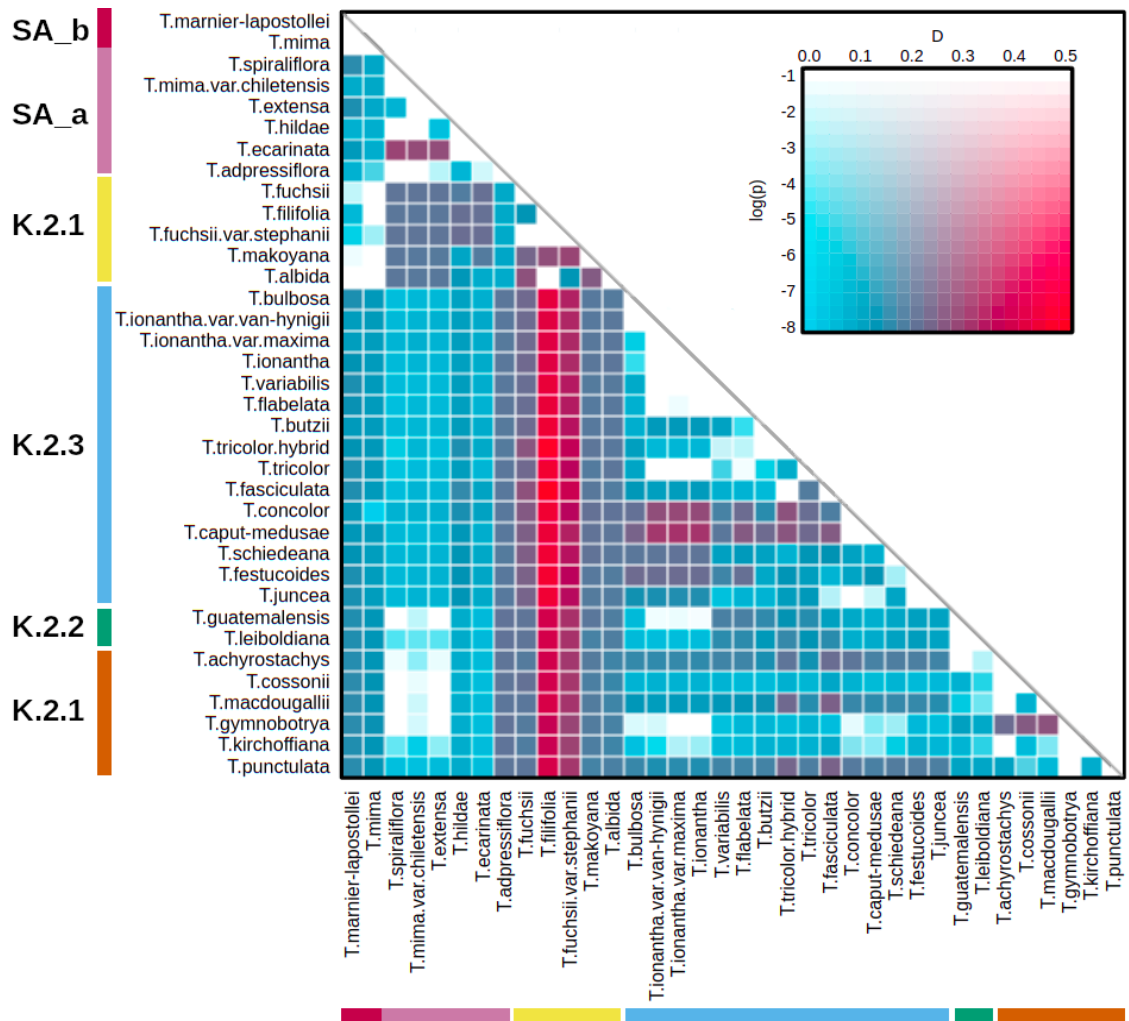


Figure 3 Heatmap summarizing 7,141 four-taxon *D*-statistic tests. *Tillandsia complanata* was used as the outgroup in all tests. The four taxa in each test have been rearranged to always obtain positive *D* values, and *P2* and *P3* are shown on the axes. Colour indicates the value of *D* and log value of *p*-value, the latter estimated using a block jackknife procedure with a 200kb window size and corrected for family wise error rate. Colours on the bars correspond to the clades in Figure 1.

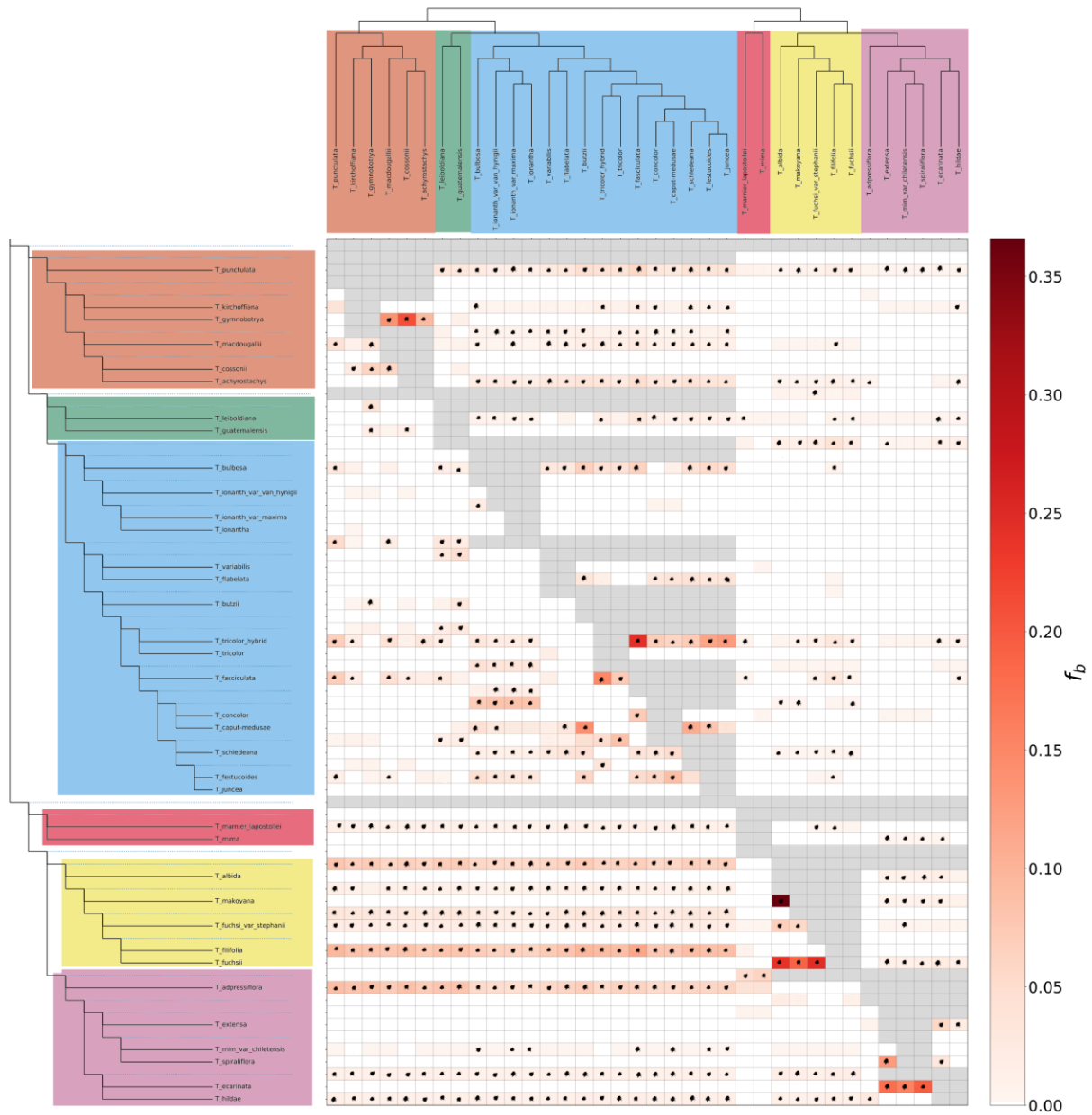


Figure 4 Heatmap summarising the statistic $f_b(C)$, where excess sharing of derived alleles is inferred between the branch of the tree on the Y axis and the species C on the X axis. The ASTRAL species tree was used as input topology for the branch statistic. The matrix is colored according to the legend of the $f_b(C)$ values and grey squares correspond to tests that are inconsistent with the ASTRAL phylogeny. Dots within the matrix denote a significant p -value, estimated using a block jackknife procedure and corrected for family wise error rate. Colours correspond to the clades in Figure 1.

The f-branch metric calculated 2,520 $f_b(C)$ scores of which 428 were significantly elevated ($P < 0.05$). 48 out of a total 70 branches showed significant excess allele sharing with at least one other species. The highest $f_b(C)$ estimate was 36.6% between *T. makoyana* and *T. albida*. Other high values were inferred especially within clades, for example between *T. gymnobotrya* and *T. cossonii* in clade K.1, and between *T. fuchsii* and three closely related species in clade K.2.1 (Figure 4). $f_b(C)$ values supported the identification of a cryptic *T. tricolor* hybrid sample (*T. tricolor.hybrid*) as a hybrid between *T. tricolor* and *T. fasciculata*. Significantly elevated scores were also inferred between the ancestor of clade K.2.1 and an ancestor of all other sampled Mexican clades with an average score of 5.9% (Figure 4). Other species that exhibited a signal of hybridization with the hypothesised ancestor of the Mexican clades are *T. filifolia* and *T. adpressiflora*, the latter being the only species within its sampled clade which is not endemic to Peru, but instead is distributed across South America from Bolivia to Northern Brazil and Colombia. Networks inferred in PhyloNet with varying numbers of reticulation events repeatedly indicated the involvement of clade K.2.1 in hybridization with the South American and Mexican clades, with higher weights assigned to the former (Figure 5). Additional reticulation events were assigned within clade K.2.3. Overall, these findings confirm the extensive violations of a tree-like structure revealed in the previous part of the analysis and suggest that these patterns are the result of many gene-flow events throughout the evolutionary history of the radiation, with high rates of gene flow between Mexican clades.

f_d and f_{dM} statistics were calculated on a total of 9,185 windows and ranged between -4.14 and 3.8 for f_d and -0.78 and 0.83 for f_{dM} . Across all chromosomes, mean values were negative at -0.19 and -0.043, respectively, fitting the expectation for higher rates of gene flow between closely related species (Supporting figure S4). Chromosome 18 was however characterized by two regions of high positive values, which were also evident when using topology weighting and coincide with regions of low genic density (Supporting figure S4). The loci exhibiting high D-statistic values contained 194 genes. Surveying the genes, we found they were associated with a variety of functions. For example, serine/threonine-protein kinase prpf4B is known to have a role in pre-mRNA splicing in yeast and humans (Eckert et al. 2016), and it is associated with stress response in millet (Parvathi et al. 2019). We also recognized S-adenosyl-L-methionine-dependent methyltransferase superfamily protein

(SAM-Mtase), a key enzyme in plant metabolic pathways like the phenylpropanoid and flavonoid pathway (Joshi and Chiang, 1998; Sistla and Rao 2004; Supporting figure S4).

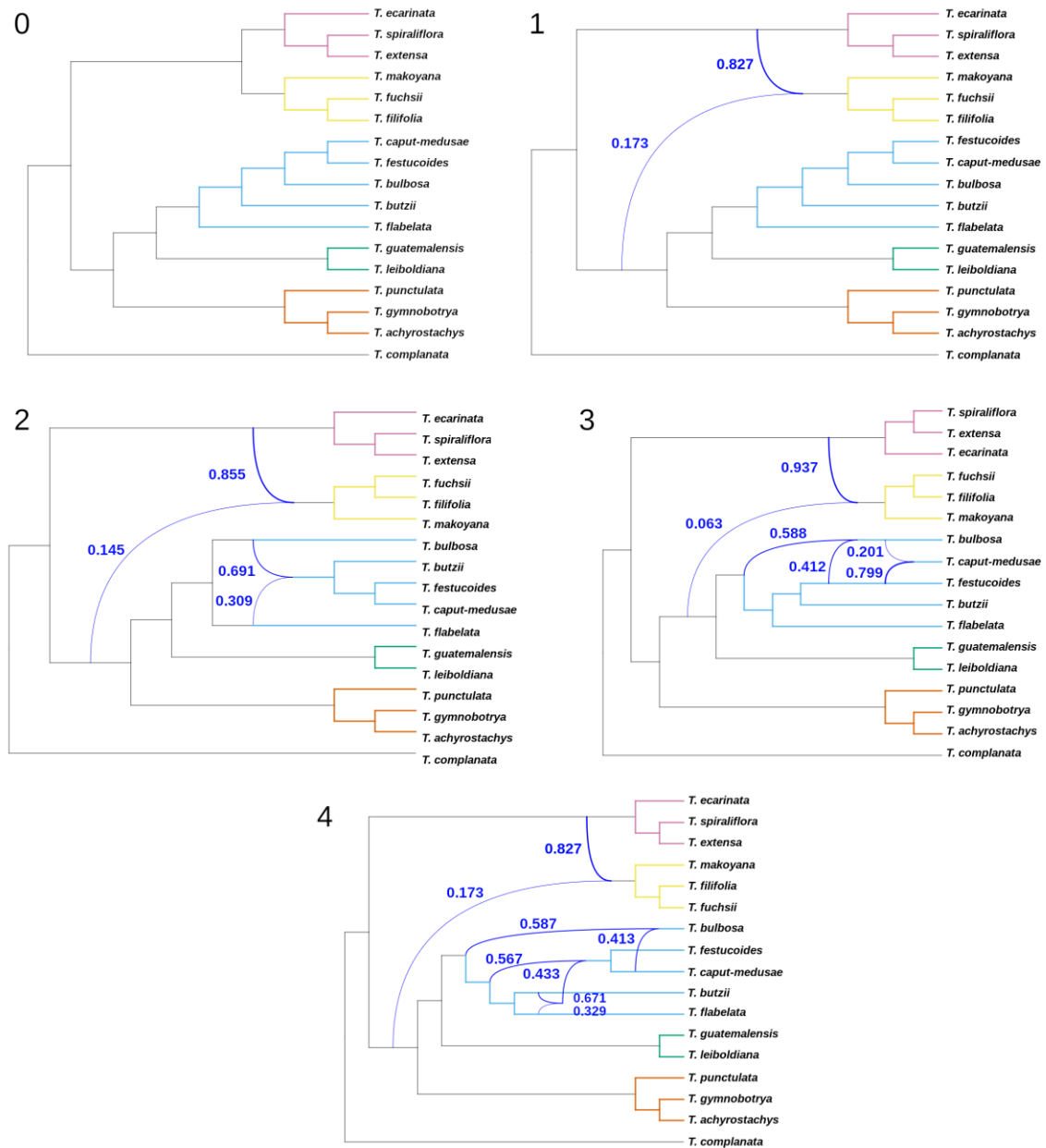


Figure 5 Best maximum pseudo-likelihood species networks inferred with PhyloNet for zero to four reticulation events. Curved branches indicate reticulation events. Numbers next to curved branches indicate inheritance probabilities for each event. Colours correspond to the clades in Figure 1.

Discussion

The generic and sub-generic classification of subfamily Tillandsioideae in general and of the genus *Tillandsia* in particular remained fluid through decades of phylogenetic research, first inferred based on morphological characters and later on a limited number of genomic markers. Our whole-genome resequencing of *Tillandsia* subgenus *Tillandsia* provides peculiar additions to our understanding of the complex evolutionary history of *Tillandsia* and offers intriguing hypotheses on the processes that accompanied diversification in this group.

The use of a large number of genomic regions spread across the genome increases the statistical confidence in resolving phylogenies while potentially revealing phylogenomic discordance (Davey et al. 2011; de La Harpe et al. 2017; Pease et al. 2018). The latter can be produced by uninformative ‘noise’ (see below) or reflect the genomic signature of cardinal molecular processes. The large amount of sequences used in the concatenated alignment and in the coalescent-based method generated high branch support for almost all nodes (Pease et al. 2016; Salichos and Rokas 2013), yet no single phylogenomic tree gave us a complete picture of the relationships between species. The branch clustering *T. mima* and *T. marnier-lapostollei* showed the most inconsistent placement between analyses. Maximum-likelihood methods perform poorly in the presence of incomplete lineage sorting (Degnan et al. 2009; Solís-Lemus et al. 2016), and the generally short internode distance coupled with high rates of discordance suggests the inconsistencies are an effect of high levels of ILS during rapid speciation (Suh et al. 2015). Regardless, ILS and hybridization can be mutually inclusive and multiple processes likely contributed to the inconsistencies along the phylogeny (Hibbins et al. 2020; Morales-Briones et al. 2021).

In stark contrast to previously inferred phylogenies, our inference retrieved the Mexican clade K as polyphyletic. Pinzón et al. (2016) sampled 108 species of *Tillandsia* with a focus on subgenus *Tillandsia* and performed analysis of plastid DNA and the external transcribed spacer (ETS) of the nuclear ribosomal DNA. In their work that also included the species *T. mima* and *T. marnier-lapostollei*, both the Mexican and South American species formed well-supported and monophyletic clades. However, they found incongruences between the plastid and ETS phylogenies and hypothesised these could be caused by plastid capture and homoplasy, in line with our main hypothesis for the incongruence between

previous studies and our phylogeny. In a phylogeny of the subfamily constructed from plastid loci and several nuclear regions, both clade K and the genus as a whole were monophyletic but poorly supported (Barfuss et al. 2016). Finally, Granados Mendoza et al. (2017) increased the sampling to represent 30% of the recognized species in the genus and constructed a phylogeny from the matK–trnK plastid region. They found a monophyletic and well-supported clade K, however the backbone of the phylogeny remained generally unresolved. Overall, phylogenies based on plastid and relatively few nuclear genomic regions offered little resolution for shallow phylogenetic in this young genus, particularly within clade K (Granados Mendoza et al. 2017).

Despite the extensive number of different topologies, a summary of the allele-sharing patterns over 3,785 genomic windows provided several points of consensus, further supported by network inference. The assignment of species to clades was highly supported as was the polyphyly of clade K, with topology weighting demonstrating that clade K.2.1 is genetically differentiated from the other K clades. Interestingly, despite their high morphological diversity, species within clades K.1, K.2.1 and K.2.2 are genetically closely related (Supporting figure S2). Each K clade formed a well supported monophyletic group and the high rates of allele sharing did not compromise clade boundaries. Departures from tree structure were recurrent and encompassed all Mexican clades and species, probably owing to the radiation's young age. Givnish et al. (2014) placed the crown age of Tillandsioideae (15.2 ± 0.42 Mya) and “core tillandsioids” (9.6 ± 0.67 Mya) in the middle or late Miocene and Barfuss et al. (2005) suggested that *Tillandsia* are phylogenetically young, based on low genetic divergence. Apart from hybridization, various molecular processes can contribute to violations of a strictly bifurcating species tree, such as ILS, paralogy and gene duplication and loss (Edwards 2009; Galtier and Daubin 2008; Smith et al. 2015). We suggest there is strong evidence for ancestral hybridization as the source of discordance between K.2.1 and other K clades: f-statistic and D-statistic tests are robust to the presence of ILS, as is PhyloNet (Malinsky et al. 2021; Martin et al. 2015; Wen et al. 2018) and ancestral population structure is unlikely to segregate through the demographic events that accompanied *Tillandsia*'s dispersal into central America (see below). $f_b(C)$ indicated that gene flow was bi-directional and involved in addition to species also internal branches or, in other words, ancestral lineages. Ancestral introgression is generally congruent with the current distribution of species in clade K.2.1, which includes several narrow endemics, likely

limiting current opportunities for extensive gene flow. The high occurrence of endemism in this clade prompts us to reject an alternative hypothesis: While the close relationship of K.2.1 and the South American clades may, in principle, be the product of recent gene flow, the clades currently share a minimal distribution. In contrast to ancestral introgression, the exceptionally high gene tree discordance within clades and specifically within the predominantly CAM clade K.2.3 probably resulted from several non-mutually exclusive processes. ILS is peculiarly associated with short inter-nodes due to rapid diversification (Galtier and Daubin 2008; Morales-Briones et al. 2021) and previous studies on *Tillandsia* subgenus *Tillandsia* found evidence for changes in population sizes and elevated rates of gene duplication and loss, specifically associated with photosynthetic syndrome shifts (de La Harpe et al. 2020; Yardeni et al. 2022).

Regardless, we present evidence that hybridization played a central role in the evolution of Mexican *Tillandsia*. Our use of D-statistics revealed prevalent recent and ancient gene flow. Notably, the obtained $f_b(C)$ ranges were in general far higher than those previously inferred for ancestral introgressions between species of Malawi cichlids (Malinsky et al. 2018) and hares (Ferreira et al. 2021), but comparable to other rapid radiations (De-Kayne et al. 2022; Suvorov et al. 2022). Networks inferred with PhyloNet repeatedly supported the hypothesis of gene flow between K.2.1 and other K clades. While the timing of the events cannot be inferred from the networks, the placement of reticulation events suggests these involve ancestral lineages. *Tillandsia* subgenus *Tillandsia* bear seeds adapted to wind dispersal and usually exhibit ornithophilous pollination syndromes, both which may facilitate high rates of inter- and intra-specific gene flow (Kessler et al. 2020; Wessinger 2021). Interestingly, the strongest $f_b(C)$ was indicated between *T. albida* and *T. makoyana* and affected an estimated 30% of the genome (Supporting figure S3), despite morphological differences and dissimilar distributions of the two species, most notably different flower colours (Espejo-Serna et al. 2004). Notable high rates of introgression with other species were also indicated for *T. filifolia*, a species with a unique flower morphology. Exhibiting small, pale lavender, non tubular flowers, this species is probably entomophilous. The generality of the obtained signal could be affected by our sampling; a wider sampling would allow us to infer if introgression involved related or ancient lineages.

Our analyses produce novel insights regarding cryptic species and the relationships between them: for example, our inference suggests that *T. mima* and *T. mima* var. *chiletensis*

are not only separate species, but also not that closely related, despite morphological similarities. Similarly, *T. fuchsii* and *T. fuchsii* var. *stephanii* did not form a monophyletic clade, raising a need to revise their species status. Interestingly, samples identified as *T. cossonii* formed a paraphyletic clade. Although both were collected at the same site in Jalisco and only 10km apart (Supporting Table S2), the samples may represent different sub-species or ecotypes, perhaps partially reproductively isolated due to differences in elevation.

Hybridization is an important factor in plant speciation and evolution, and *Tillandsia* presents an excellent opportunity to study its role in rapid diversification. Speciation in *Tillandsia* seems to have been reticulated and advanced in the presence of recurrent gene flow, rather than by refined reproductive barriers. This study therefore provides another example to the mounting evidence of the generality of diversification with gene flow (Arnold et al. 2016; Filiault et al. 2018; Keller et al. 2013; Nosil 2008; Seehausen et al. 2014; Stankowski et al. 2019; D. Zhang et al. 2021). While genetic variation introduced through hybridization can be maladaptive or neutral, the potential adaptive role of hybridization had been demonstrated in numerous animal and plant taxa (Arnold et al. 2016; Dasmahapatra et al. 2012; Lamichhaney et al. 2015; Taylor and Larson 2019). The high rates of accumulation of morphological and physiological disparity suggest that radiation in *Tillandsia* may have been adaptive: furthermore, since adaptive syndromes require the accumulation of mutations in different genic regions, hybridization through porous species boundaries may drive adaptation by continuously introducing novel alleles into populations. The current study was conducted with an incomplete taxon sampling, hence we cannot yet draw observations about adaptive introgression in *Tillandsia*. Future studies in *Tillandsia*, especially those employing wider sampling and deeper coverage, will gain by considering the role of hybridization. Intriguingly, the introgressed region we reported on chromosome 18 occurred in a region of low genic density and at a chromosome edge. Theory suggests that introgression may avoid the immediate consequences of natural selection and ‘go under the radar’ when alleles with mildly deleterious effect introgress into regions of low gene density and that barriers to introgression may be weaker towards chromosome ends (Barton and Bengtsson, 1986; Martin and Jiggins 2017; Sankararaman et al. 2014).

Our results portray a Neotropical plant radiation that proceeded in several episodes of dispersal, diversification and gene flow. The characteristic phenotypic syndrome which allowed *Tillandsia* to expand into humid and xeric habitats had been gained and lost

repeatedly, possibly involving recurrent gene flow between closely related taxa, a likely scenario especially in the light of our results. Previous studies proposed that the ancestor of *Tillandsia* subgenus *Tillandsia* may have been South American (Barfuss et al. 2016; Granados Mendoza et al. 2017), and we complement the tale by suggesting that different Mexican taxa may have had different geographical origins. It's possible that several migration events into Central America proceeded with strong population bottlenecks and episodes of gene-flow, resulting in relatively long internal branches. In summary, whole-genome resequencing data provided clear advantages over the use of few highly conserved markers yet ultimately, the complex evolutionary history of *Tillandsia* remains elusive. Future phylogenomic investigations will require a combination of analytical approaches and deeper taxonomical sampling to uncover the interplay of processes that drove this rapid diversification.

Acknowledgements

This research was supported with funding from the Christian Lexer professorship start-up BE772002 at the University of Vienna, and the FWF START grant Y661-B16 to O.P. The analyses took advantage of the Vienna Scientific Cluster (VSC). We thank Clara Groot Crego, Magnus Nordborg, James Pease and Hannes Svoldal for insightful discussions and advice. We thank the members of Swiss SNSF Sinergia project CRSII3_147630 and Juan P. Pinzón for accession sampling.

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.

- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Arnold B.J., Lahner B., DaCosta J.M., Weisman C.M., Hollister J.D., Salt D.E., Bomblies K., Yant L. 2016. Borrowed alleles and convergence in serpentine adaptation. *Proc. Natl. Acad. Sci.* 113:8320–8325.
- Barfuss M.H.J., Samuel R., Till W., Stuessy T.F. 2005. Phylogenetic relationships in subfamily Tillandsioideae (Bromeliaceae) based on DNA sequence data from seven plastid regions. *Am. J. Bot.* 92:337–351.
- Barfuss M.H.J., Till W., Leme E.M.C., Pinzón J.P., Manzanares J.M., Halbritter H., Samuel R., Brown G.K. 2016. Taxonomic revision of Bromeliaceae subfam. Tillandsioideae based on a multi-locus DNA sequence phylogeny and morphology. *Phytotaxa*. 279:1–97.
- Barnett D.W., Garrison E.K., Quinlan A.R., Strömberg M.P., Marth G.T. 2011. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics*. 27:1691–1692.
- Barth J.M.I., Gubili C., Matschiner M., Tørresen O.K., Watanabe S., Egger B., Han Y.-S., Feunteun E., Sommaruga R., Jehle R., Schabetsberger R. 2020. Stable species boundaries despite ten million years of hybridization in tropical eels. *Nat. Commun.* 11:1433.
- Barton N., Bengtsson B.O. 1986. The barrier to genetic exchange between hybridising populations. *Heredity*. 57:357–376.
- Benzing D.H. 2000. *Bromeliaceae: profile of an adaptive radiation*. Cambridge University Press.
- Boschman L.M., Condamine F.L. 2022. Mountain radiations are not only rapid and recent: Ancient diversification of South American frog and lizard families related to Paleogene Andean orogeny and Cenozoic climate variations. *Glob. Planet. Change*. 208:103704.
- Cao Z., Zhu J., Nakhleh L. 2019. Empirical performance of tree-based inference of phylogenetic networks. *19th Int. Workshop Algorithms Bioinformatics (WABI 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
- Chew T., De Luna E., González D. 2010. Phylogenetic relationships of the pseudobulbous *Tillandsia* species (Bromeliaceae) inferred from cladistic analyses of ITS 2, 5.8 S ribosomal RNA gene, and ETS sequences. *Syst. Bot.* 35:86–95.

Cingolani P., Platts A., Wang L.L., Coon M., Nguyen T., Wang L., Land S.J., Lu X., Ruden D.M. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly*. 6:80–92.

Cody S., Richardson J.E., Rull V., Ellis C., Pennington R.T. 2010. The Great American Biotic Interchange revisited. *Ecography*. 33:326–332.

Crayn D.M., Winter K., Schulte K., Smith J.A.C. 2015. Photosynthetic pathways in Bromeliaceae: phylogenetic and ecological significance of CAM and C3 based on carbon isotope ratios for 1893 species. *Bot. J. Linn. Soc.* 178:169–221.

Dasmahapatra K.K., Walters J.R., Briscoe A.D., Davey J.W., Whibley A., Nadeau N.J., Zimin A.V., Hughes D.S.T., Ferguson L.C., Martin S.H., Salazar C., Lewis J.J., Adler S., Ahn S.-J., Baker D.A., Baxter S.W., Chamberlain N.L., Chauhan R., Counterman B.A., Dalmay T., Gilbert L.E., Gordon K., Heckel D.G., Hines H.M., Hoff K.J., Holland P.W.H., Jacquin-Joly E., Jiggins F.M., Jones R.T., Kapan D.D., Kersey P., Lamas G., Lawson D., Mapleson D., Maroja L.S., Martin A., Moxon S., Palmer W.J., Papa R., Papanicolaou A., Pauchet Y., Ray D.A., Rosser N., Salzberg S.L., Supple M.A., Surridge A., Tenger-Trolander A., Vogel H., Wilkinson P.A., Wilson D., Yorke J.A., Yuan F., Balmuth A.L., Eland C., Gharbi K., Thomson M., Gibbs R.A., Han Y., Jayaseelan J.C., Kovar C., Mathew T., Muzny D.M., Onger F., Pu L.-L., Qu J., Thornton R.L., Worley K.C., Wu Y.-Q., Linares M., Blaxter M.L., French-Constant R.H., Joron M., Kronforst M.R., Mullen S.P., Reed R.D., Scherer S.E., Richards S., Mallet J., Owen McMillan W., Jiggins C.D., The Heliconius Genome Consortium. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*. 487:94–98.

Davey J.W., Hohenlohe P.A., Etter P.D., Boone J.Q., Catchen J.M., Blaxter M.L. 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12:499–510.

Degnan J.H., DeGiorgio M., Bryant D., Rosenberg N.A. 2009. Properties of consensus methods for inferring species trees from gene trees. *Syst. Biol.* 58:35–54.

De-Kayne R., Selz O.M., Marques D.A., Frei D., Seehausen O., Feulner P.G.D. 2022. Hybridization and a mixture of small and large-effect loci facilitate adaptive radiation. *bioRxiv*:2022.02.18.481029.

DePristo M.A., Banks E., Poplin R., Garimella K.V., Maguire J.R., Hartl C., Philippakis A.A., del Angel G., Rivas M.A., Hanna M., McKenna A., Fennell T.J., Kernysky A.M., Sivachenko A.Y., Cibulskis K., Gabriel S.B., Altshuler D., Daly M.J. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 43:491–498.

Doyle J.J., Doyle J.L., editors. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19:11–15.

Drummond C.S., Eastwood R.J., Miotto S.T.S., Hughes C.E. 2012. Multiple continental radiations and correlates of diversification in *Lupinus* (Leguminosae): testing for key innovation with incomplete taxon sampling. *Syst. Biol.* 61:443–460.

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

Eckert D., Andrée N., Razanau A., Zock-Emmenthal S., Lützelberger M., Plath S., Schmidt H., Guerra-Moreno A., Cozzuto L., Ayté J., Käufer N.F. 2016. Prp4 kinase grants the license to splice: control of weak splice sites during spliceosome activation. *PLOS Genet.* 12:e1005768.

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., Challis R., Kumar S., Moreira G.R.P., Salazar C., Chouteau M., Counterman B.A., Papa R., Blaxter M., Reed R.D., Dasmahapatra K.K., Kronforst M., Joron M., Jiggins C.D., McMillan W.O., Di Palma F., Blumberg A.J., Wakeley J., Jaffe D., Mallet J. 2019. Genomic architecture and introgression shape a butterfly radiation. *Science.* 366:594–599.

Edwards S.V. 2009. Is a new and general theory of molecular systematics merging? *Evolution.* 63:1–19.

Espejo-Serna A., López-Ferrari A.R., Ramírez-morillo I., Holst B.K., Luther H.E., Till W. 2004. Checklist of Mexican Bromeliaceae with notes on species distribution and levels of endemism. *Selbyana.* 25:33–86.

Ferreira M.S., Jones M.R., Callahan C.M., Farelo L., Tolesa Z., Suchentrunk F., Boursot P., Mills L.S., Alves P.C., Good J.M., Melo-Ferreira J. 2021. The legacy of recurrent introgression during the radiation of hares. *Syst. Biol.* 70:593–607.

Filiault D.L., Ballerini E.S., Mandáková T., Aköz G., Derieg N.J., Schmutz J., Jenkins J., Grimwood J., Shu S., Hayes R.D., Hellsten U., Barry K., Yan J., Mihaltcheva S., Karafiátová M., Nizhynska V., Kramer E.M., Lysak M.A., Hodges S.A., Nordborg M. 2018. The *Aquilegia* genome provides insight into adaptive radiation and reveals an extraordinarily polymorphic chromosome with a unique history. *eLife.* 7:e36426.

Galtier N., Daubin V. 2008. Dealing with incongruence in phylogenomic analyses. *Philos. Trans. R. Soc. B Biol. Sci.* 363:4023–4029.

Giarla T.C., Esselstyn J.A. 2015. The challenges of resolving a rapid, recent radiation: empirical and simulated phylogenomics of Philippine shrews. *Syst. Biol.* 64:727–740.

Gill N., Hans C.S., Jackson S. 2008. An overview of plant chromosome structure. *Cytogenet. Genome Res.* 120:194–201.

Givnish T.J. 2015. Adaptive radiation versus ‘radiation’ and ‘explosive diversification’: why conceptual distinctions are fundamental to understanding evolution. *New Phytol.* 207:297–303.

Givnish T.J., Barfuss M.H.J., Ee B.V., Riina R., Schulte K., Horres R., Gonsiska P.A., Jabaily R.S., Crayn D.M., Smith J.A.C., Winter K., Brown G.K., Evans T.M., Holst B.K., Luther H., Till W., Zizka G., Berry P.E., Sytsma K.J. 2011. Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: Insights from an eight-locus plastid phylogeny. *Am. J. Bot.* 98:872–895.

Givnish T.J., Barfuss M.H.J., Ee B.V., Riina R., Schulte K., Horres R., Gonsiska P.A., Jabaily R.S., Crayn D.M., Smith J.A.C., Winter K., Brown G.K., Evans T.M., Holst B.K., Luther H., Till W., Zizka G., Berry P.E., Sytsma K.J. 2014. Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Mol. Phylogenet. Evol.* 71:55–78.

Granados Mendoza C., Granados-Aguilar X., Donadio S., Salazar G.A., Flores-Cruz M., Hágsater E., Starr J.R., Ibarra-Manríquez G., Fragoso-Martínez I., Magallón S. 2017. Geographic structure in two highly diverse lineages of *Tillandsia* (Bromeliaceae). *Botany*. 95:641–651.

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., Malaspina 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., Gušić I., Doronichev V.B., Golovanova L.V., Lalueva-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.

Harrison R.G., Larson E.L. 2014. Hybridization, introgression, and the nature of species boundaries. *J. Hered.* 105:795–809.

- Hibbins M.S., Gibson M.J., Hahn M.W. 2020. Determining the probability of hemiplasy in the presence of incomplete lineage sorting and introgression. *eLife*. 9:e63753.
- Hudlow W.R., Krieger R., Meusel M., Sehhat J.C., Timken M.D., Buoncristiani M.R. 2011. The NucleoSpin® DNA Clean-up XS kit for the concentration and purification of genomic DNA extracts: an alternative to microdialysis filtration. *Forensic Sci. Int. Genet.* 5:226–230.
- Huerta-Cepas J., Serra F., Bork P. 2016. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol. Biol. Evol.* 33:1635–1638.
- Hughes C.E., Atchison G.W. 2015. The ubiquity of alpine plant radiations: from the Andes to the Hengduan Mountains. *New Phytol.* 207:275–282.
- Hughes C.E., Nyffeler R., Linder H.P. 2015. Evolutionary plant radiations: where, when, why and how? *New Phytol.* 207:249–253.
- Joshi C.P., Chiang V.L. 1998. Conserved sequence motifs in plant S-adenosyl-L-methionine-dependent methyltransferases. *Plant Mol. Biol.* 37:663–674.
- Junier T., Zdobnov E.M. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinforma. Oxf. Engl.* 26:1669–1670.
- Keller I., Wagner C.E., Greuter L., Mwaiko S., Selz O.M., Sivasundar A., Wittwer S., Seehausen O. 2013. Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. *Mol. Ecol.* 22:2848–2863.
- Kessler M., Abrahamczyk S., Krömer T. 2020. The role of hummingbirds in the evolution and diversification of Bromeliaceae: unsupported claims and untested hypotheses. *Bot. J. Linn. Soc.* 192:592–608.
- Kircher M., Sawyer S., Meyer M. 2012. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res.* 40:e3.
- Krueger F. 2019. Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. *Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom*.
- de La Harpe M., Paris M., Hess J., Barfuss M.H.J., Serrano-Serrano M.L., Ghatak A., Chaturvedi P., Weckwerth W., Till W., Salamin N., Wai C.M., Ming R., Lexer C. 2020. Genomic footprints of repeated evolution of CAM photosynthesis in a Neotropical species radiation. *Plant Cell Environ.* 43:2987–3001.

de La Harpe M., Paris M., Karger D.N., Rolland J., Kessler M., Salamin N., Lexer C. 2017. Molecular ecology studies of species radiations: current research gaps, opportunities and challenges. *Mol. Ecol.* 26:2608–2622.

Lagomarsino L.P., Condamine F.L., Antonelli A., Mulch A., Davis C.C. 2016. The abiotic and biotic drivers of rapid diversification in Andean bellflowers (Campanulaceae). *New Phytol.* 210:1430–1442.

Lamichhaney S., Berglund J., Almén M.S., Maqbool K., Grabherr M., Martinez-Barrio A., Promerová M., Rubin C.-J., Wang C., Zamani N., Grant B.R., Grant P.R., Webster M.T., Andersson L. 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature.* 518:371–375.

Langmead B., Salzberg S.L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* 9:357–359.

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., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics.* 25:2078–2079.

Linck E., Battey C.J. 2019. On the relative ease of speciation with periodic gene flow. *bioRxiv.*:758664.

Lindenbaum P. 2015. Jvarkit: java-based utilities for Bioinformatics. [Computer software]. <https://doi.org/10.6084/m9.figshare.1425030>

Linder H.P. 2008. Plant species radiations: where, when, why? *Philos. Trans. R. Soc. B Biol. Sci.* 363:3097–3105.

Loiseau O., Olivares I., Paris M., de La Harpe M., Weigand A., Koubínová D., Rolland J., Bacon C.D., Balslev H., Borchsenius F., Cano A., Couvreur T.L.P., Delnatte C., Fardin F., Gayot M., Mejía F., Mota-Machado T., Perret M., Roncal J., Sanin M.J., Stauffer F., Lexer C., Kessler M., Salamin N. 2019. Targeted capture of hundreds of nuclear genes unravels phylogenetic relationships of the diverse Neotropical palm tribe Geonomateae. *Front. Plant Sci.* 10:864.

- 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.
- Mallet J., Besansky N., Hahn M.W. 2016. How reticulated are species? *BioEssays*. 38:140–149.
- Manichaikul A., Mychaleckyj J.C., Rich S.S., Daly K., Sale M., Chen W.-M. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 26:2867–2873.
- Martin S.H., Davey J.W., Jiggins C.D. 2015. Evaluating the use of ABBA–BABA statistics to locate introgressed loci. *Mol. Biol. Evol.* 32:244–257.
- Martin S.H., Jiggins C.D. 2017. Interpreting the genomic landscape of introgression. *Curr. Opin. Genet. Dev.* 47:69–74.
- Martin S.H., Van Belleghem S.M. 2017. Exploring evolutionary relationships across the genome using topology weighting. *Genetics*. 206:429–438.
- Meier J.I., Stelkens R.B., Joyce D.A., Mwaiko S., Phiri N., Schlieven U.K., Selz O.M., Wagner C.E., Katongo C., Seehausen O. 2019. The coincidence of ecological opportunity with hybridization explains rapid adaptive radiation in Lake Mweru cichlid fishes. *Nat. Commun.* 10:5391.
- Messerschmid T.F.E., Wehling J., Bobon N., Kahmen A., Klak C., Los J.A., Nelson D.B., dos Santos P., de Vos J.M., Kadereit G. 2021. Carbon isotope composition of plant photosynthetic tissues reflects a Crassulacean Acid Metabolism (CAM) continuum in the majority of CAM lineages. *Perspect. Plant Ecol. Evol. Syst.* 51:125619.
- Meyer M., Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* 2010:pdb.prot5448.
- Mirarab S. 2019. Species tree estimation using ASTRAL: practical considerations. *ArXiv*:1904.03826.

- Morales-Briones D.F., Kadereit G., Tefarikis D.T., Moore M.J., Smith S.A., Brockington S.F., Timoneda A., Yim W.C., Cushman J.C., Yang Y. 2021. Disentangling sources of gene tree discordance in phylogenomic data sets: testing ancient hybridizations in Amaranthaceae s.l. *Syst. Biol.* 70:219–235.
- Naciri Y., Linder H.P. 2020. The genetics of evolutionary radiations. *Biol. Rev.* 95:1055–1072.
- Nguyen L.-T., 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.
- Nosil P. 2008. Speciation with gene flow could be common. *Mol Ecol.* 17:2103–2106.
- Oliver J.C. 2013. Microevolutionary processes generate phylogenomic discordance at ancient divergences. *Evolution.* 67:1823–1830.
- Ortiz, E.M. 2019. vcf2phyliip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis. DOI:10.5281/zenodo.2540861
- Parins-Fukuchi C., Stull G.W., Smith S.A. 2021. Phylogenomic conflict coincides with rapid morphological innovation. *Proc. Natl. Acad. Sci.* 118:e2023058118.
- Parvathi M.S., Nataraja K.N., Reddy Y.A., Naika M.B., Gowda M.V. 2019. Transcriptome analysis of finger millet (*Eleusine coracana* (L.) Gaertn.) reveals unique drought responsive genes. *J. Genet.* 98:1–12.
- Paun, O., Turner, B., Trucchi, E., Munzinger, J., Chase, M. W., & Samuel, R. (2016). Processes driving the adaptive radiation of a tropical tree (Diospyros, Ebenaceae) in New Caledonia, a biodiversity hotspot. *Systematic Biology*, 65(2), 212–227. <https://doi.org/10.1093/sysbio/syv076>
- 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.
- 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.

Pérez-Escobar O.A., Chomicki G., Condamine F.L., Karremans A.P., Bogarín D., Matzke N.J., Silvestro D., Antonelli A. 2017. Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytol.* 215:891–905.

Picard Toolkit. 2018. Broad Institute, GitHub Repository: <http://broadinstitute.github.io/picard/>. *Broad Institute*.

Pinzón J.P., Ramírez-Morillo I.M., Carnevali G., Barfuss M.H., Till W., Tun J., Ortiz-Díaz J.J. 2016. Phylogenetics and evolution of the *Tillandsia utriculata* complex (Bromeliaceae, Tillandsioideae) inferred from three plastid DNA markers and the ETS of the nuclear ribosomal DNA. *Bot. J. Linn. Soc.* 181:362–390.

Poplin R., Ruano-Rubio V., DePristo M.A., Fennell T.J., Carneiro M.O., Auwera G.A.V. der, Kling D.E., Gauthier L.D., Levy-Moonshine A., Roazen D., Shakir K., Thibault J., Chandran S., Whelan C., Lek M., Gabriel S., Daly M.J., Neale B., MacArthur D.G., Banks E. 2018. Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*:201178.

Qian H., Ricklefs R.E. 2000. Large-scale processes and the Asian bias in species diversity of temperate plants. *Nature*. 407:180–182.

Quinlan A.R., Hall I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 26:841–842.

R Core Team. 2020. R: A language and environment for statistical computing. *Vienna, Austria: R Foundation for Statistical Computing*.

Renaud G., Stenzel U., Maricic T., Wiebe V., Kelso J. 2015. deML: robust demultiplexing of Illumina sequences using a likelihood-based approach. *Bioinforma. Oxf. Engl.* 31:770–772.

Revell L.J. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3:217–223.

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.

Rivera N.L. 2019. *Exploring genomic relationships in Tillandsia subgenus Tillandsia*. [Master's Thesis]. University of Vienna. <https://doi.org/10.25365/thesis.57906>

Rundell R.J., Price T.D. 2009. Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends Ecol. Evol.* 24:394–399.

- Salichos L., Rokas A. 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. *Nature*. 497:327–331.
- Sankararaman S., Mallick S., Dannemann M., Prüfer K., Kelso J., Pääbo S., Patterson N., Reich D. 2014. The genomic landscape of Neanderthal ancestry in present-day humans. *Nature*. 507:354–357.
- Schluter D. 2000. *The ecology of adaptive radiation*. OUP Oxford.
- 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. 2014. Genomics and the origin of species. *Nat. Rev. Genet.* 15:176–192.
- Sistla S., Rao D.N. 2004. S-Adenosyl-L-methionine-dependent restriction enzymes. *Crit. Rev. Biochem. Mol. Biol.* 39:1–19.
- Smith S.A., Moore M.J., Brown J.W., Yang Y. 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. *BMC Evol. Biol.* 15:150.
- Solís-Lemus C., Yang M., Ané C. 2016. Inconsistency of species tree methods under gene flow. *Syst. Biol.* 65:843–851.
- Soltis P.S., Folk R.A., Soltis D.E. 2019. Darwin review: angiosperm phylogeny and evolutionary radiations. *Proc. R. Soc. B Biol. Sci.* 286:20190099.
- Stankowski S., Chase M.A., Fuiten A.M., Rodrigues M.F., Ralph P.L., Streisfeld M.A. 2019. Widespread selection and gene flow shape the genomic landscape during a radiation of monkeyflowers. *PLoS Biol.* 17:e3000391.
- Straub S.C.K., Moore M.J., Soltis P.S., Soltis D.E., Liston A., Livshultz T. 2014. Phylogenetic signal detection from an ancient rapid radiation: Effects of noise reduction, long-branch attraction, and model selection in crown clade Apocynaceae. *Mol. Phylogenet. Evol.* 80:169–185.
- Stroud J.T., Losos J.B. 2016. Ecological opportunity and adaptive radiation. *Annu. Rev. Ecol. Evol. Syst.* 47:507–532.
- Suarez-Gonzalez A., Lexer C., Cronk Q.C.B. 2018. Adaptive introgression: a plant perspective. *Biol. Lett.* 14:20170688.

- Suh A., Smeds L., Ellegren H. 2015. The dynamics of incomplete lineage sorting across the ancient adaptive radiation of neoavian birds. *PLOS Biol.* 13:e1002224.
- Suvorov A., Kim B.Y., Wang J., Armstrong E.E., Peede D., D'Agostino E.R.R., Price D.K., Waddell P.J., Lang M., Courtier-Ordogozo V., David J.R., Petrov D., Matute D.R., Schrider D.R., Comeault A.A. 2022. Widespread introgression across a phylogeny of 155 *Drosophila* genomes. *Curr. Biol.* 32:111-123.e5.
- Taylor S.A., Larson E.L. 2019. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nat. Ecol. Evol.* 3:170–177.
- Than C., Ruths D., Nakhleh L. 2008. PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics.* 9:1–16.
- Till W. 2000. *Tillandsioideae*. D.H. Benzing, ed. Bromeliaceae: profile of an adaptive radiation. New York, New York: Cambridge University Press. p. Pp. 555-571.
- Valente L.M., Savolainen V., Vargas P. 2010. Unparalleled rates of species diversification in Europe. *Proc. R. Soc. B Biol. Sci.* 277:1489–1496.
- Wang L.-G., Lam T.T.-Y., Xu S., Dai Z., Zhou L., Feng T., Guo P., Dunn C.W., Jones B.R., Bradley T., Zhu H., Guan Y., Jiang Y., Yu G. 2020. Treeio: an R package for phylogenetic tree input and output with richly annotated and associated data. *Mol. Biol. Evol.* 37:599–603.
- Wessinger C.A. 2021. From pollen dispersal to plant diversification: genetic consequences of pollination mode. *New Phytol.* 229:3125–3132.
- Whitfield J.B., Lockhart P.J. 2007. Deciphering ancient rapid radiations. *Trends Ecol. Evol.* 22:258–265.
- Winkler S. 1990. Zur Evolution der Gattung *Tillandsia* L. *Bot. Jahrb. Syst.* 112:43–77.
- Yan H., Jiang J. 2007. Rice as a model for centromere and heterochromatin research. *Chromosome Res.* 15:77–84.
- Yardeni G., Viruel J., Paris M., Hess J., Groot Crego C., de La Harpe M., Rivera N., Barfuss M.H.J., Till W., Guzmán-Jacob V., Krömer T., Lexer C., Paun O., Leroy T. 2022. Taxon-specific or universal? Using target capture to study the evolutionary history of rapid radiations. *Mol. Ecol. Resour.* 22:927–945.

Yu G. 2020. Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinforma.* 69:e96.

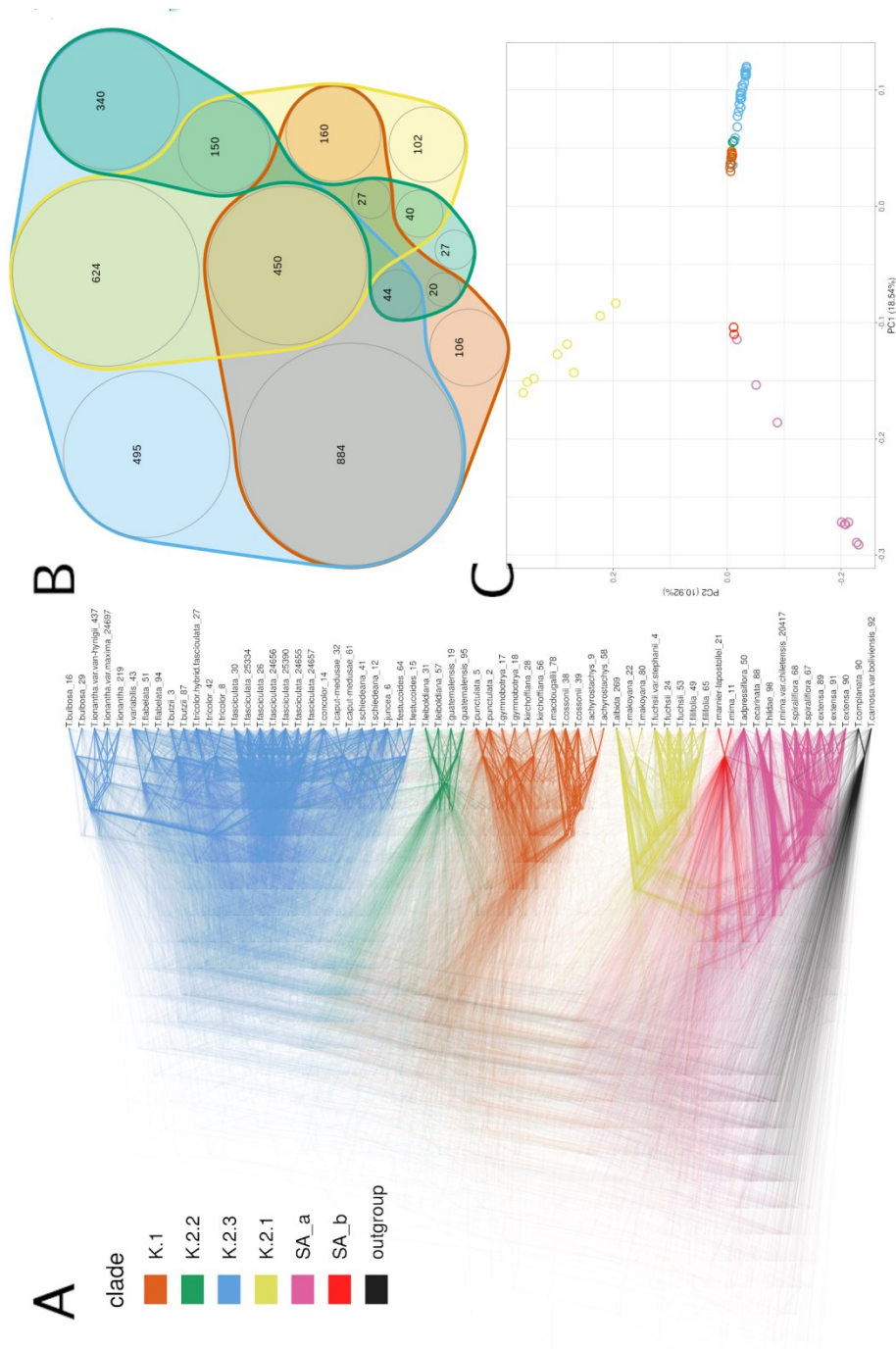
Yu Y., Nakhleh L. 2015. A maximum pseudo-likelihood approach for phylogenetic networks. *BMC Genomics.* 16:S10.

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

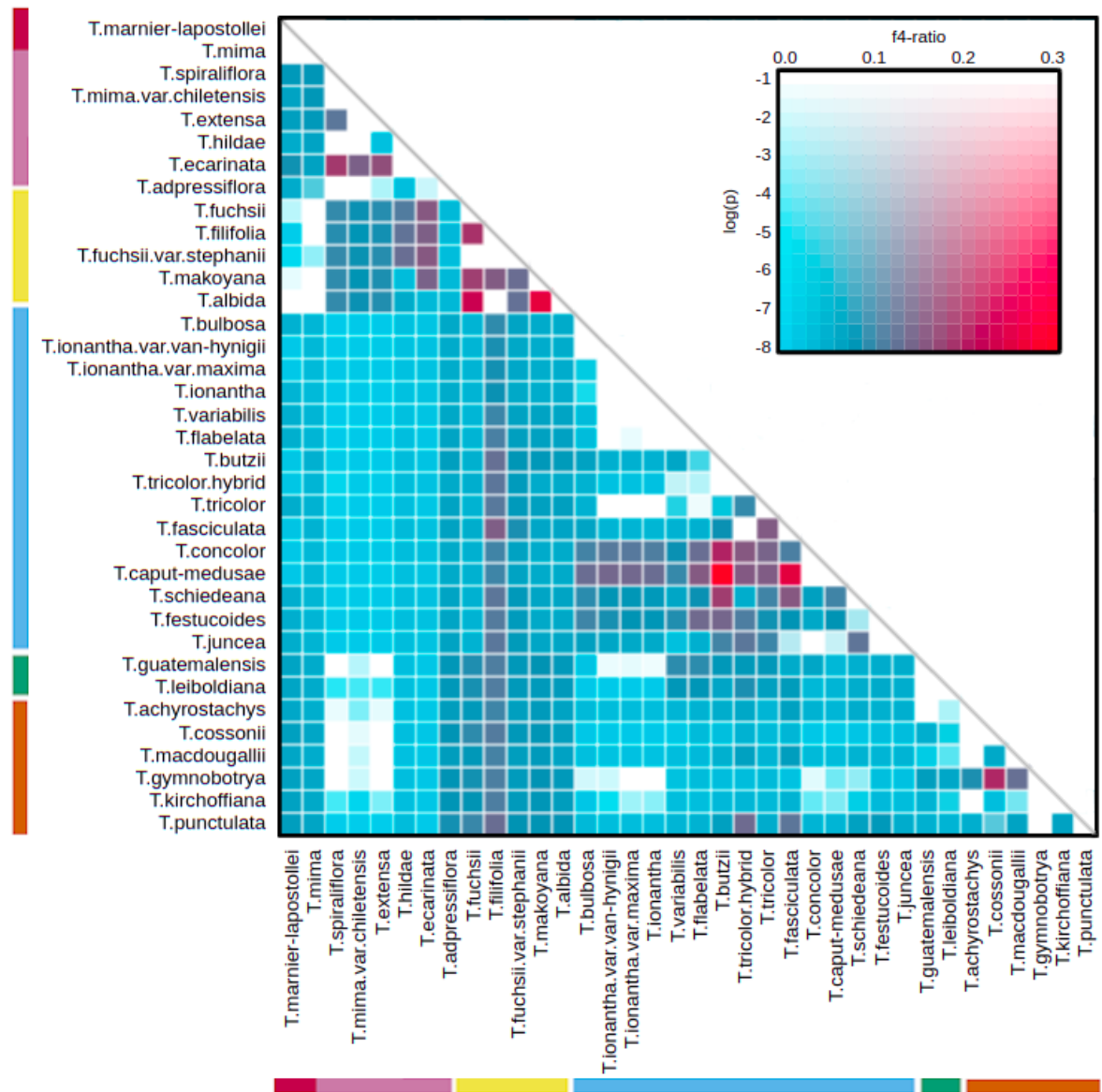
Zhang D., Rheindt F.E., She H., Cheng Y., Song G., Jia C., Qu Y., Alström P., Lei F. 2021. Most genomic loci misrepresent the phylogeny of an avian radiation because of ancient gene flow. *Syst. Biol.* 70:961–975.

Zheng X., Levine D., Shen J., Gogarten S.M., Laurie C., Weir B.S. 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics.* 28:3326–3328.

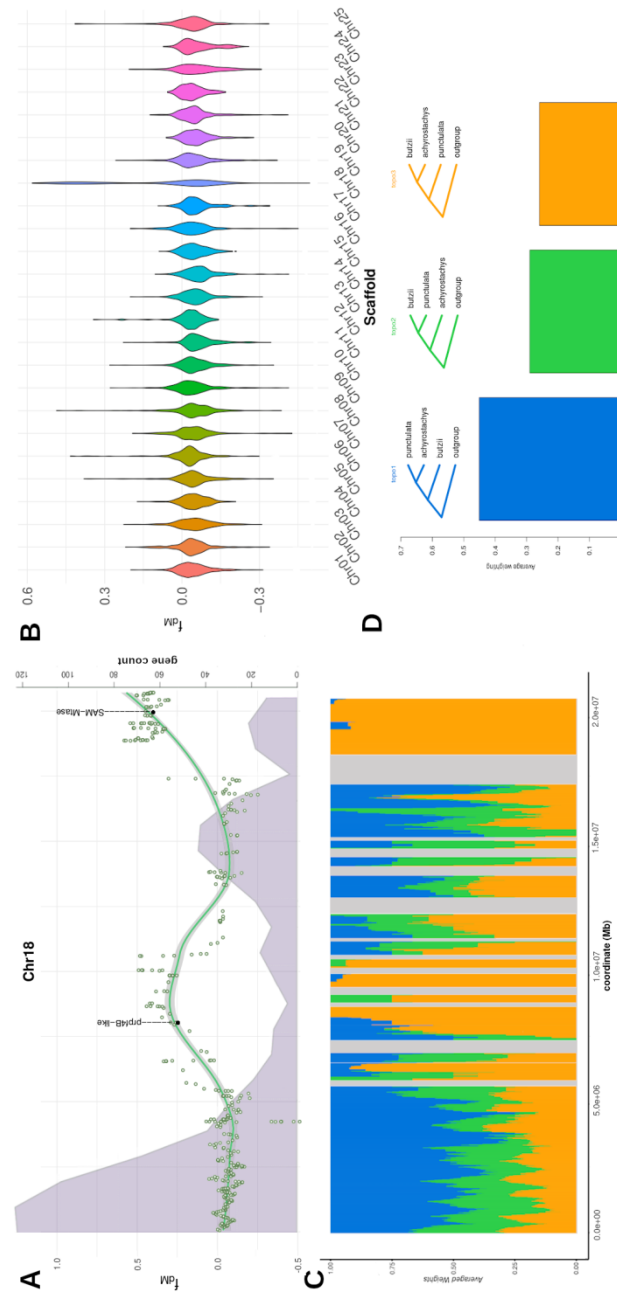
Supporting Information, including Supporting Figures, Supporting note, Supporting files and Supporting Tables



Supporting Figure S2 Widespread deviations from tree-like structure in *Tillandsia* phylogeny. Colours follow the legend and Figure 1. **A**, Claudogram of 3,785 ML phylogenetic trees inferred from non-overlapping 100kb genomic windows. Windows were forced to an ultrametric shape for visualisation purposes. **B**, Derived allele conveys deviations from tree structure affecting the K clades. The Euler diagram represents the proportion of significantly elevated Dmin scores for comparisons in trios including species from within or between clades. Numbers indicate the total number of significantly elevated Dmin for each overlap. **C**, PCA analysis for all clades on bi-allelic, distance-pruned 16,204 SNPs, including maximum 10% missing data.



Supporting Figure S3 Heatmap summarising 7,141 f_4 -ratio tests. *Tillandsia complanata* was used as the outgroup for all tests. Colour indicates the value of f_4 and log value of p -value, the latter estimated using a block jackknife procedure with a 200kb window size and corrected for family wise error rate. Colours correspond to the clades in Figure 1.

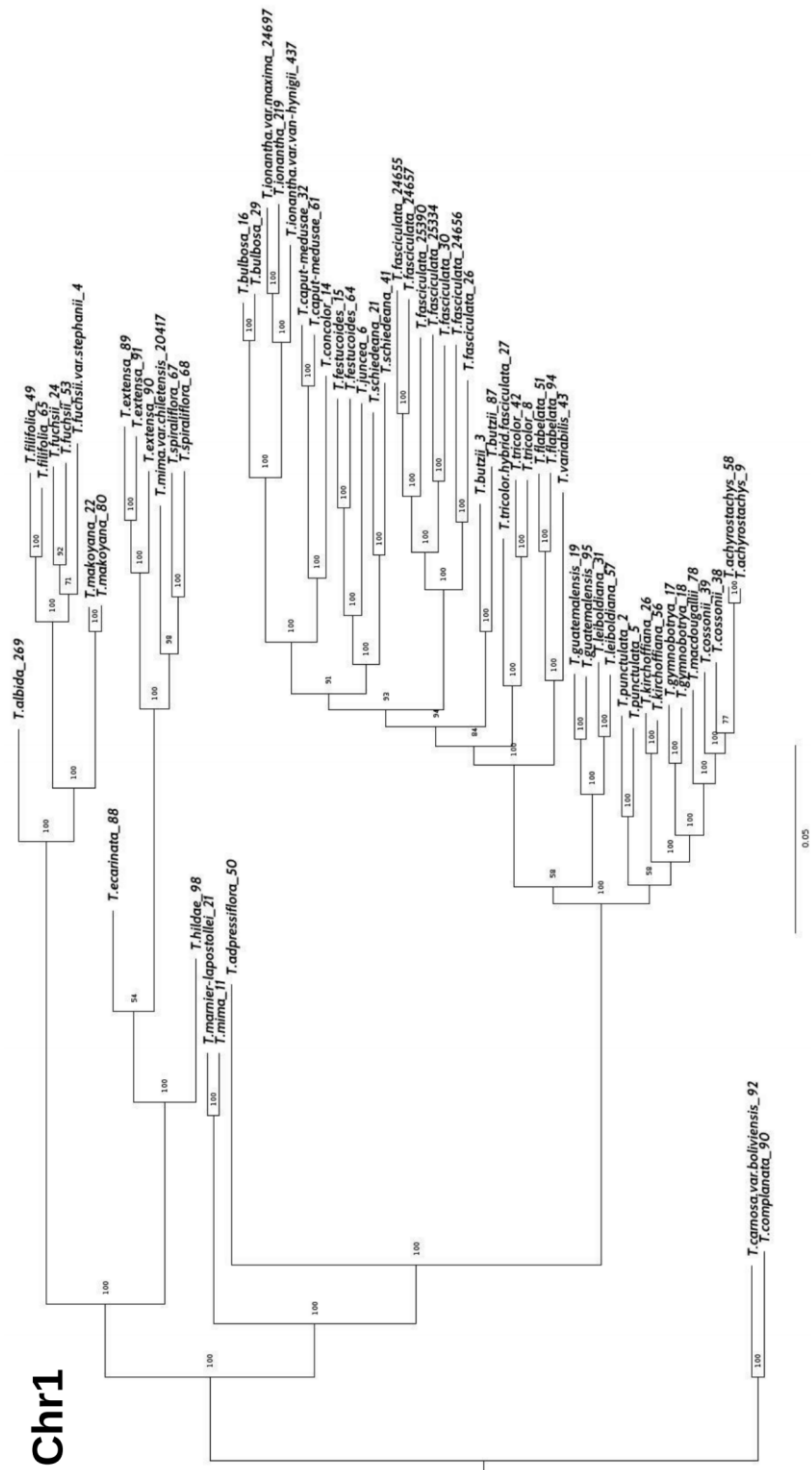


Supporting Figure S4 Signature of introgression and topology weighting on chromosome 18 between *T. punctulata*, *T. butzii* and *T. achyrostachys* (P1, P2 and P3; *T. complanata* was used as the outgroup). **A**, f_{DM} statistic (green dots and smoothed with a green line, with the scale on the right side Y axis) on genomic windows. Analysis was performed on windows of 50 SNPs with a step size of ten SNPs. Shared variation is quantified as positive values when shared between P2 and P3 and as negative values when shared between P3 and P1. Gene content is shown with lavender shape, scale on the right side. **B**, values of f_{DM} statistic for each chromosome for *T. punctulata*, *T. butzii* and *T. achyrostachys*. **C**, Topology weighting by iterative sampling of subtrees in genomic windows of 50 SNPs, using Twisst. Colours represent the frequency of each topology in **D** along chromosomal position in each chromosome: grey lines indicate regions excluded due to high levels of missing data.



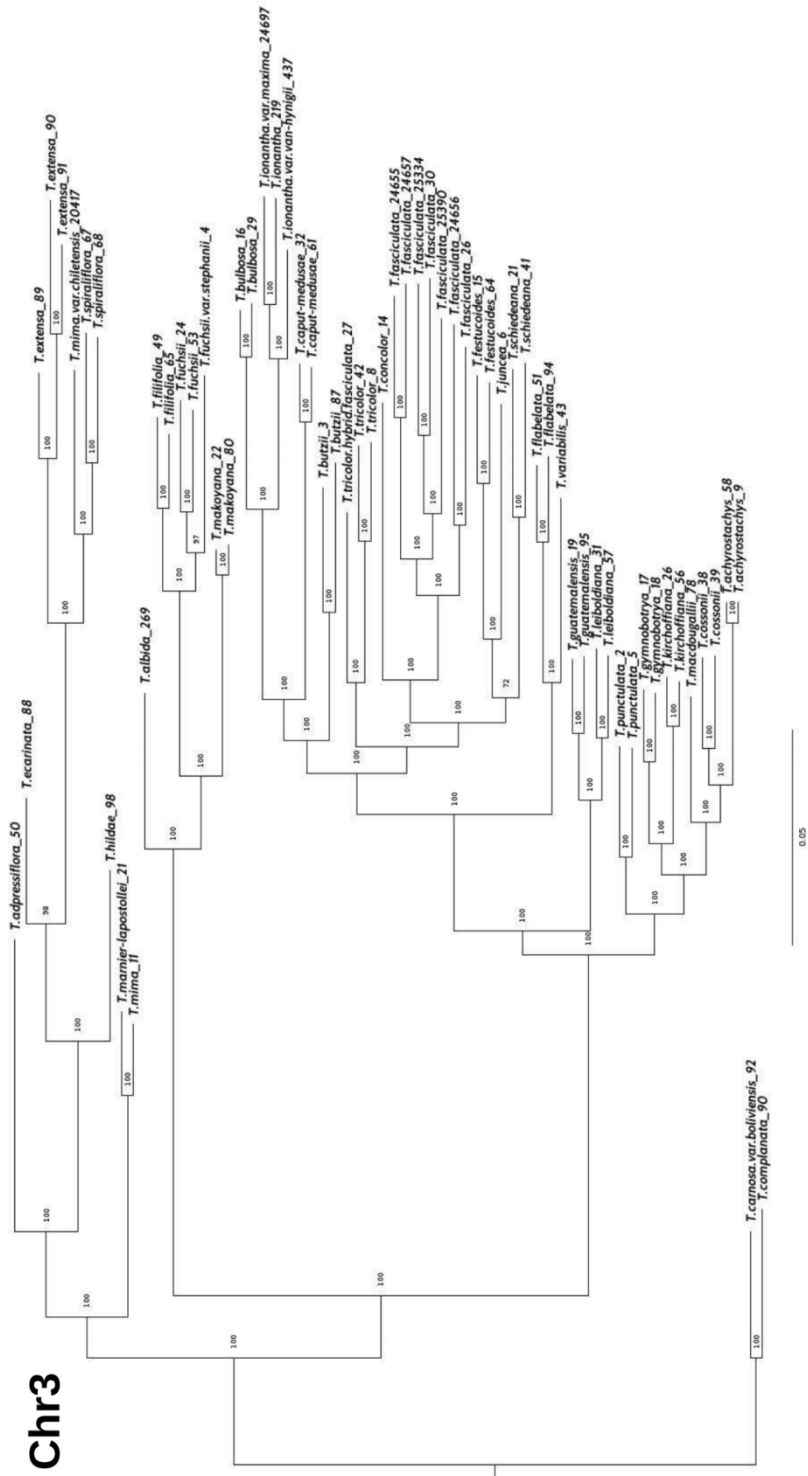
Supporting file 1 – maximum likelihood trees constructed for each of the 25 chromosomes and for subsets of the data, excluding different clades: a species tree was constructed (i) containing exclusively K clades (ii) excluding clade K.2.1. (iii) excluding the species *T. mima* and *T. marnier-lapostollei*.

Chr1

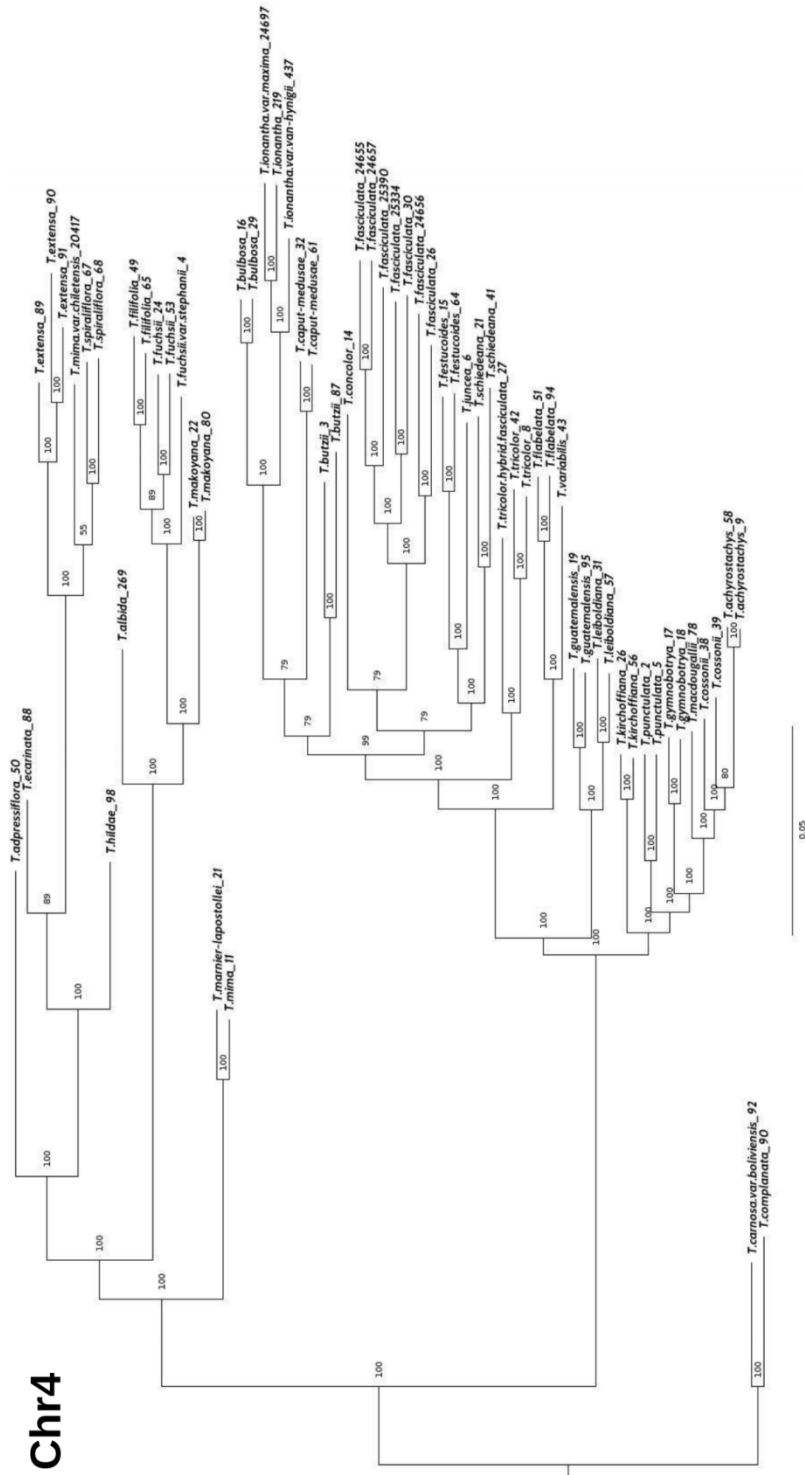


[illegible]

Chr3



Chr4



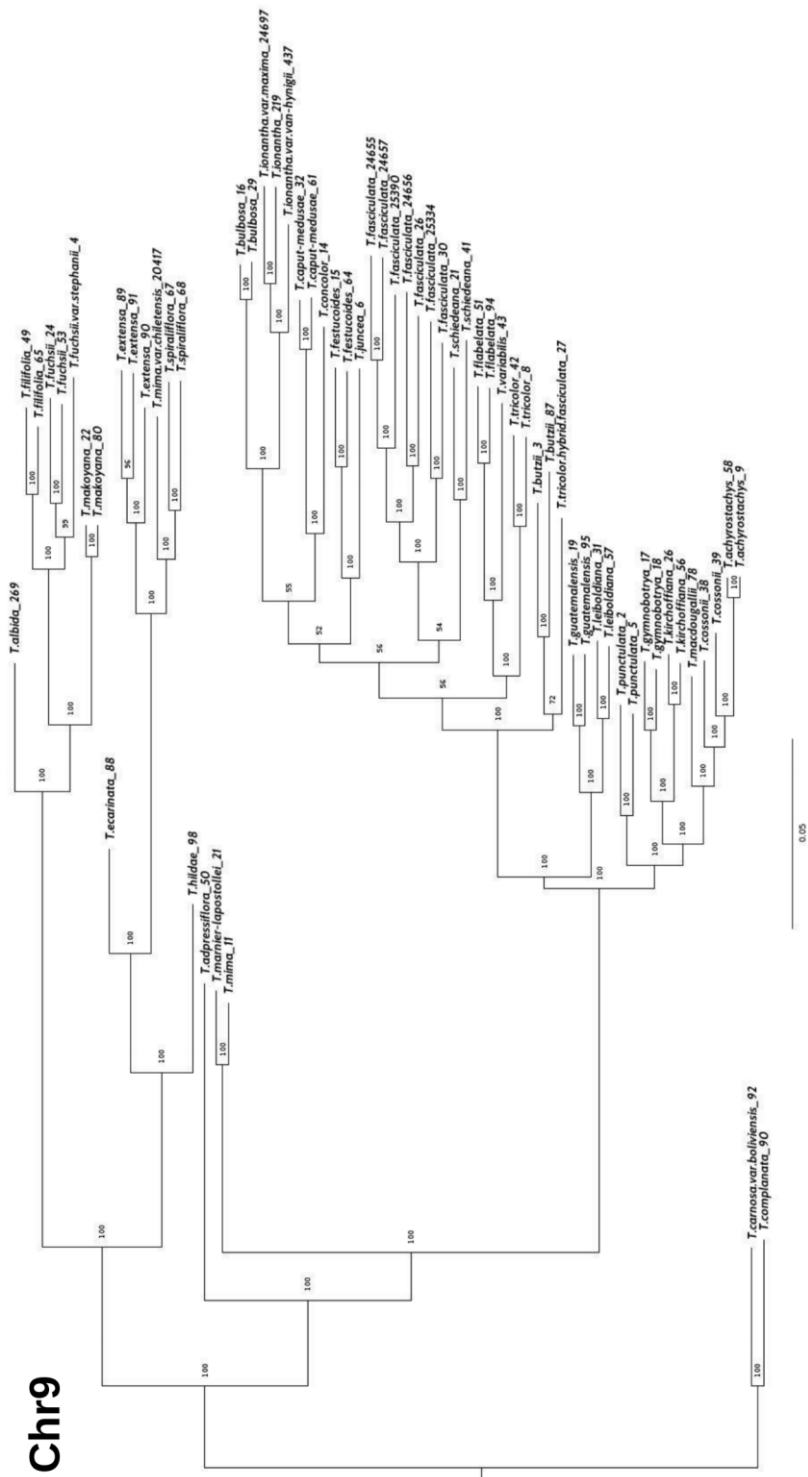
[illegible]

[illegible]

[illegible]

[illegible]

Chr9



[illegible]

200

Chr11

Phylogenetic tree showing relationships between various *Trichostema* species, with bootstrap values indicated at the nodes. The scale bar represents 0.05 substitutions per site.

Species included (from top to bottom):

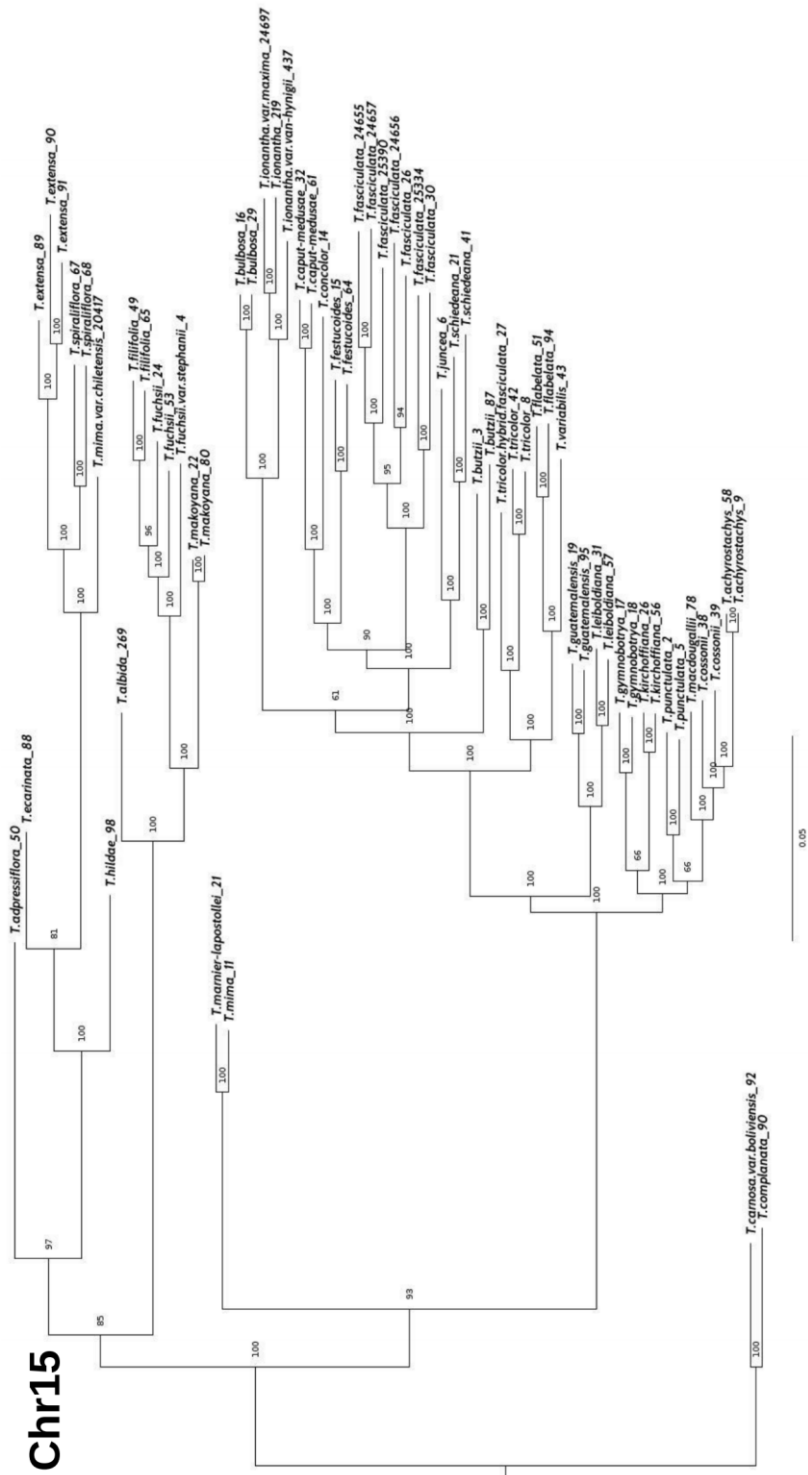
- T. albida_269*
- T. filifolia_49*
- T. filifolia_65*
- T. fuchsi_24*
- T. fuchsi_53*
- T. makoyana_22*
- T. makoyana_80*
- T. extensa_89*
- T. extensa_90*
- T. spirallifera_67*
- T. spirallifera_68*
- T. minima var. chiletensis_20417*
- T. minima var. minima_24697*
- T. lonanthe var. van-huyngii_437*
- T. bulbosa_16*
- T. bulbosa_29*
- T. caput-medusae_32*
- T. caput-medusae_61*
- T. concolor_44*
- T. festuoides_15*
- T. festuoides_64*
- T. schreckiana_21*
- T. schreckiana_41*
- T. fasciculata_24655*
- T. fasciculata_24657*
- T. fasciculata_2535*
- T. fasciculata_23394*
- T. fasciculata_30*
- T. fasciculata_24656*
- T. fasciculata_26*
- T. juncea_6*
- T. butzii_3*
- T. butzii_87*
- T. tricolor_42*
- T. tricolor_8*
- T. tricolor_51*
- T. flabellata_94*
- T. variabilis_43*
- T. guatemalensis_19*
- T. guatemalensis_35*
- T. leiboldiana_57*
- T. kirchoffiana_26*
- T. kirchoffiana_36*
- T. punctulata_5*
- T. punctulata_3*
- T. gymnobanya_17*
- T. gymnobanya_18*
- T. macedougalli_78*
- T. cossonii_38*
- T. cossonii_39*
- T. taylorii_92*
- T. taylorii_93*
- T. taylorii_94*
- T. taylorii_95*
- T. taylorii_96*
- T. taylorii_97*
- T. taylorii_98*
- T. taylorii_99*
- T. taylorii_100*
- T. taylorii_101*
- T. taylorii_102*
- T. taylorii_103*
- T. taylorii_104*
- T. taylorii_105*
- T. taylorii_106*
- T. taylorii_107*
- T. taylorii_108*
- T. taylorii_109*
- T. taylorii_110*
- T. taylorii_111*
- T. taylorii_112*
- T. taylorii_113*
- T. taylorii_114*
- T. taylorii_115*
- T. taylorii_116*
- T. taylorii_117*
- T. taylorii_118*
- T. taylorii_119*
- T. taylorii_120*
- T. taylorii_121*
- T. taylorii_122*
- T. taylorii_123*
- T. taylorii_124*
- T. taylorii_125*
- T. taylorii_126*
- T. taylorii_127*
- T. taylorii_128*
- T. taylorii_129*
- T. taylorii_130*
- T. taylorii_131*
- T. taylorii_132*
- T. taylorii_133*
- T. taylorii_134*
- T. taylorii_135*
- T. taylorii_136*
- T. taylorii_137*
- T. taylorii_138*
- T. taylorii_139*
- T. taylorii_140*
- T. taylorii_141*
- T. taylorii_142*
- T. taylorii_143*
- T. taylorii_144*
- T. taylorii_145*
- T. taylorii_146*
- T. taylorii_147*
- T. taylorii_148*
- T. taylorii_149*
- T. taylorii_150*
- T. taylorii_151*
- T. taylorii_152*
- T. taylorii_153*
- T. taylorii_154*
- T. taylorii_155*
- T. taylorii_156*
- T. taylorii_157*
- T. taylorii_158*
- T. taylorii_159*
- T. taylorii_160*
- T. taylorii_161*
- T. taylorii_162*
- T. taylorii_163*
- T. taylorii_164*
- T. taylorii_165*
- T. taylorii_166*
- T. taylorii_167*
- T. taylorii_168*
- T. taylorii_169*
- T. taylorii_170*
- T. taylorii_171*
- T. taylorii_172*
- T. taylorii_173*
- T. taylorii_174*
- T. taylorii_175*
- T. taylorii_176*
- T. taylorii_177*
- T. taylorii_178*
- T. taylorii_179*
- T. taylorii_180*
- T. taylorii_181*
- T. taylorii_182*
- T. taylorii_183*
- T. taylorii_184*
- T. taylorii_185*
- T. taylorii_186*
- T. taylorii_187*
- T. taylorii_188*
- T. taylorii_189*
- T. taylorii_190*
- T. taylorii_191*
- T. taylorii_192*
- T. taylorii_193*
- T. taylorii_194*
- T. taylorii_195*
- T. taylorii_196*
- T. taylorii_197*
- T. taylorii_198*
- T. taylorii_199*
- T. taylorii_200*
- T. taylorii_201*
- T. taylorii_202*
- T. taylorii_203*
- T. taylorii_204*
- T. taylorii_205*
- T. taylorii_206*
- T. taylorii_207*
- T. taylorii_208*
- T. taylorii_209*
- T. taylorii_210*
- T. taylorii_211*
- T. taylorii_212*
- T. taylorii_213*
- T. taylorii_214*
- T. taylorii_215*
- T. taylorii_216*
- T. taylorii_217*
- T. taylorii_218*
- T. taylorii_219*
- T. taylorii_220*
- T. taylorii_221*
- T. taylorii_222*
- T. taylorii_223*
- T. taylorii_224*
- T. taylorii_225*
- T. taylorii_226*
- T. taylorii_227*
- T. taylorii_228*
- T. taylorii_229*
- T. taylorii_230*
- T. taylorii_231*
- T. taylorii_232*
- T. taylorii_233*
- T. taylorii_234*
- T. taylorii_235*
- T. taylorii_236*
- T. taylorii_237*
- T. taylorii_238*
- T. taylorii_239*
- T. taylorii_240*
- T. taylorii_241*
- T. taylorii_242*
- T. taylorii_243*
- T. taylorii_244*
- T. taylorii_245*
- T. taylorii_246*
- T. taylorii_247*
- T. taylorii_248*
- T. taylorii_249*
- T. taylorii_250*
- T. taylorii_251*
- T. taylorii_252*
- T. taylorii_253*
- T. taylorii_254*
- T. taylorii_255*
- T. taylorii_256*
- T. taylorii_257*
- T. taylorii_258*
- T. taylorii_259*
- T. taylorii_260*
- T. taylorii_261*
- T. taylorii_262*
- T. taylorii_263*
-

[illegible]

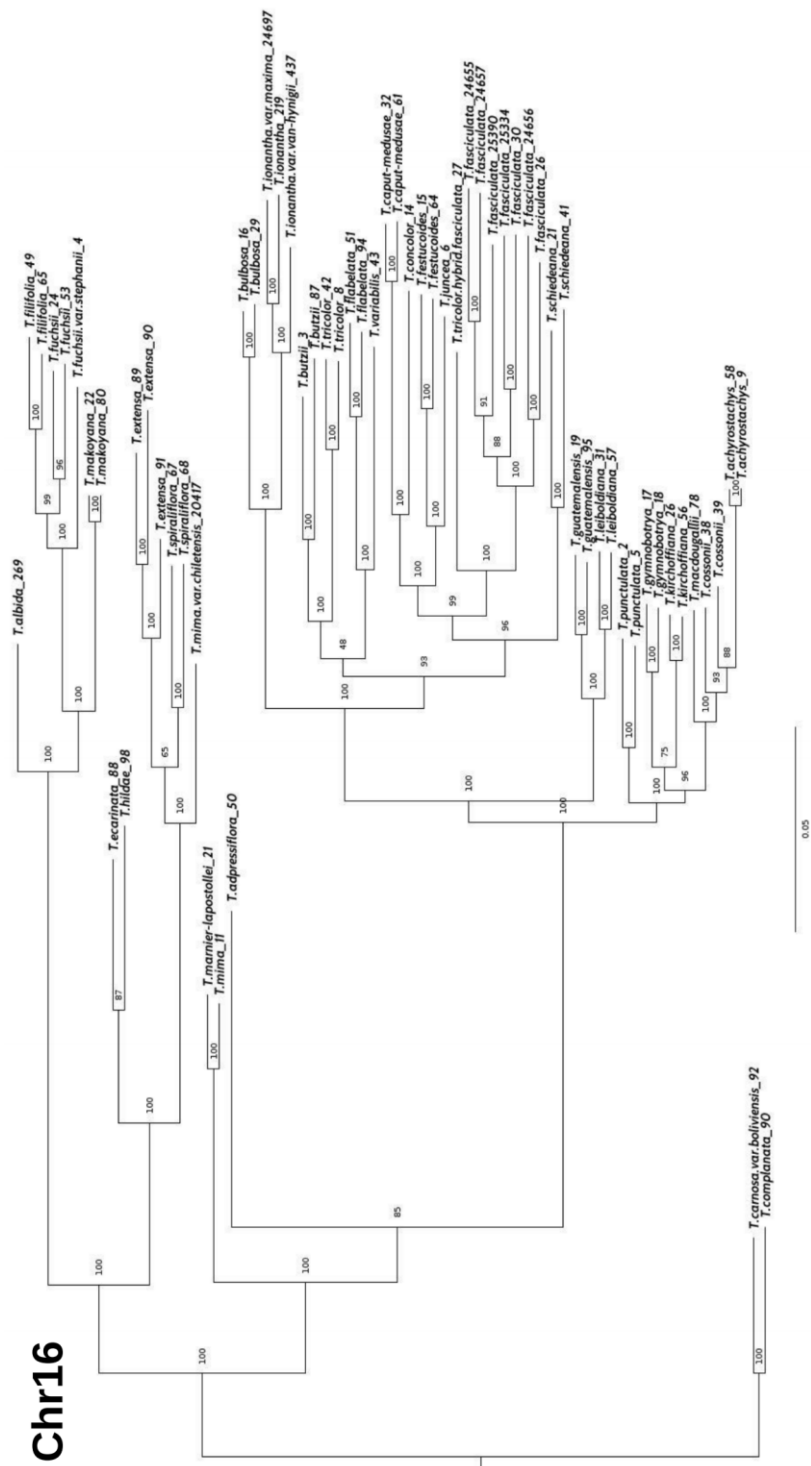
[illegible]

[illegible]

Chr15



Chr16



107



[illegible]

*T. albida*_269



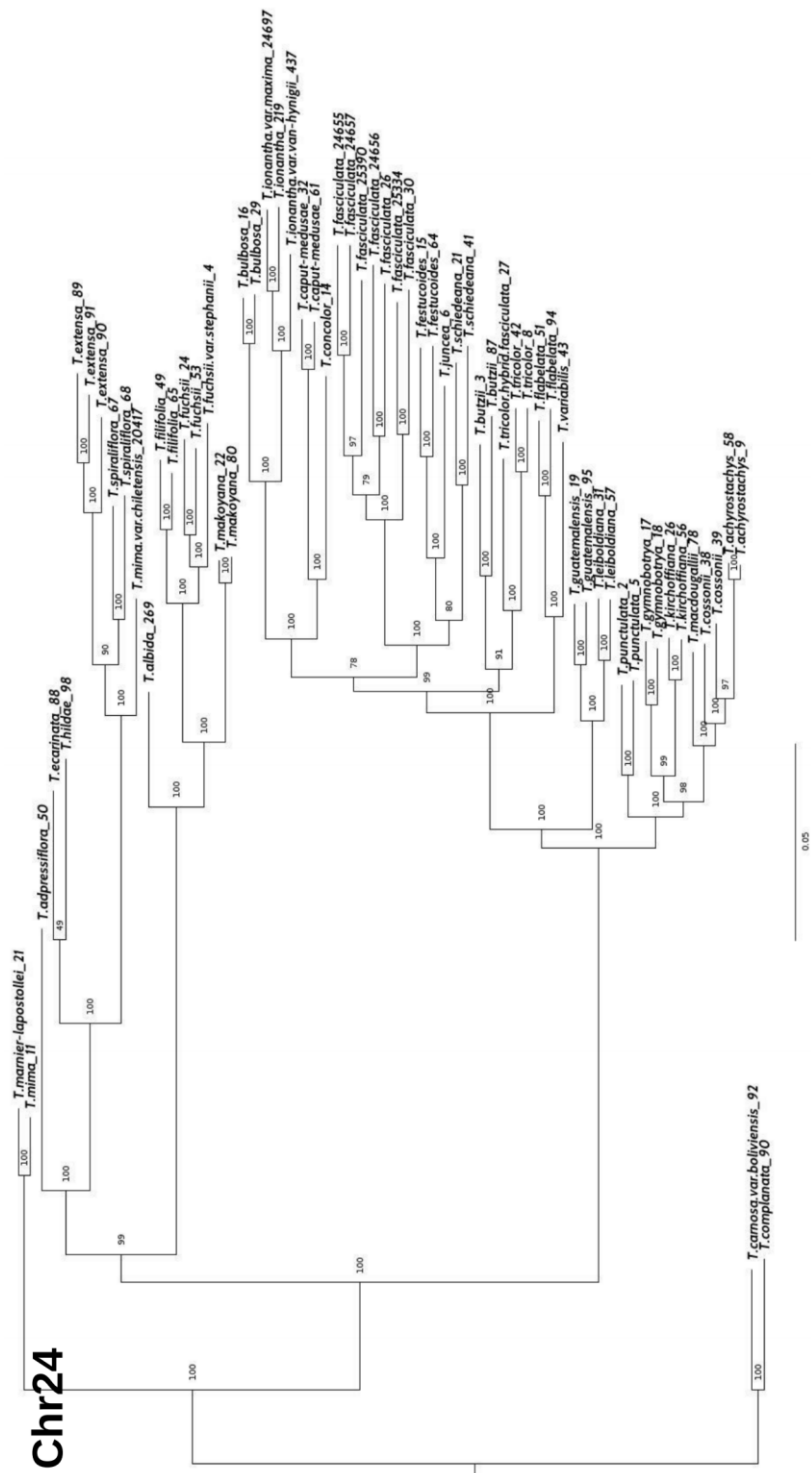
[illegible]

[illegible]

[illegible]

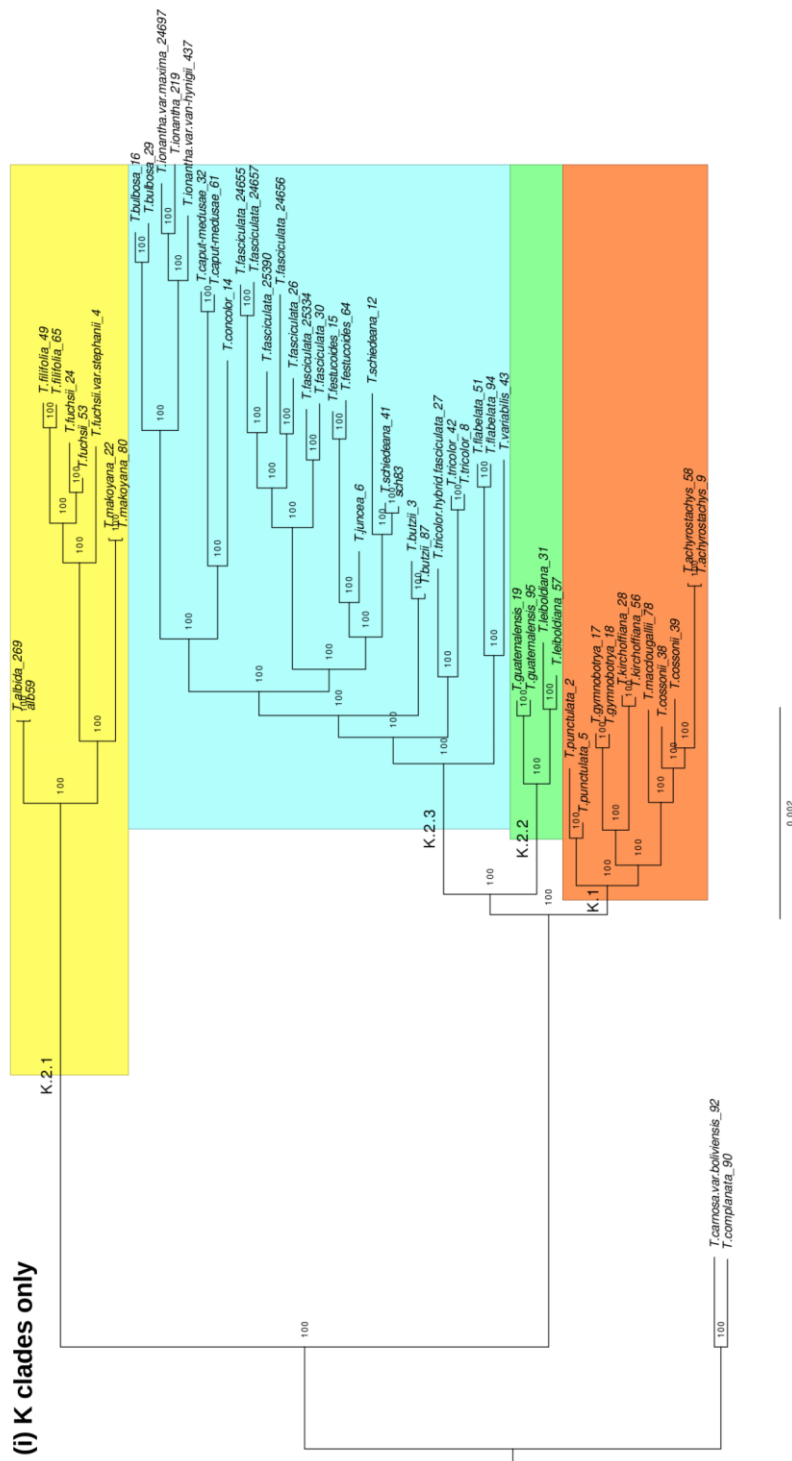
[illegible]

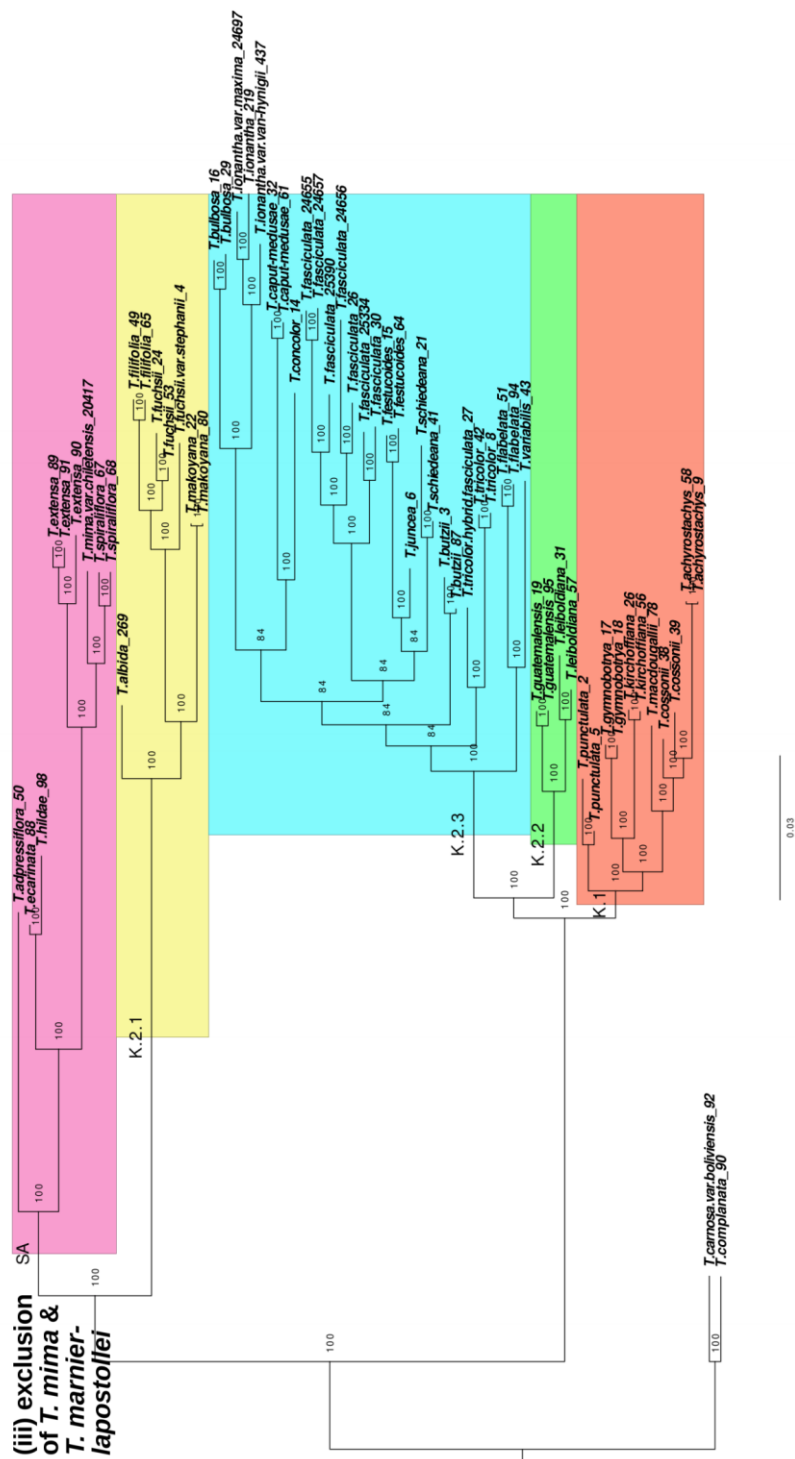
Chr24



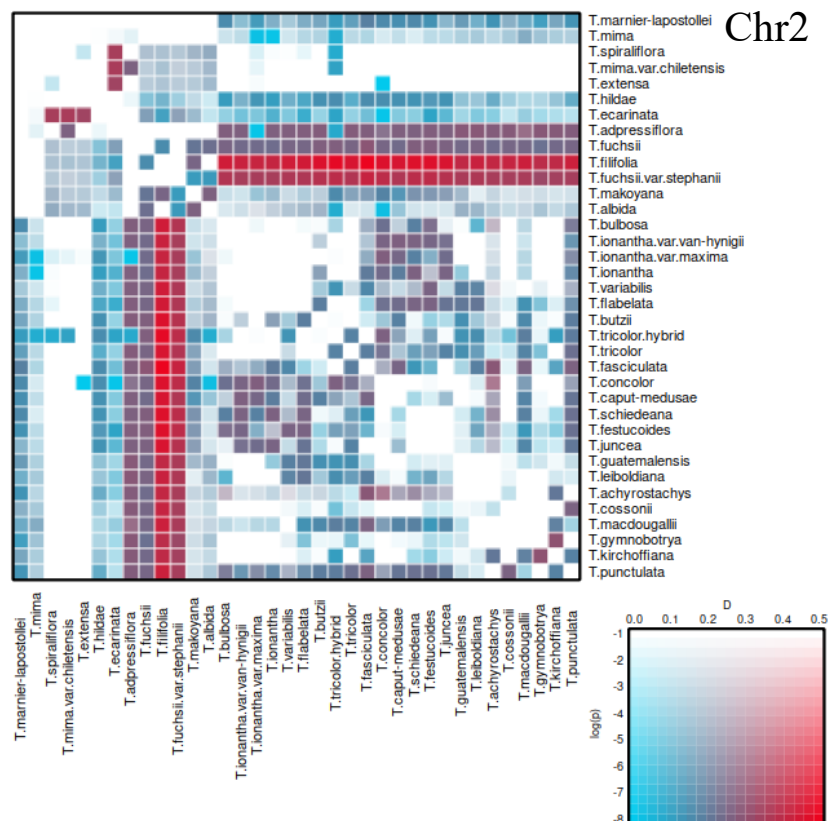
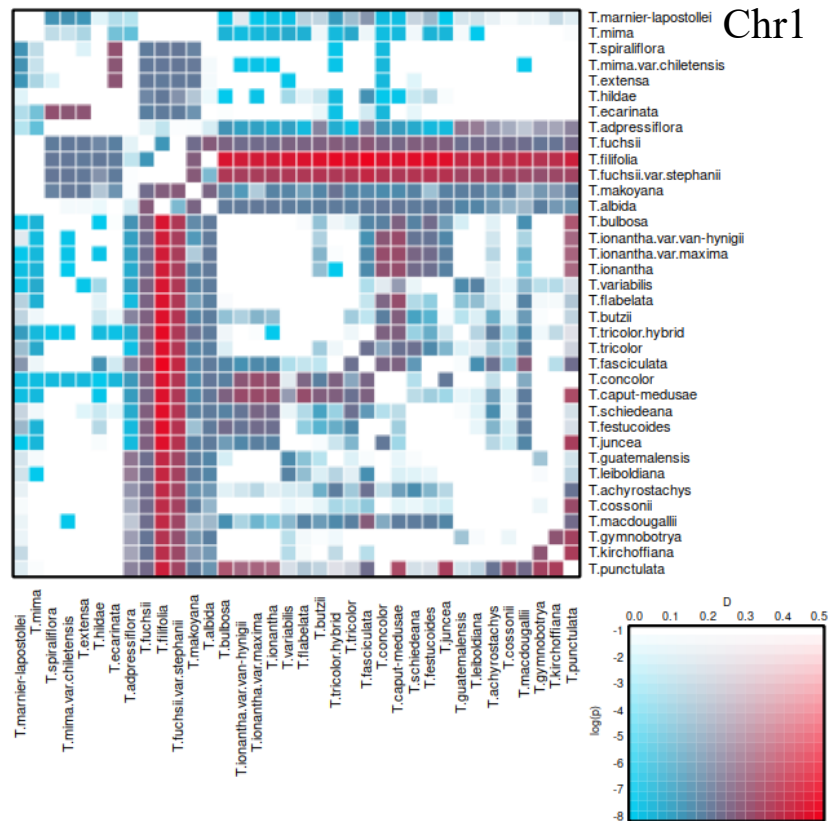
[illegible]

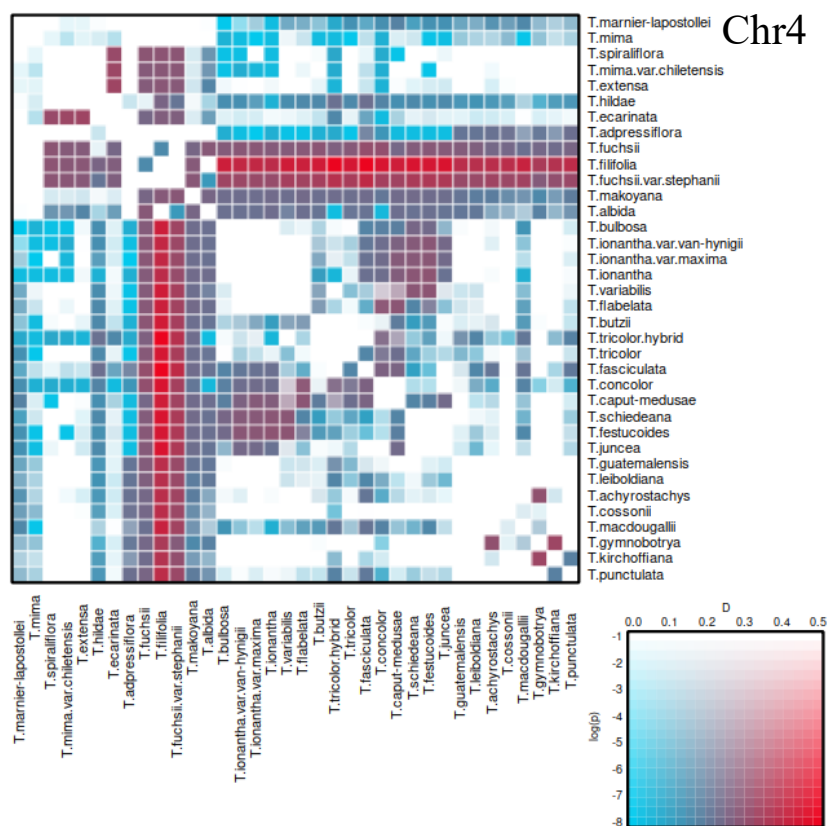
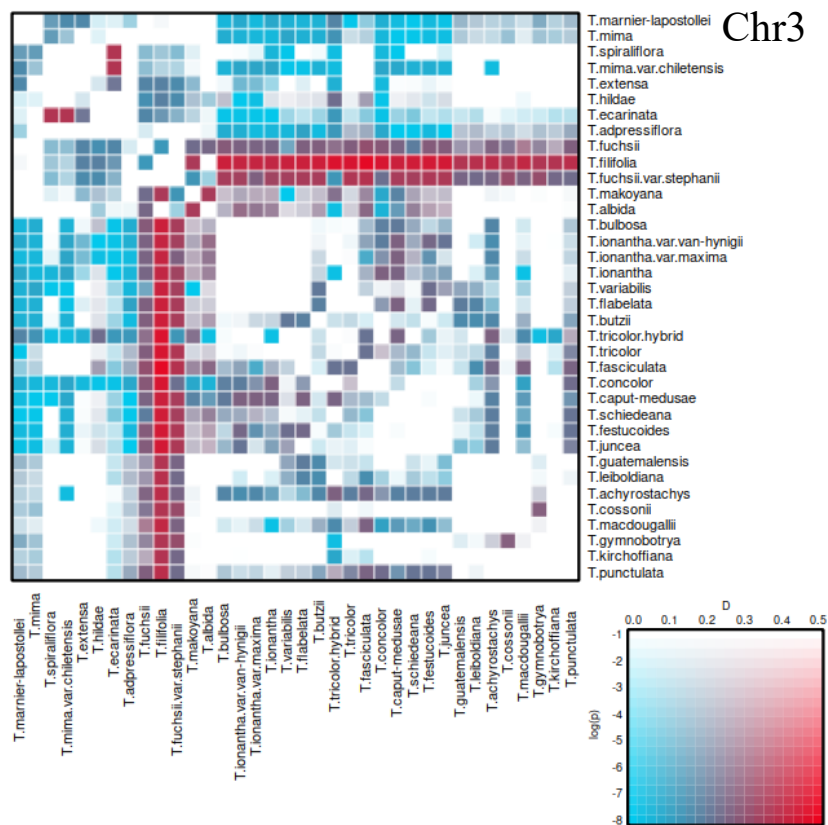
(i) K clades only

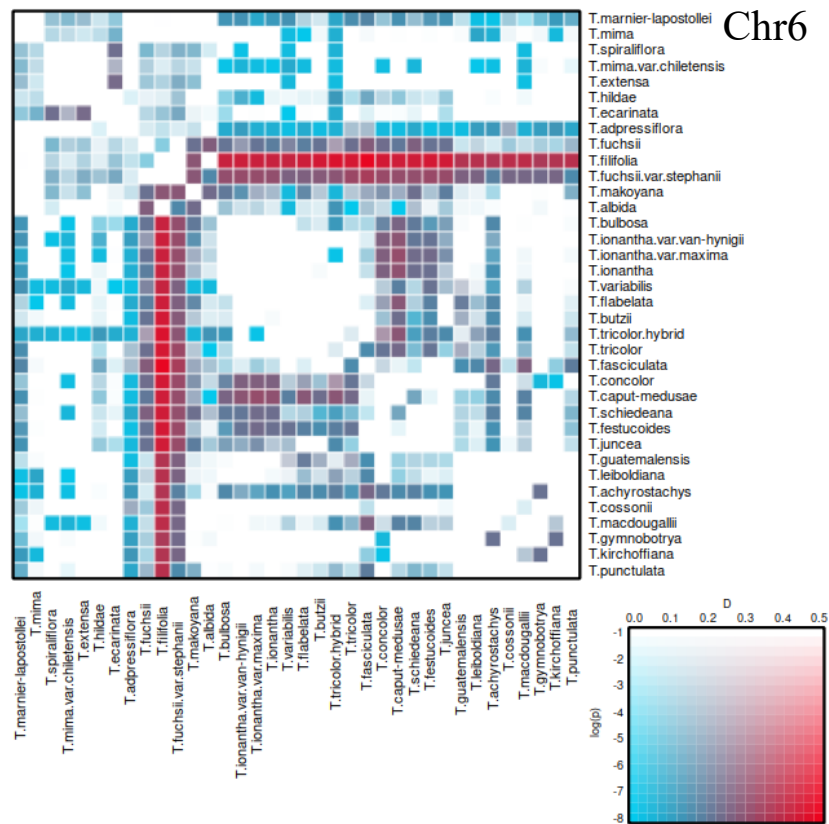
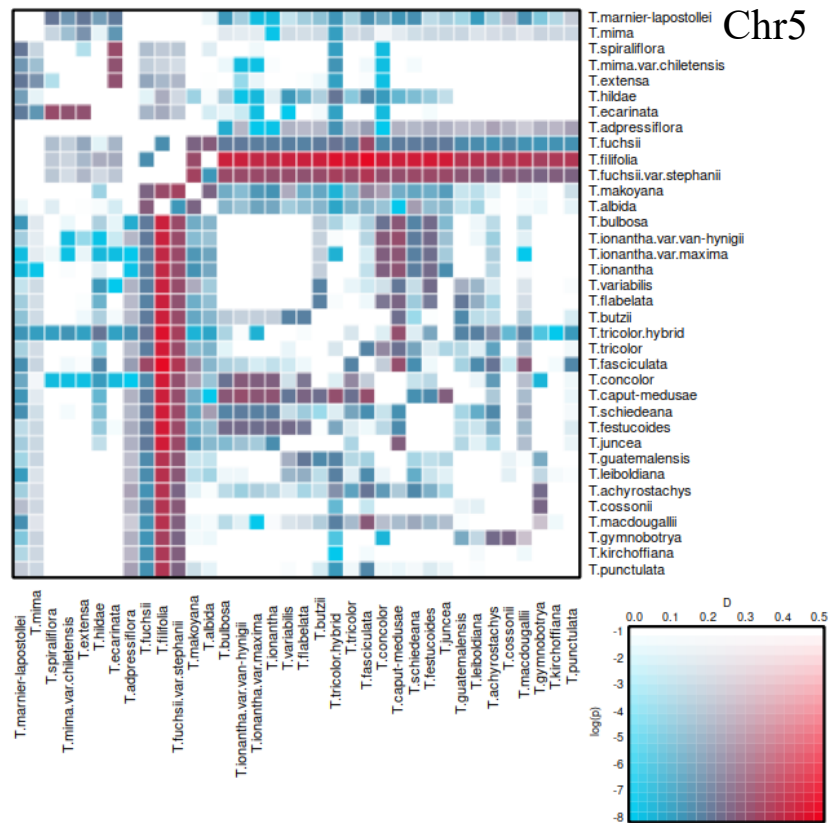


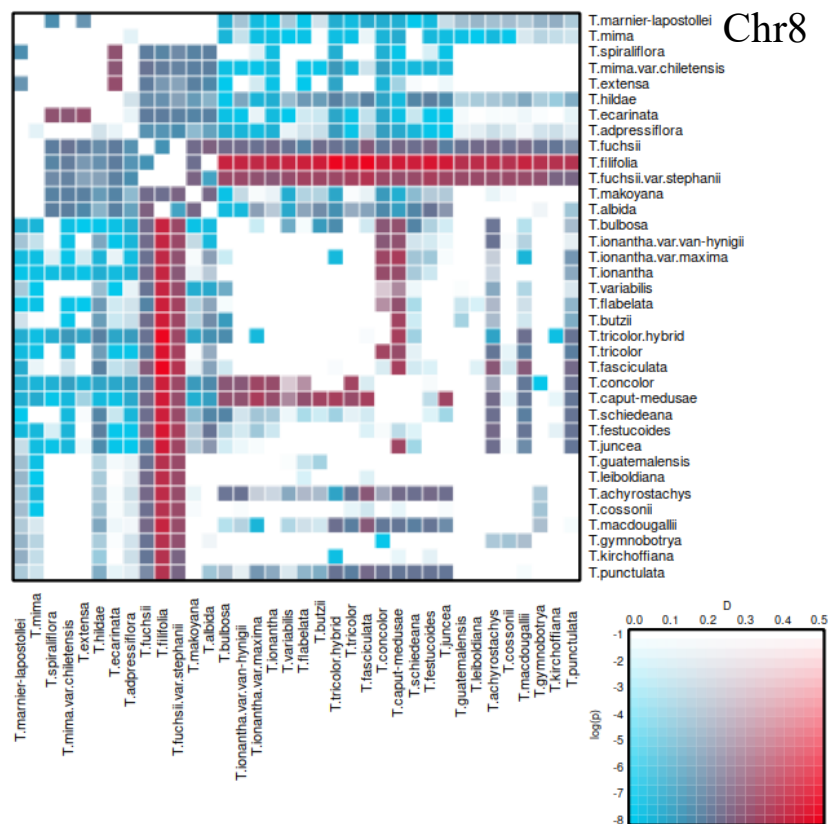
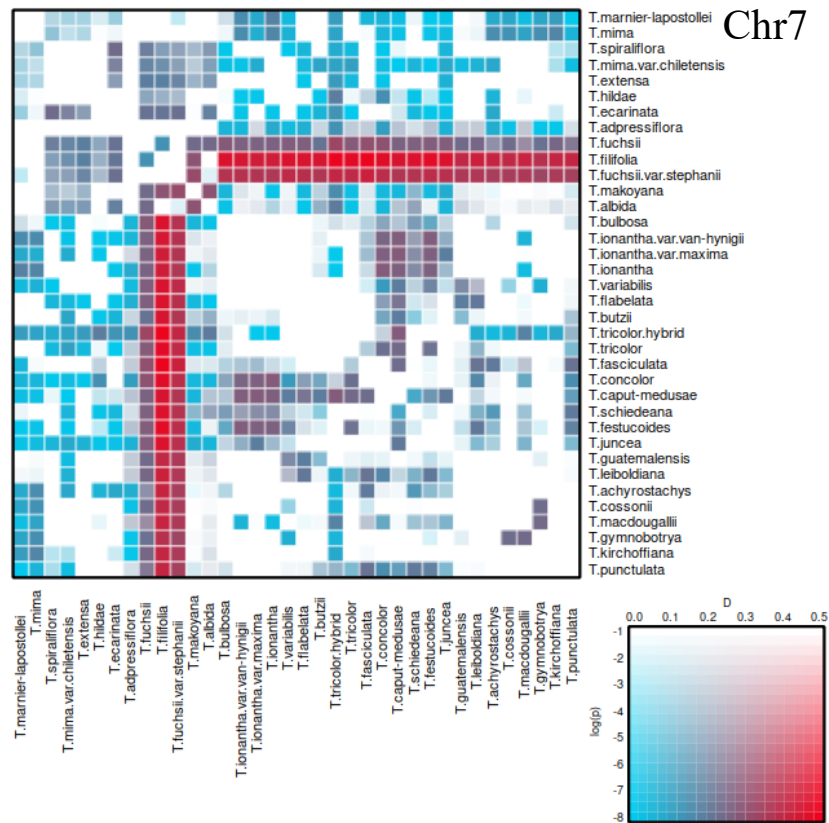


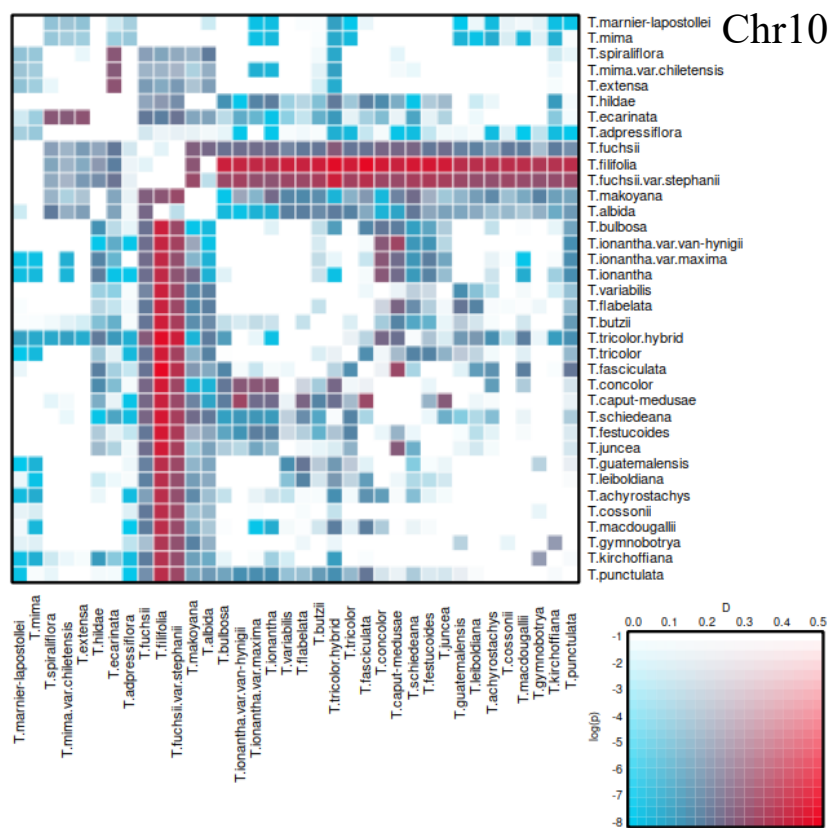
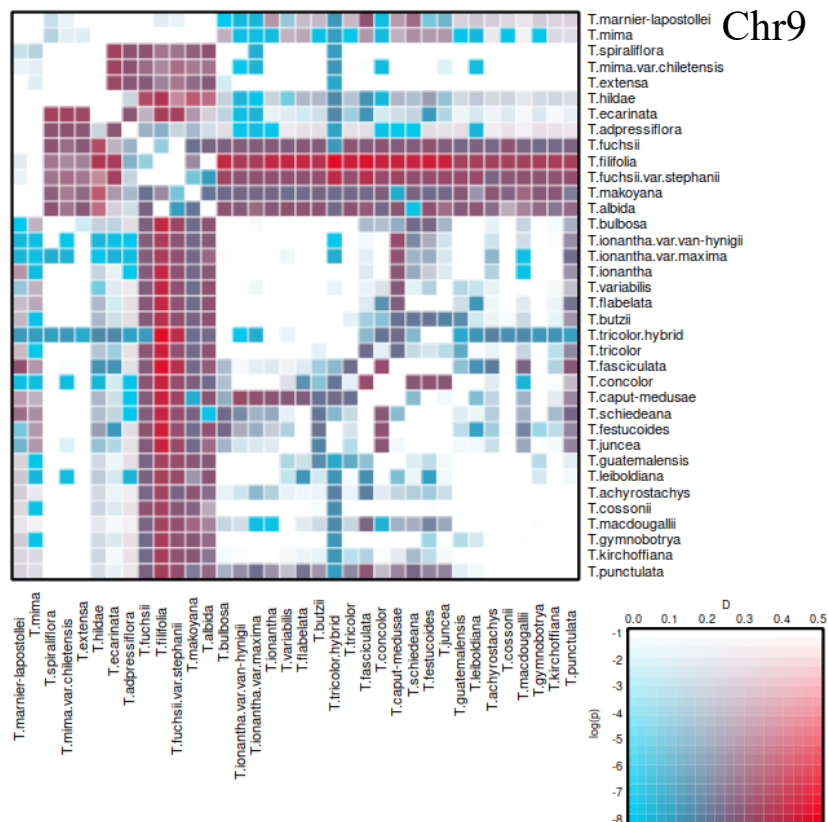
Supporting file 2 Heatmaps summarizing 7,141 four-taxon D-statistic tests for each of the 25 reference chromosomes, indicate on each figure. *Tillandsia complanata* was used as the outgroup in all tests. The four taxa in each test have been rearranged to always obtain positive D values, and P2 and P3 are shown on the axes. Colour indicates the value of D and log value of p-value, as appears in legend (bottom right).

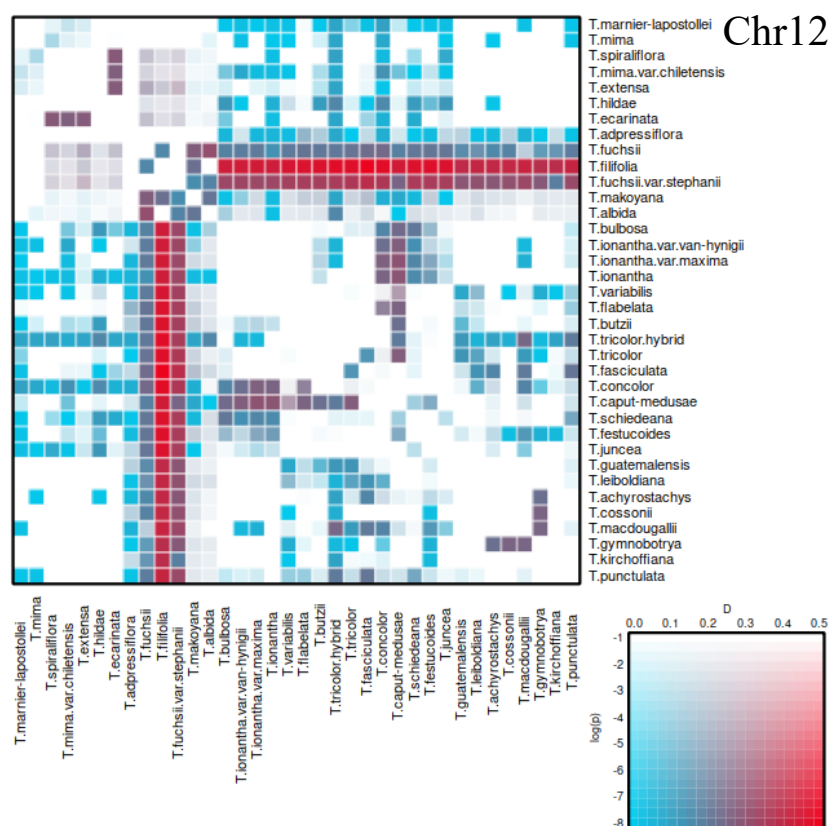
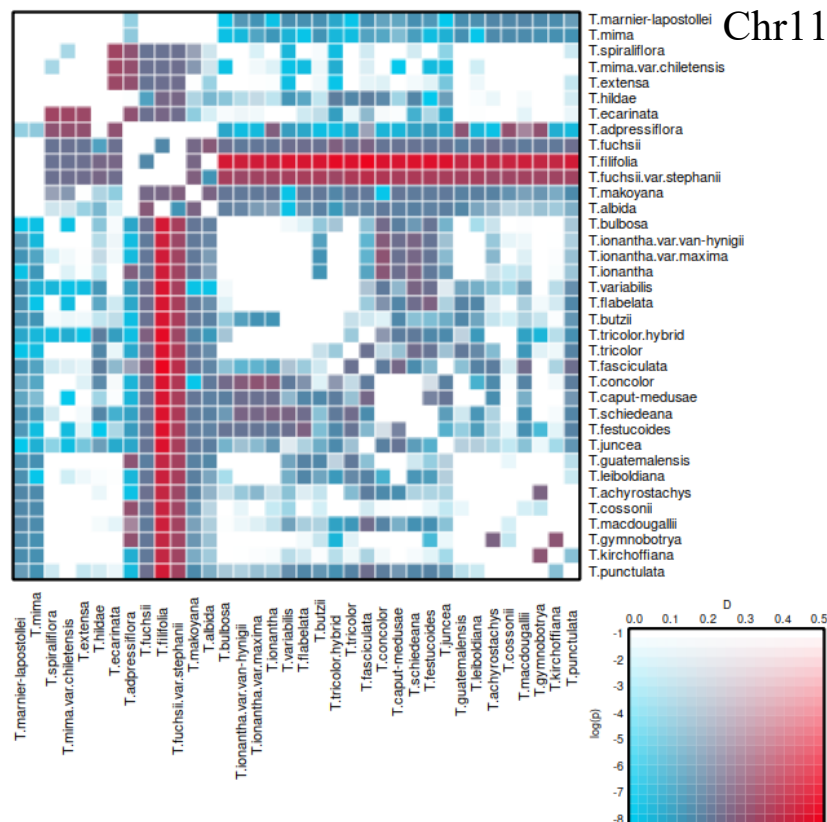


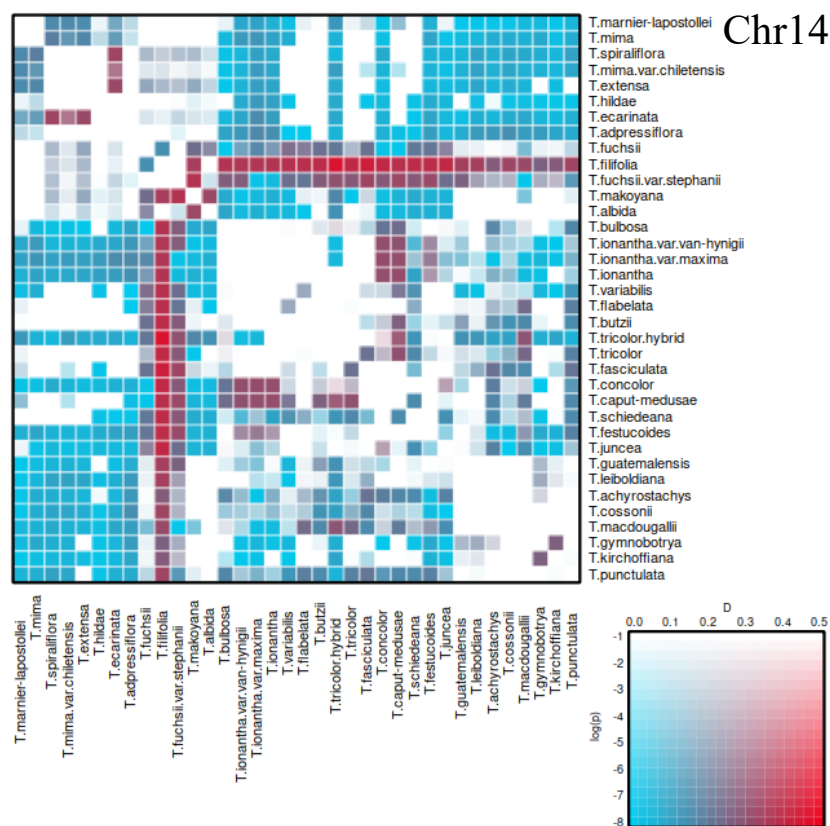
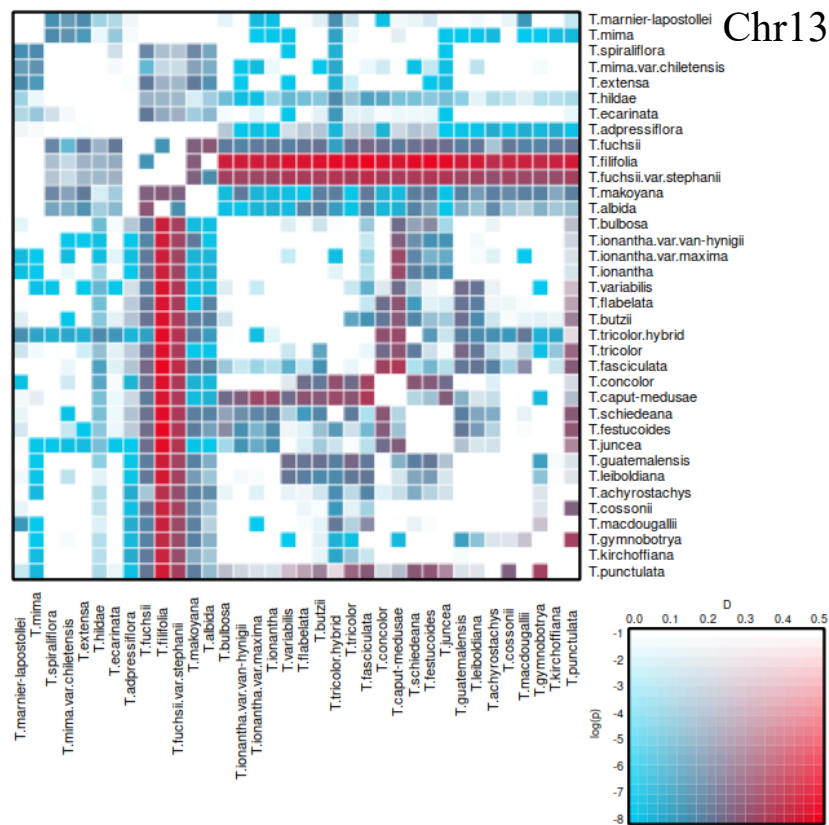


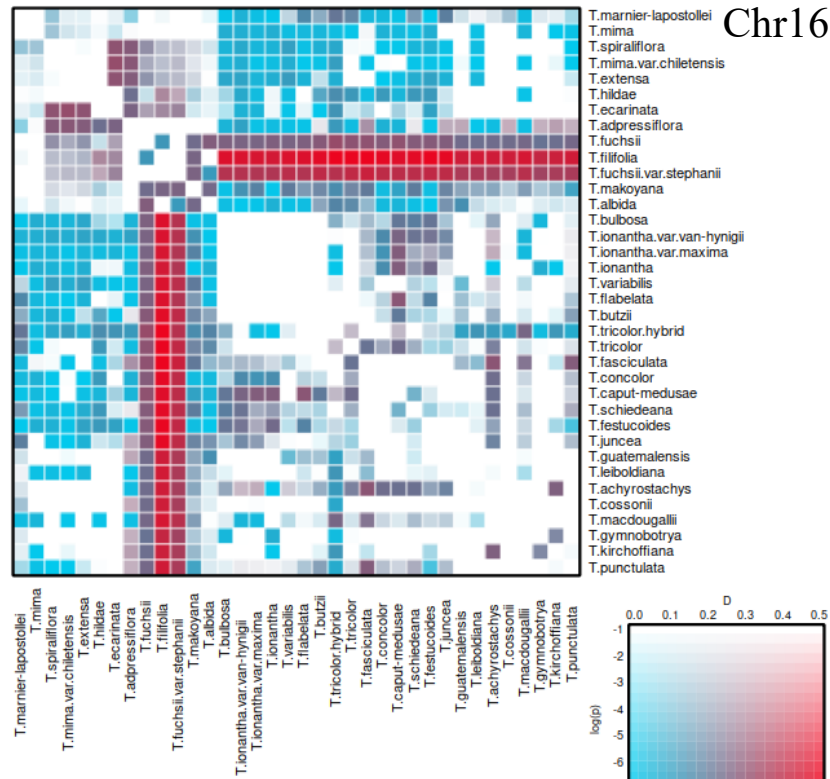
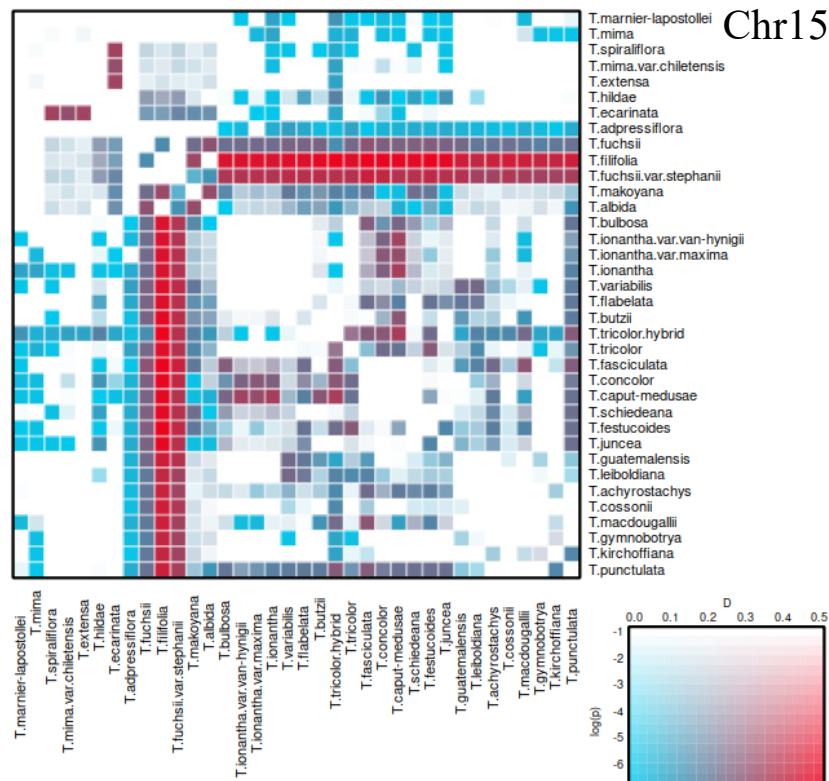


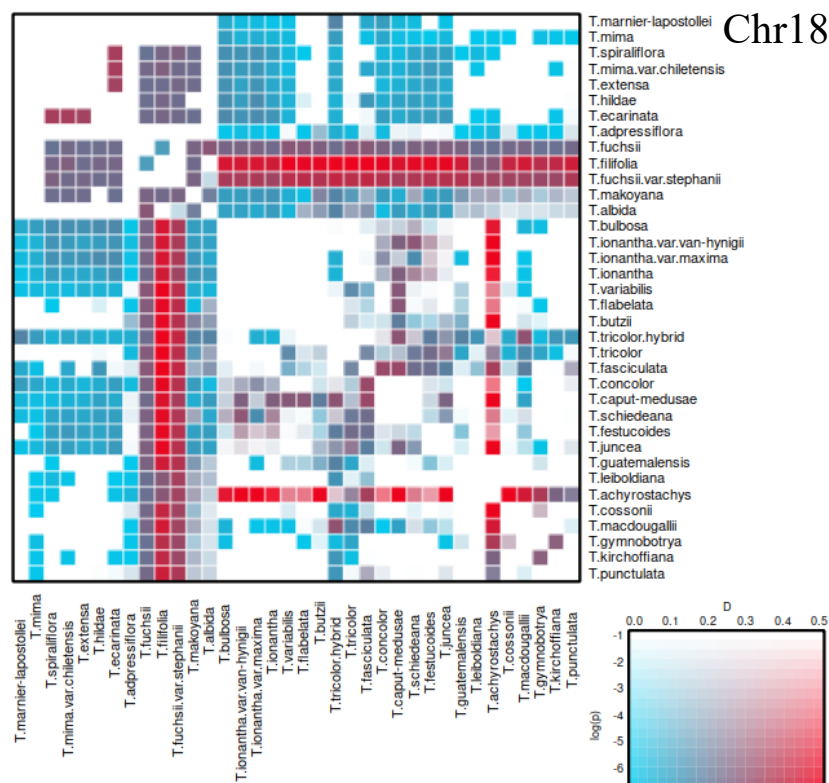
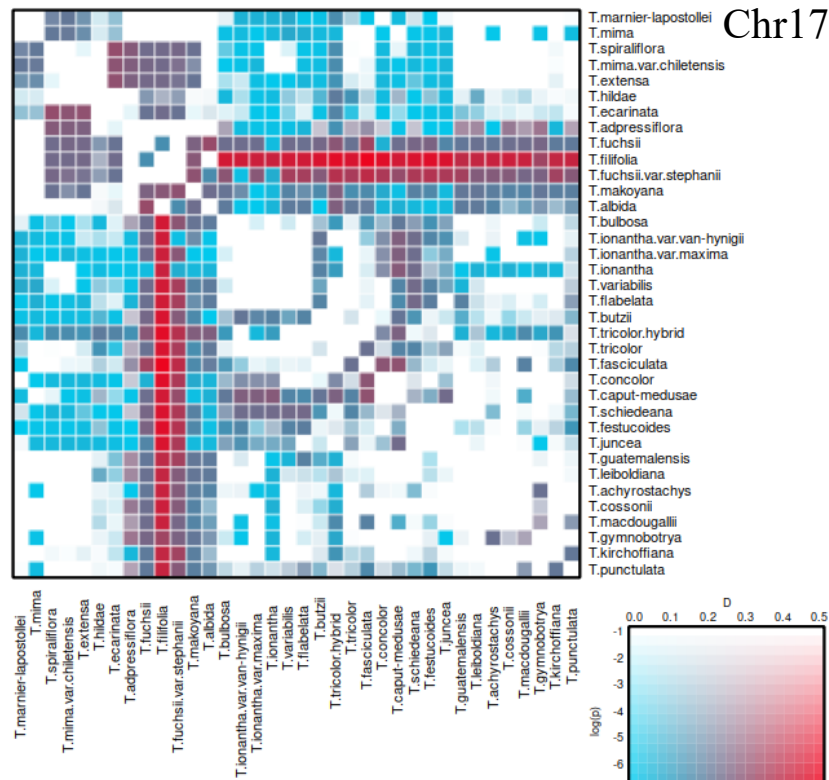


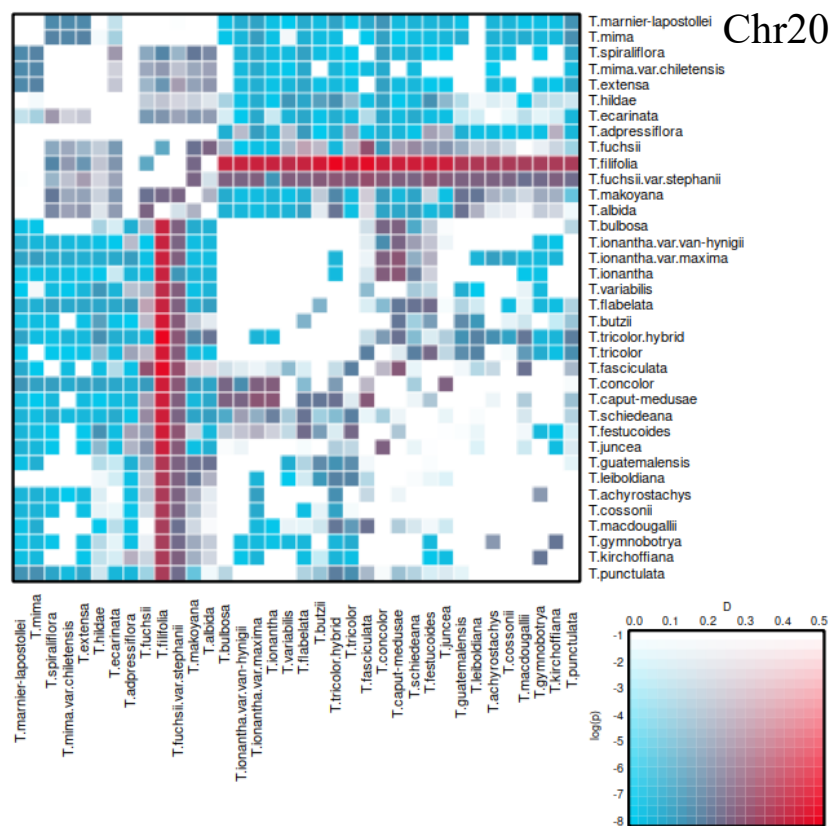
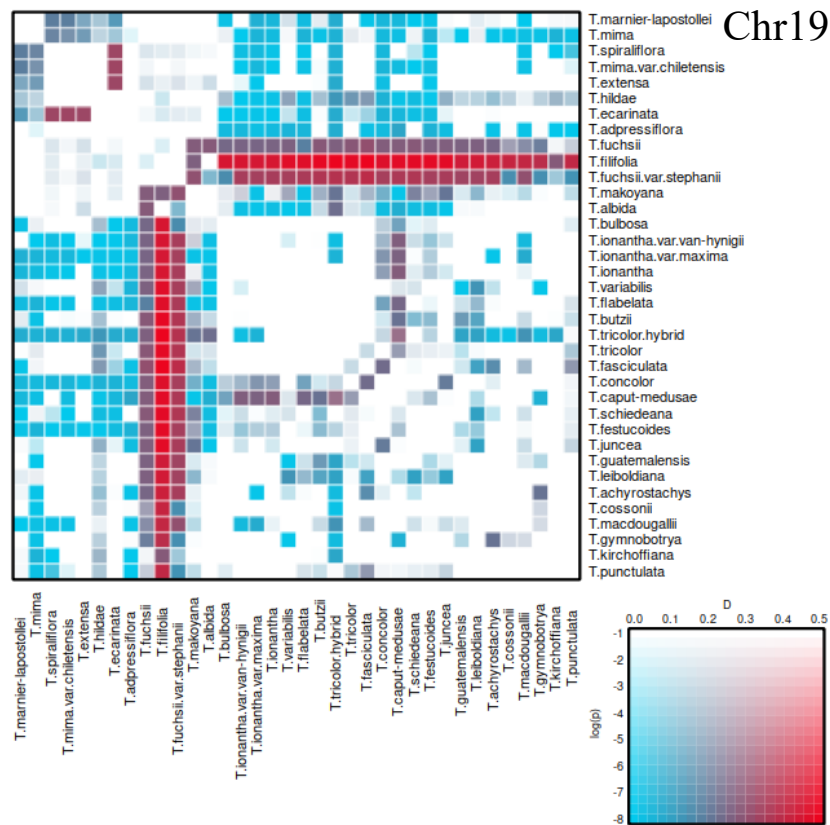


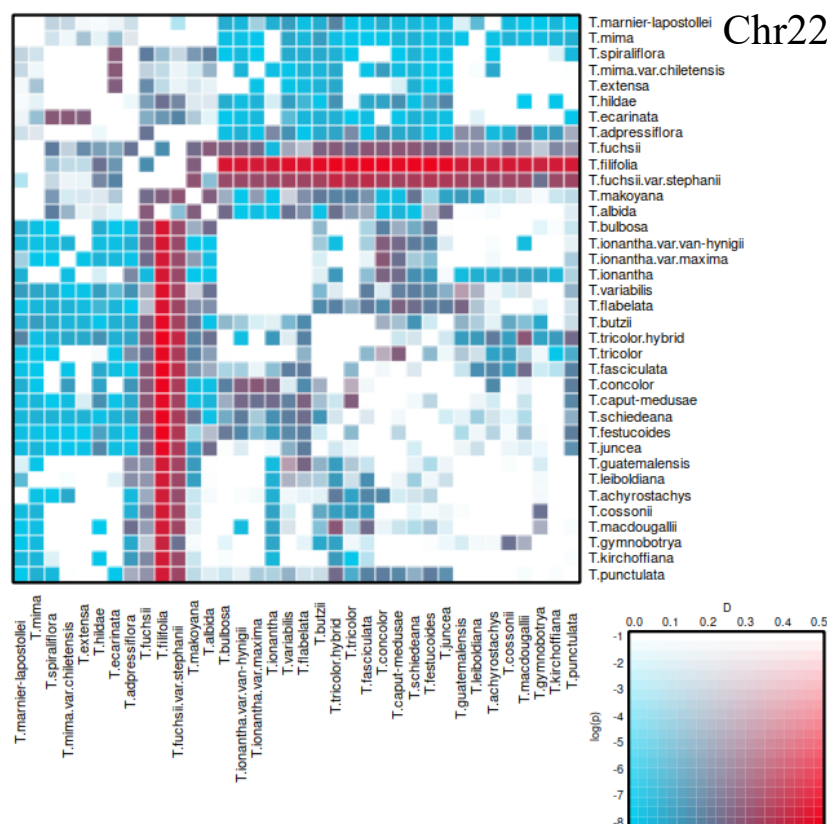
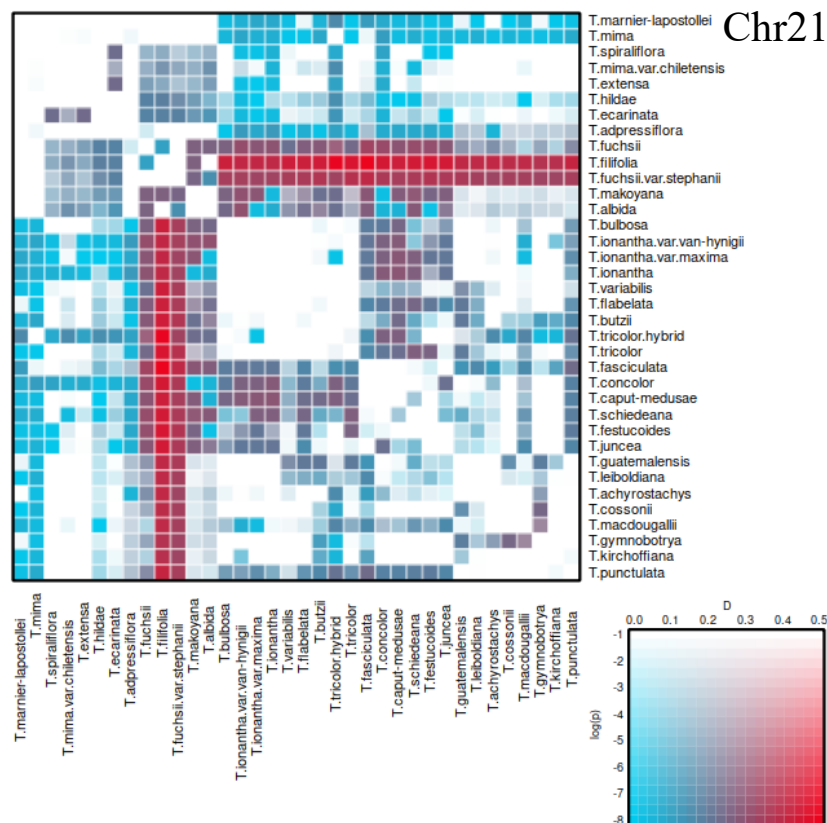


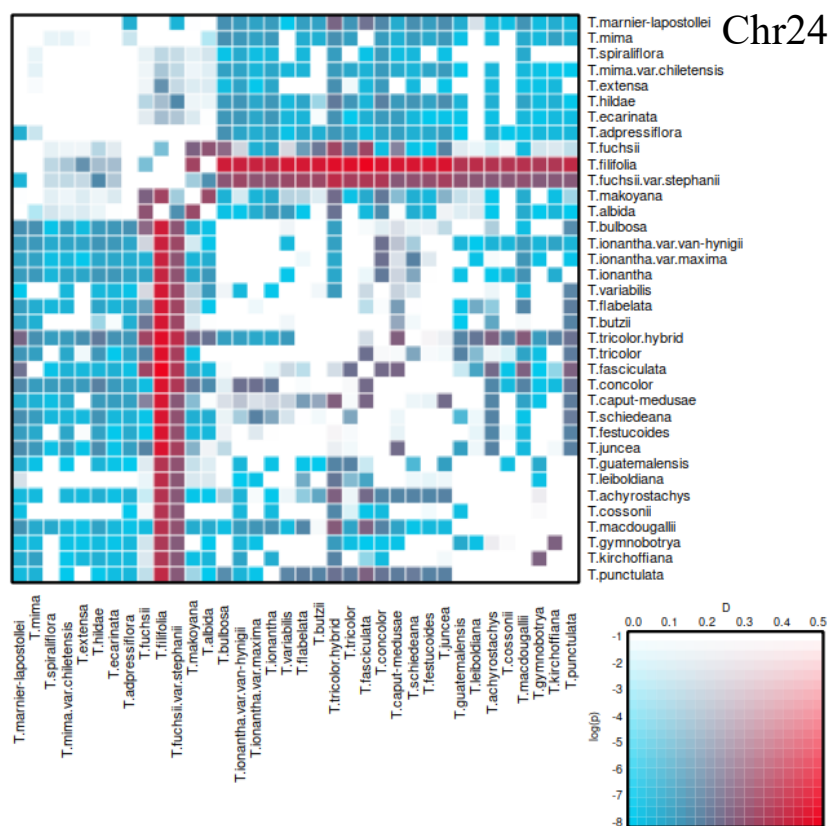
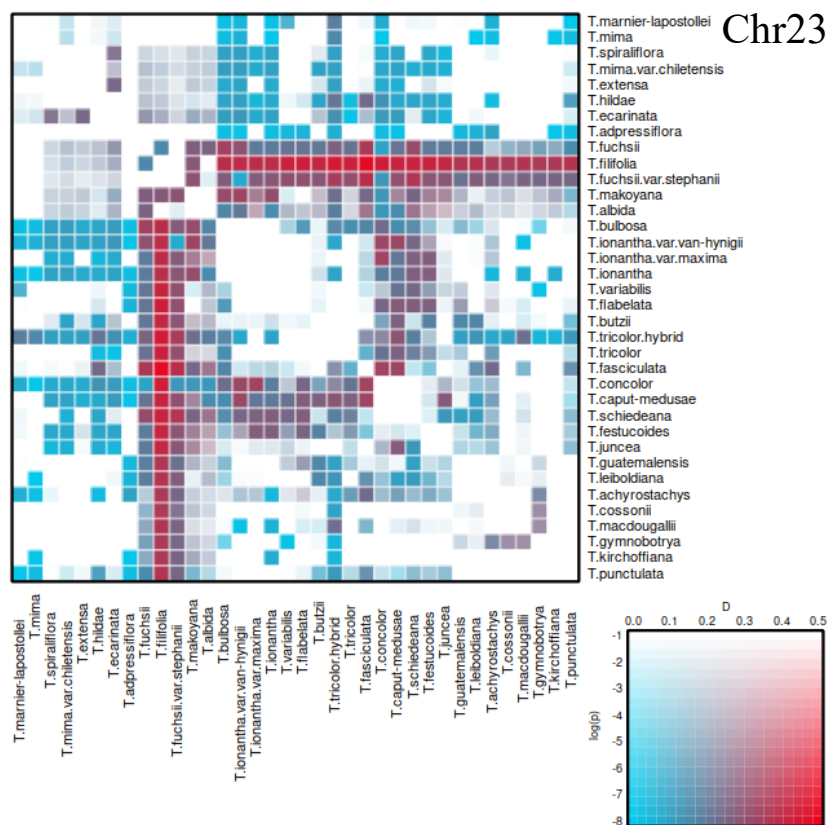


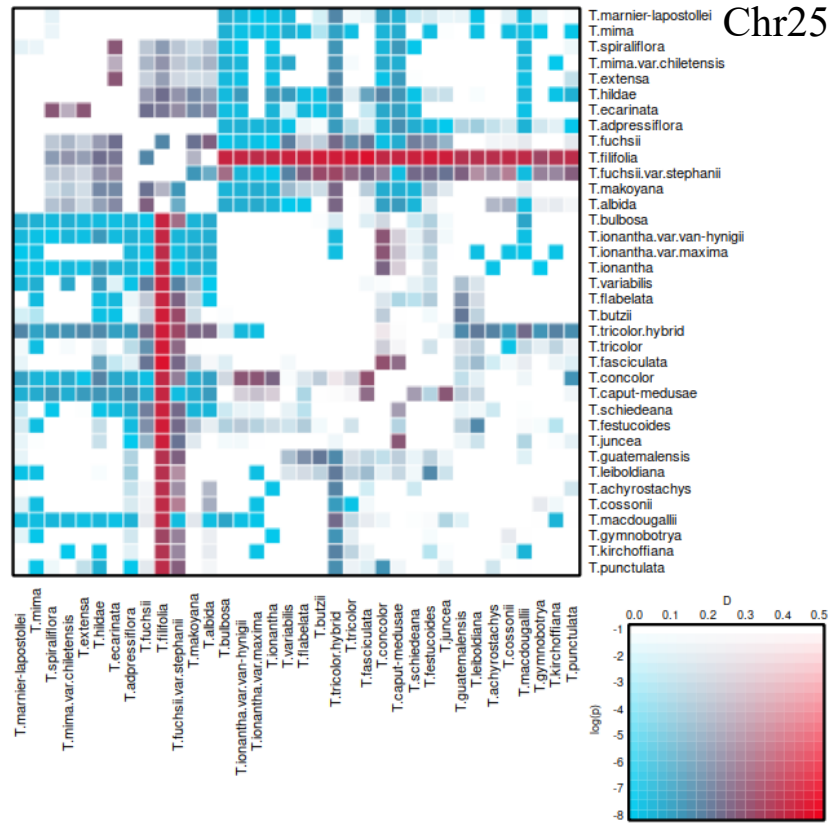












Supporting table S1 list of species sampled for this study with clade assignment, phenotype and ecology (if known). References to corresponding literature are detailed in sheet 2.

species	clade	habit (epiphyte, terrestrial, saxicolous)	leaf succulence/hydroderma	Photosynthesis (CAM / C3) (%)	isotope ratio $\delta^{13}C$ (‰)	leaf	trichome	water tank	bract colors	petals/flower colors	distribution	Ecology (if known)	elevation (meters)	reproductive system	Other	Conservation status	endemism/widespread
<i>T. achyrocyathus</i>	K.1	Epiphytic [1]	rather no	CAM[13]	-14.7[13]	dark green, short and thick[4]	yes[4]	maybe(WT), small[4]	Pink[2]	green (citron)[2]	Mexico[1]	oak forests[3]/desert scrub[3]/palm grove[3], xeric [4]	330-2350 [6]	ornithophilous, SI[7], outcrossing[3]	occur on 35-1 % of B. copallifera trees[2]	vulnerable	Endemic [1] [6]
<i>T. adpressiflora</i>	SA	Epiphytic [14]	no	C3[13][14]	-26.2[13]	green thin leaves 10-15 cm wide[19]	no[7][19]	yes[19]	red/yellow[19]	purple[19]	Bolivia, Venezuela, Colombia, Ecuador, Peru, French Guiana, northern Brazil[15]	epiphyte lower montane forest	0-1000[14]	ornithophilous			widespread[15]
<i>T. albidia</i>	K.2.1	Epiphytic [19]	yes	C3[13]	-21.9[13]	white, smooth leaf surfaces [18]	yes[18]	Intermediate [18]	yellow[1N]	cream[1N]	Mexico[15]	xerophytic [17]	1380-2150[6]	ornithophilous	xerophytic [17]	LC	Endemic [6]
<i>T. bulbosa</i>	K.2.3	epiphytic [20]	yes	CAM[13]	-12.4[13]	thick, green, hard, crooked[16]	yes[21]	no[20]	green[20]	purple (lila) [20]	Belize, Brazil, Caribbean, Colombia, Costa Rica, Ecuador, El Salvador, French Guiana, Guatemala, Honduras, Mexico, Nicaragua, Panama, Suriname, US, Venezuela	Myrmecophyte [16], xenomorphic [20]	0-1030[6]	ornithophilous	Myrmecophyte [16]	LC	widespread
<i>T. butzii</i>	K.2.3	Epiphytic [22] in mostly dry open habitats	yes[22]	CAM[13]	-13.1[13]	thin, hard[22]	yes[22]	no[22]	red/green[1N]	purple (lila)[1N]	Mexico, Panama, Honduras, Guatemala, Nicaragua, El Salvador, Costa Rica, Belize		800-2810[6]	ornithophilous	Myrmecophyte [22]		widespread
<i>T. caput-medusae</i>	K.2.3	epiphytic [7]	yes[22]	CAM[13]	-15.4[13]	thick, twisting, bulbous[22]	yes[22]	no[7][22]	pink[11]	purple (lila)[11]	Belize, Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama	ornithophilous; SI[7]	0-2300[6]	ornithophilous; SI[7]	Myrmecophyte [22]		widespread
<i>T. concolor</i>	K.2.3	Epiphytic [23]	yes	CAM[13]	-18.6[13]	thick, greyish green, red in sun(GY), white, smooth leaf surfaces[18]	yes [18]	small, intermediate[18]	green red (GY)	red pink(GY)	El Salvador, Mexico [6]	exposed habitats, coastal[23]	0-1090 [6]	ornithophilous		LC	Endemic [6]
<i>T. cossonii</i>	K.1	epiphytic	maybe	C3[13]	-22.7[13]	Green-grey, wide and long	yes	yes(GY)/intermediate(WT)	pink[1N]	yellow-green (citron)[1N]	Mexico[24]		1500-2600[6], 3900[24]	ornithophilous			Endemic [6]
<i>T. ecarinata</i>	SA	lithophytic [25]	no	C3[1N]	NA	green, wide, shiny[1N]	no [1N]	Yes [1N]	red	blue-violet	Peru[15]	dry forest[25]	300-700 [25]	ornithophilous		LC	Endemic [25]
<i>T. extensa</i>	SA	saxicolous[25]	yes	CAM[13]	-18.8[13]	thin, wide, green, develop purple foliage[25]	NA	yes[1N]	green/red[25]	dark violet[25]	Peru[25]	xerophytic [25]	500-700[13], 2500 m[25]	ornithophilous		Endemic to N. Peru[25]	
<i>T. fasciculata</i>	K.2.3	Epiphytic [3]	yes[12]	CAM[8][12] [13]	-16.1[12] - 19.9[13]	Greyish thin leaves, thick cuticle[12]	yes[12]	no[12]	red/yellow[3]	purple (lila)[9]	Mexico, Central America, West Indies, N. South America,[15]	juniper forests[3]/palm grove[3]mature forests[10]	0-1860[6]	ornithophilous	xenomorphic [3], occurs in juniper forest[3], mature forests[10]	LC	widespread[6]
<i>T. fuscoides</i>	K.2.3	epiphytic [3]	rather yes	CAM[13]	-14.6[13]	light green, thin and grassy[BS]	NA	no [BS]	red[1N]	purple (lila)[1N]	Belize, Caribbean, Costa Rica, Guatemala, Honduras, Mexico, Nicaragua, Panama [15]	deciduous, evergreen and semi-evergreen	25-1400[6]	ornithophilous			widespread
<i>T. filifolia</i>	K.2.1	epiphytic [28]	No [BS]	C3[13]	-27.2[13]	many thin green leaves, filiform[29]	yes [BS]	no [BS]	green[29]	pale lavender, white[29]	Belize, Costa Rica, Guatemala, Honduras, Mexico, Nicaragua [15]		0-1750[6]	entomophilous	non-tubular flowers		widespread
<i>T. flabellata</i>	K.2.3	epiphytic [1N]	no	CAM[13]	-15.1[13]	green, hard	NA	intermediate	red[1N]	purple (lila)[1N]	Mexico, Guatemala, El Salvador, Honduras, Nicaragua [15]		380-1500[6]	ornithophilous			widespread
<i>T. fuchsii</i>	K.2.1	epiphytic [BS]	somewhat [BS]	C3[13]/ CAM	-21.1[13]	many thin and greyish	yes	no	red[BS]	purple[BS]	Mexico, Guatemala[15]		400-2016[6]	ornithophilous		endangered [BS]	widespread; Endemic [BS]
<i>T. fuchsii</i> , var. <i>steph.</i>	K.2.1	epiphytic [BS]	somewhat	CAM	NA	many thin and greyish	yes	no	red[BS]	purple[BS]	Mexico[15]	deciduous-leaf shrubland	200	ornithophilous		vulnerable	Endemic [15]
<i>T. guatemalensis</i>	K.2.2	epiphytic [30]	no[BS]	C3[13]	-25.3[13]	green, wide and long[BS]	no[BS]	yes[31]	red[BS]	purple (lila)[1N]	Costa Rica, El Salvador, Guatemala, Honduras, Mexico Northeast, Mexico Southeast, Mexico Southwest, Nicaragua, Panama[15]		900-2800[6]	ornithophilous/entomophilous			Widespread [25]
<i>T. gymnoboraya</i>	K.1	Saxicolous and epiphytic, cliffs and forest	no	C3[13]	-23.9[13]	wide, thin, light green greyish	NA	maybe	pink[15]	purple[15]	Mexico - Costa Rica [15]		100-2800[6]	ornithophilous/entomophilous		vulnerable	Endemic [6]
<i>T. liliifolia</i>	SA	Saxicolous[25]	yes	prob. CAM	NA	white cross-banded leaves[BS]	NA		green[BS]	blue-violet[BS]	Peru[25]	dry forest[25], dry valleys[BS]	1000-1100[25]	ornithophilous	xerophyte	LC	Endemic [25]
<i>T. ionantha</i>	K.2.3	epiphytic, occasionally saxicolous[5] or terrestrial [12]	yes[12]	CAM[12]	-13.9[12]	yes[33]	yes[12]	no[12]	red[12]	purple[12]	central Mexico to Nicaragua[5]	desert scrubs[3], xenomorphic [20]	0-1750[6]	ornithophilous; outcrossing[5] or mixed, exerted genitalia[5]	semi-xeric [5]		widespread
<i>T. juncea</i>	K.2.3	Epiphytic and terrestria[12]	yes[12]	CAM[12]	-15.3[12]	thick, narrow, grassy, bluish green	yes[34]	no[12]	green red[3]	purple (lila)[3] [12], pink	Bolivia, Brazil, Caribbean, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Peru, US, Venezuela[41]	juniper forest[3]/ oak forest[3]	0-2500[6]	ornithophilous; Self & cross. hummingbird pollinated[34]	increase in abundance in disturbed forests, cross-hatched pastures, coffee plantations		Widespread [25]

species	clade	habit (epiphyte, terrestrial, saxicolous)	leaf succulence/ hydroderma	photosynthesis (CAM / C3)	isotope ratio $\delta^{13}\text{C}$ (‰)	leaf	trichome	water tank	bract colors	petals/ flower colors	distribution	ecology (if known)	elevation [meters]	reproductive system	Other	Conservation status	endemic/ widespread
<i>T. kirchhoffiana</i>	K.1	epiphytic [GY]	no	C3[13]	-22.9[13]	Green sheets, long fillform. [36]	no[IN]	intermediate[36]	green red[3]	purple (lila)[3]	Mexico[15]	mesic [BS]	1000-2100[6]	ornithophilous	increase in abundance in disturbed forests, isolated trees in pastures,coffee plantations		Endemic [6]
<i>T. leiboldiana</i>	K.2.2	Epiphytic and terrestrial [12]	no[12]	C3[12]	-31.3[12]	soft, green, wide[IN]	no[12]	yes[12]	red[12]	purple[12]	Belize, Costa Rica, El Salvador, Guatemala, Honduras, Mexico Central, Mexico Gulf, Mexico Southeast, Mexico Southwest, Nicaragua, Panamá[15]		800-1900[6]	ornithophilous/? entomophilous			widespread
<i>T. macdougalii</i>	K.1	epiphytic [37]	no[39]	C.37[13]	-20.8[13]	semi-fleshy, ashen-green	yes[37]	intermediate	pink[3]	purple[3]	Mexico[37]	oak forest[3], pine forest, oak-pine forest[37]	1800-3300[6]	ornithophilous; mixed[38]		LC	Endemic [6]
<i>T. makoyana</i>	K.2.1	epiphytic [26]	rather yes	CAM[13]	-13.9[13]	greyish, glaucous-grey	yes[IN]	yes[26]	green red [3]	purple (lila)[3], citron stamens[3]	Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua[15]	oak forest[3]	50-2320[6]	monocarpic, Sc [L], hummingbird pollinated[26]			Endemic [6]
<i>T. marlier-lapostollei</i>	SA	lithophyte	rather yes	CAM[13]	-14.6[13]	large, green-greyish, wide	NA	yes	green	blue-violet	Ecuador[25]	saxicolous	500-1880[13]	ornithophilous		LC	Endemic
<i>T. mina</i>	SA	lithophyte or epiphyte[BS]	yes[BS]	CAM[13]	-15[13]	large, hard, greyish-green[BS]	yes[BS]	yes	carmine-red	violet	Colombia, Ecuador, Peru[15]	saxicolous	650-1900 [BS]	ornithophilous		LC	Endemic
<i>T. mina var. chilensis</i>	SA	Saxicolous[25]	yes	CAM	NA		yes	intermediate	carmine-red	dark-violet[25]	Peru[25]	terrestrial in dry valley	800m[25]	ornithophilous	xyrophyte	LC	Endemic [25]
<i>T. punctulata</i>	K.1	Epiphytic [40]	no	C3[13]	-24.5[13]	thick, narrow, grassy, bluish green	maybe	intermediate[32]	green[BS]	dark purple,black, white apex[BS]	Belize, Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua, Panamá[15]	cloud forest [40]	200-2200[6]	ornithophilous; pref. Outcrops	increase in abundance in disturbed forests, isolated trees in pastures,coffee plantations		widespread
<i>T. schiedana</i>	K.2.3	epiphytic [41]	yes[41]	CAM[13]	-12.1[13]	thin, hard, greyish[BS]	yes[35]	no[7]	red[3]	yellow/green[3]	Belize, Caribbean, Colombia, Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Venezuela[15]	juniper forest[3]/oak forest[3] desert scrub[3]/ palm grove. [3]	0-2230[6]	ornithophilous and SC, high-selfing[7]			Endemic [6]
<i>T. spiratiflora</i>	SA	Saxicolous [25]	rather no	CAM	NA	thin, wide, greyish	NA	intermediate	green[IN]	purple[IN]	Peru[25]	dry forest[25]	1700[25]	ornithophilous		LC	Endemic [25]
<i>T. tricolor</i>	K.2.3	epiphytic	yes[BS]	CAM[13]	-15.2[13]	narrow, green or red, greyish[BS]	yes[BS]	no	red yellow green[BS]	purple[BS]	Belize, Costa Rica, Guatemala, Honduras, Mexico, Nicaragua, Panamá[15]		250-1820[6]	ornithophilous			widespread
<i>T. variabilis</i>	K.2.3	epiphytic [15]	no	CAM[13]	-12.7[13]	soft	no[BS]	yes[42]	green red[BS]	lavender to blue[BS]	Belize, Bolivia, Colombia, Costa Rica, Cuba, Ecuador, Guatemala, Haiti, Honduras, Jamaica, Mexico Central-Southwest, Nicaragua, Panamá, Puerto Rico, Venezuela[15]	amazon forest, Premontane Humid Forest[25]	550-1170[6]	ornithophilous; possibly cleistogamous[29]		Threatened	widespread
<i>T. carnosa var. boliviensis</i>	oungroup	lithophytic [25]	no	C3	NA	thick, green	NA	yes[42]	green	blue-violet	Bolivia[15]	lithophyte	3000	ornithophilous		LC	Endemic
<i>T. complanata</i>	oungroup	epiphytic, lithophytic	no	C3	NA	thin, green	no	yes[42]	red or green	blue (and purple-red)	Bolivia, Peru, northern Brazil, Colombia, Guyana, Costa Rica, Panamá, Cuba, Jamaica, Trinidad, Venezuela, Ecuador [15]		750-3000 [IN]	ornithophilous/ entomophilous	non-terminal blooming bromeliad, axillary inflorescences,	widespread	widespread
														SI - self incompatible		LC - Least concern	
														SC - self compatible			

Corresponding literature for supporting table S1.

code	reference
1	Smith LB, Downs RJ. 1977. Tillandsioideae (Bromeliaceae). Flora Neotropica Monograph, New York: Hafner Press, 14: 663–1494.
2	González-Astorga, J., Cruz-Angon, A., Flores-Palacios, A., & Vovides, A. P. (2004). Diversity and genetic structure of the Mexican endemic epiphyte <i>Tillandsia achyrostachys</i> E. Morr. ex Baker var. <i>achyrostachys</i> (Bromeliaceae). <i>Annals of Botany</i> , 94(4), 545-551.
3	Gomez-Escamilla, N., Espejo-Serna, A., López-Ferrari, A. R., & Krömer, T. (2018). The Epiphytic Bromeliad Flora of the RTP Cerros Negro-Yucaño, Oaxaca, Mexico. <i>Journal of the Bromeliad Society</i> , 68(1), 45-57.
4	Benzing, D. H., & Burt, K. M. (1970). Foliar permeability among twenty species of the Bromeliaceae. <i>Bulletin of the Torrey Botanical Club</i> , 269-279.
5	Soltis, D. E., Gilmartin, A. J., Rieseberg, L., & Gardner, S. (1987). Genetic variation in the epiphytes <i>Tillandsia ionantha</i> and <i>T. recurvata</i> (Bromeliaceae). <i>American Journal of Botany</i> , 74(4), 531-537.
6	Espejo-Serna, A., López-Ferrari, A. R., Ramírez-Morillo, I., Holst, B. K., Luther, H. E., & Till, W. (2004). Checklist of Mexican Bromeliaceae with notes on species distribution and levels of endemism. <i>Selbyana</i> , 33-86.
7	Orozco-Ibarrola, O. A., Flores-Hernández, P. S., Victoriano-Romero, E., Corona-López, A. M., & Flores-Palacios, A. (2015). Are breeding system and florivory associated with the abundance of <i>Tillandsia</i> species (Bromeliaceae)? <i>Botanical Journal of the Linnean Society</i> , 177(1), 50-65.
8	Benzing, D. H. (2000). <i>Bromeliaceae: profile of an adaptive radiation</i> . Cambridge University Press.
9	Cascante-Marín, A., De Jong, M., Borg, E. D., Oostermeijer, J. G. B., Wolf, J. H., & den Nijs, J. C. (2006). Reproductive strategies and colonizing ability of two sympatric epiphytic bromeliads in a tropical premontane area. <i>International Journal of Plant Sciences</i> , 167(6), 1187-1195.
10	Cascante-Marín, A., Wolf, J. H., Oostermeijer, J. G. B., Den Nijs, J. C. M., Sanahuja, O., & Durán-Apuy, A. (2006). Epiphytic bromeliad communities in secondary and mature forest in a tropical premontane area. <i>Basic and Applied Ecology</i> , 7(6), 520-532.
11	Flores-Palacios, A., Bustamante-Molina, A. B., Corona-López, A. M., & Valencia-Díaz, S. (2015). Seed number, germination and longevity in wild dry forest <i>Tillandsia</i> species of horticultural value. <i>Scientia Horticulturae</i> , 187, 72-79.
12	De La Harpe, M., Paris, M., Hess, J., Barfuss, M. H. J., Serrano-Serrano, M. L., Ghatak, A., ... & Lexer, C. (2020). Genomic footprints of repeated evolution of CAM photosynthesis in a Neotropical species radiation. <i>Plant, Cell & Environment</i> , 43(12), 2987-3001.
13	Crayn, D. M., Winter, K., Schulte, K., & Smith, J. A. C. (2015). Photosynthetic pathways in Bromeliaceae: phylogenetic and ecological significance of CAM and C3 based on carbon isotope ratios for 1893 species. <i>Botanical Journal of the Linnean Society</i> , 178(2), 169-221.
14	Medina, E. (1974). Dark CO2 fixation, habitat preference and evolution within the Bromeliaceae. <i>Evolution</i> , 677-686.
15	Kew world checklist - https://powo.science.kew.org/
16	Mooney, J. (2013). A Bromeliad Hotspot in French Guiana. <i>Journal of the Bromeliad Society</i> , 63(3-6), 180-231.
17	Kováčik, J., Klejdus, B., Štork, F., & Hedbavny, J. (2012). Physiological responses of <i>Tillandsia albida</i> (Bromeliaceae) to long-term foliar metal application. <i>Journal of hazardous materials</i> , 239, 175-182.
18	Pierce, S. (2007). The jeweled armor of <i>Tillandsia</i> —Multifaceted or elongated trichomes provide photoprotection. <i>Aliso: A Journal of Systematic and Evolutionary Botany</i> , 23(1), 44-52.
19	Koch, A. K., Monteiro, R. F., & Ilkiu-Borges, A. L. (2015). Checklist de Bromeliaceae da região da volta grande do Xingu, Pará, Brasil. <i>Rodriguésia</i> , 66, 455-464.
20	Guarçoni, E. A., Sousa, J. D., Ferreira, A. W., Junior, W. R., Oliveira, M. S., & Silva, E. O. (2020). New occurrence records of Bromeliaceae for the state of Maranhão, northeastern Brazil. <i>Check List</i> , 16, 1575.
21	Sugden, A. M. (1981). Aspects of the ecology of vascular epiphytes in two Colombian cloud forests. II. Habitat preferences of Bromeliaceae in the Serranía de Macuira. <i>Selbyana</i> , 5(3/4), 264-273.
22	Benzing, D. H. (1970). An investigation of two bromeliad myrmecophytes: <i>Tillandsia butzii</i> Mez, <i>T. caput-medusae</i> E. Morren, and their ants. <i>Bulletin of the Torrey Botanical Club</i> , 109-115.
23	García-Franco, J. G. (1996). Distribución de epífitas vasculares en matorrales costeros de Veracruz, México. <i>Acta Botánica Mexicana</i> , (37), 1-9.
24	Ceja-Romero, J. (Ed.). (2012). <i>Epífitas vasculares del Bajío y de regiones adyacentes</i> . Instituto de Ecología AC, Centro Regional del Bajío.
25	León, B., & Alva, A. S. (2008). General overview of <i>Tillandsia</i> subgenus <i>Tillandsia</i> in Peru: the three-pinnate species and the case of two endemic species. <i>Revista peruana de biología</i> , 15(1), 25-30.
26	Sáyago, R., Quesada, M., Aguilar, R., Ashworth, L., Lopezariza-Mikel, M., & Martén-Rodríguez, S. (2018). Consequences of habitat fragmentation on the reproductive success of two <i>Tillandsia</i> species with contrasting life history strategies. <i>AoB Plants</i> , 10(4), ply038.
27	Olmsted, I., & Juárez, M. G. (1996). Distribution and conservation of epiphytes on the Yucatan Peninsula. <i>Selbyana</i> , 58-70.
28	Krömer, T., Acebey, A., Kluge, J., & Kessler, M. (2013). Effects of altitude and climate in determining elevational plant species richness patterns: a case study from Los Tuxtlas, Mexico. <i>Flora-Morphology, Distribution, Functional Ecology of Plants</i> , 208(3), 197-210.
29	Gardner, C. S. (1986). Inferences about pollination in <i>Tillandsia</i> (Bromeliaceae). <i>Selbyana</i> , 76-87.
30	Wolf, J. H., & Konings, C. J. (2001). Toward the sustainable harvesting of epiphytic bromeliads: a pilot study from the highlands of Chiapas, Mexico. <i>Biological conservation</i> , 101(1), 23-31.
31	Aranda-Coello, J. M., Mendoza Velázquez, O. M., Gómez Cruz, A., Cervantes Díaz, J., Villalobos Escobar, C. A., Moctezuma Hernández, N., & Montoya Cabrera, E. A. (2018). Ausencia de anfibios en bromelias de tanque en la Estación Biológica San José, Chiapas, México. <i>Cuadernos de Investigación UNED</i> , 10(1), 227-229.
32	Hietz, P., Ausserer, J., & Schindler, G. (2002). Growth, maturation and survival of epiphytic bromeliads in a Mexican humid montane forest. <i>Journal of Tropical Ecology</i> , 18(2), 177-191.
33	Nowak, E. J., & Martin, C. E. (1997). Physiological and anatomical responses to water deficits in the CAM epiphyte <i>Tillandsia ionantha</i> (Bromeliaceae). <i>International Journal of Plant Sciences</i> , 158(6), 818-826.
34	Mosti, S., Ross Friedman, C., Pacini, E., Brighigna, L., & Papini, A. (2013). Nectary ultrastructure and secretory modes in three species of <i>Tillandsia</i> (Bromeliaceae) that have different pollinators. <i>Botany</i> , 91(11), 786-798.
35	Chilpa-Galván, N., Márquez-Guzmán, J., Zotz, G., Echevarría-Machado, I., Andrade, J. L., Espadas-Manrique, C., & Reyes-García, C. (2018). Seed traits favouring dispersal and establishment of six epiphytic <i>Tillandsia</i> (Bromeliaceae) species. <i>Seed Science Research</i> , 28(4), 349-359.
36	Guzmán-Jacob, V. (2013). Diversidad de macroartrópodos asociados a tres especies de tillandsia (Bromeliaceae) en Tlalnelhuayocan, Veracruz.
37	Vázquez Hurtado, N. B. (2014). Germinación y propagación de <i>Tillandsia macdougallii</i> LB Sm y <i>Tillandsia violacea</i> Baker (Bromeliaceae) con fines de aprovechamiento sustentable (Master's thesis).
38	Mondragon, D., Valverde, T., & Hernández-Apolinar, M. (2015). Population ecology of epiphytic angiosperms: a review. <i>Tropical Ecology</i> , 56(1).
39	Ticktin, T., Mondragón, D., & Gaoe, O. G. (2016). Host genus and rainfall drive the population dynamics of a vascular epiphyte. <i>Ecosphere</i> , 7(11), e01580.
40	Toledo-Aceves, T., García-Franco, J. G., Lozada, S. L., Mateos, M. L. L., & MacMillan, K. (2012). Germination and seedling survivorship of three <i>Tillandsia</i> species in the cloud-forest canopy. <i>Journal of Tropical Ecology</i> , 28(4), 423-426.
41	Loeschen, V. S., Martin, C. E., Smith, M., & Eder, S. L. (1993). Leaf anatomy and CO2 recycling during crassulacean acid metabolism in twelve epiphytic species of <i>Tillandsia</i> (Bromeliaceae). <i>International Journal of Plant Sciences</i> , 154(1), 100-106.
42	Judith, C., Schneider, J. V., Schmidt, M., Ortega, R., Gaviria, J., & Zizka, G. (2013). Using high-resolution remote sensing data for habitat suitability models of Bromeliaceae in the city of Mérida, Venezuela. <i>Landscape and urban planning</i> , 120, 107-118.
IN	blog/forum
BS	bromeliad society website and articles
GY	field observations, G. Yardeni
none	Walter Till, private correspondence

Supporting table S2 sampled accessions in this study with collection, locality and voucher information. samples marked grey were removed from analysis (see table 3)

name_in_project	Species	clade	Collection_name	DNA_code in M. Baruss collection	Garden_code	Det.	Collector	Country	State	Locality	Elevation (m)	Coordinates	notes
ach58	<i>T. achyrostachys</i>	K.1	LSM_3708	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1400	19°39'20.17" N, 104°19'45.86" W	Camino Viejo a Corralitos
ach9	<i>T. achyrostachys</i>	K.1	LSM_370A	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1400	19°39'20.17" N, 104°19'45.86" W	Camino Viejo a Corralitos
adp50	<i>T. adpressiflora</i>	SA	BR3.2015 (8951)	na	na	na	na	Brazil					
alb59	<i>T. albida</i>	K.2.1	0000071[B12/B7]		BG Wien (HBV) 0000271 (B12/B7)	W. Till, M.H.J. Barfuss 2017	R. & F. Wolf	Mexico					
alb269	<i>T. albida</i>	K.2.1	0000269[B2/91]	MHJB-B370	BG Wien (HBV) 0000269 (B2/91)	W. Till, M.H.J. Barfuss 2017	R. & F. Wolf	Mexico	Hidalgo	Venados Valley	no data		
bul16	<i>T. bulbosa</i>	K.2.3	OL010_f	na	na	na	na	Costa Rica		La Selva Biological Station	77	10.43375, -77 84.0124444	Hromadnik; HR18 (Blass: s. n.)
bul29	<i>T. bulbosa</i>	K.2.3	HBVSN/B1222	MHJB-B1222	Hromadnik HR18	L. Hromadnik 2017	no data	Colombia	no data	Reserva de la Biósfera El Triunfo, Chiapas	no data	15°38'07.3" N, 89°24'48.14.3" W	Senderos Costa, Prusia
but3	<i>T. butzii</i>	K.2.3	LSM_132C	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Chiapas		1800	15°38'07.3" N, 89°24'48.14.3" W	
but87	<i>T. butzii</i>	K.2.3	B833	MHJB-B833	(field collection)	J.P. Pinzón	J.P. Pinzón	Mexico	Chiapas	Al O de Teopisca			
cap32	<i>T. caput-medusae</i>	K.2.3	LSM_369C	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1400	19°39'20.17" N, 104°19'45.86" W	Camino Viejo a Corralitos
cap61	<i>T. caput-medusae</i>	K.2.3	LSM_369E	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1400	19°39'20.17" N, 104°19'45.86" W	Camino Viejo a Corralitos
car92	<i>T. carnosa</i> var. <i>boliviensis</i>	SA	UDI0002	MHJB-B1865	BG Wien (HBV) 0029144	A. Seringer, W. Till, M.H.J. Barfuss 2017	A. Seringer	Bolivia	La Paz	Mecapca			
cat27	<i>T. cf. Tricolor.fasciculata</i>	K.2.3	GY084	na	na	W. Till, Jul 2017	G. Yardeni	Mexico	Veracruz	Cerro Macuilteptl		-4.112886, -90.78968189	putative hybrid
com90	<i>T. complanata</i>	outgroup	IO041	na	na	I. Olivares	I. Olivares	Ecuador		Bombuscaro	1200	19°35'58.0" N, 78°22'49.1" W	
con14	<i>T. concolor</i>	K.2.3	GY133	na	na	G. Yardeni	G. Yardeni	Mexico	Veracruz	La Mancha biological station			
cos38	<i>T. cossonii</i>	K.1	LSM_381C	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	2000	19°35'11.9" N, 104°15'57.9" W	Las Joyas
cos39	<i>T. cossonii</i>	K.1	LSM_365C	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1500	19°39'05.2" N, 104°20'06" W	Tepehuaje, camino viejo a Corralitos
eca88	<i>T. ecarinata</i>	SA	UDI0006	MHJB-B1871	Hromadnik HR23020	L. Hromadnik	H. & L. Hromadnik HR 23020	Peru	Cajamarca	Rock wall above Pucara	1100		
ext89	<i>T. extensa</i>	SA	UDI0005	MHJB-B1870	Hromadnik HR4198 (2)	L. Hromadnik	H. & L. Hromadnik HR 4198	Peru	Anazonas	Abra Chanchillo - Balsas	2000		
ext90	<i>T. extensa</i>	SA	UDI0007	MHJB-B1878	BG Wien (HBV) 0024653 (B223/91)	no data	no data	Peru	no data	no data	no data		
ext91	<i>T. extensa</i>	SA	UDI0010	MHJB-B1882	Hromadnik HR4198 (1)	L. Hromadnik	H. & L. Hromadnik HR 4198	Peru	Anazonas	Abra Chanchillo - Balsas	2000		
fas26	<i>T. fasciculata</i>	K.2.3	GY004	na	na	G. Yardeni	G. Yardeni	Mexico	Veracruz	Agüita Fría			disturbed forest
fas30	<i>T. fasciculata</i>	K.2.3	HBV26950	MHJB-B1839	BG Wien (HBV) 0026950 (B103/94)	W. Till, M.H.J. Barfuss 2017	G. Noller 9106	Guatemala	Sololá	Lago de Atitlán	no data		
fas24655	<i>T. fasciculata</i>	K.2.3	HBV24655	MHJB-B1841	BG Wien (HBV) 0024655 (B99/91)	W. Till, M.H.J. Barfuss 2017	W. & S. Till WT 7004	Costa Rica	Heredia	Barva north of Heredia	1200		
fas24656	<i>T. fasciculata</i>	K.2.3	HBV24656	MHJB-B360	BG Wien (HBV) 0024656 (B108/94)	W. Till, M.H.J. Barfuss 2017	R. Ehlers EM 890701	Mexico	Chiapas	Sumidero Canon at Tuxtla Gutiérrez, El Mirador km 17	1200		
fas24657	<i>T. fasciculata</i>	K.2.3	HBV24657	MHJB-B1840	BG Wien (HBV) 0024657 (B179/91)	W. Till, M.H.J. Barfuss 2017	W. & S. Till WT 7116	Costa Rica	Puntarenas	SW declivities of Cordillera de Tilaram, along the road from Sta. Elena to Rancho Grande, at Guaría	750		

name_in_project	Species	clade	Collection_name	DNA_code in M. Barfuss collection	Garden_code	Det.	Collector	Country	State	Locality	Elevation (m)	Coordinates	notes
fas25334	<i>T. fasciculata</i>	K.2.3	HBV25334	MHJB-B1838	BG Wien (HBV) 0025334 (B99B53-1)	W. Till, M.H.J. Barfuss	Kamm	Honduras	no data	no data	no data		
fas25390	<i>T. fasciculata</i>	K.2.3	HBV25390	MHJB-B1837	BG Wien (HBV) 0025390 (B977/94)	W. Till, M.H.J. Barfuss	W. Weiz 3132	Guatemala	Escuintla	Sta. Teresa Esquintla	no data	10.433750, -77.84.012444	
fes15	<i>T. festuoides</i>	K.2.3	OL009c	na	na	O. Loiseau	O. Loiseau	Costa Rica	Heredia	La Selva Biological Station	77	10.433750, -77.84.012444	
fes64	<i>T. festuoides</i>	K.2.3	OL009e	na	na	O. Loiseau	O. Loiseau	Costa Rica	Heredia	La Selva Biological Station	77	10.433750, -77.84.012444	
fil49	<i>T. filifolia</i>	K.2.1	MHB08b	na	na	M. de la Harpe	M. de la Harpe	Mexico	Oaxaca	San Juan Bautista Valle Nacional	300	17.824333, -96.2437	
fil65	<i>T. filifolia</i>	K.2.1	MHB08a	na	na	M. de la Harpe	M. de la Harpe	Mexico	Oaxaca	San Juan Bautista Valle Nacional	300	17.824333, -96.2437	
fla51	<i>T. flabelata</i>	K.	LSM_273c	na	na	E. Espejo-Serna, Jan. 2017	K. Silva-Mijangos	Mexico	Chiapas	Reserva de la Biósfera El Triunfo, Chiapas	1250	15°42'55" N, 92°51'21.3" W	
fla94	<i>T. flabelata</i>	K.	LSM_273a	na	na	E. Espejo-Serna, Jan. 2017	K. Silva-Mijangos	Mexico	Chiapas	Reserva de la Biósfera El Triunfo, Chiapas	1250	15°42'55" N, 92°51'21.3" W	
fuc24	<i>T. fuchsii</i>	K.2.1	LSM_170A	na	na	E. Espejo-Serna, Jan. 2017	L. Silva-Mijangos	Mexico	Chiapas	Reserva de la Biósfera El Triunfo, Chiapas	1600	15°37'49.7" N, 92°48'12.1" W	UTRECHT: E.J. Gouda (Bg. U. Staff), 32
fuc53	<i>T. fuchsii</i>	K.2.1	B1726	MHJB-B1726	BG Utrecht, 1997GR00709	U: 1997GR00709		Mexico					
fus4	<i>T. fuchsii var. stephanii</i>	K.2.1	LSM_368F	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1400	19°39'20.17" N, 104°19'45.86" W	
gua19	<i>T. guatemalensis</i>	K.2.2	LSM_15D	na	na	E. Espejo-Serna, Jan. 2017	L. Silva-Mijangos	Mexico	Chiapas	Reserva de la Biósfera El Triunfo, Chiapas	2000	15°39'32.4" N, 92°48'28.5" W	
gua95	<i>T. guatemalensis</i>	K.2.2	LSM_15C	na	na	E. Espejo-Serna, Jan. 2017	L. Silva-Mijangos	Mexico	Chiapas	Reserva de la Biósfera El Triunfo, Chiapas	2000	15°39'32.4" N, 92°48'28.5" W	
gym17	<i>T. gymnobotrya</i>	K.1	MHD27C	na	na	M. de la Harpe	M. de la Harpe	Mexico	Veracruz	Coatepec	2040	19.498694, -96.97147	
gym18	<i>T. gymnobotrya</i>	K.1	MH124c	na	na	M. de la Harpe	M. de la Harpe	Mexico	Oaxaca	San Felipe Usila	2210	17.68052, -96.5533	
hil98	<i>T. hildae</i>	SA	UDI009	MHJB-B1880	BG Wien (HBV) 0027606 (B279/95)	W. Till, M.H.J. Barfuss	no data	Peru	Cajamarca	no data	no data		BG Linz
ion219	<i>T. ionantha</i>	K.2.3	HBV219	MHJB-B1845	BG Wien (HBV) 0030230 (B219/91)	W. Till, M.H.J. Barfuss	R. Ehlers EM 911306	Mexico	no data	100 km befor Altamirano	no data		
iom24697	<i>T. ionantha var. maxima</i>	K.2.3	HBV24697/B279	MHJB-B1844	BG Wien (HBV) 0024697 (B279/95)	W. Till, M.H.J. Barfuss	J. Lautner	Mexico	Oaxaca	Near the town of Tehuantepec, Rio Huamelula near San Pedro Huamelula, epilitnic	100		
iov437	<i>T. ionantha var. van-hymn</i>	K.2.3	HBV437	MHJB-B439	BG Wien (HBV) 0025358 (437/82)	W. Till, M.H.J. Barfuss	no data	Mexico	Chiapas	Rio Grijalva	no data		
jun6	<i>T. juncea</i>	K.2.3	LSM366D	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1500	19°39'05.2" N, 104°20'06" W	Tepehuaje, camino viejo a Corralitos
kir28	<i>T. kirchoffiana</i>	K.1	B1308	MHJB-B1308	BG Wien (HBV) 0026412 (B95/9)	W. Till, M.H.J. Barfuss	K-D. & R. Ehlers EM942601	Mexico	Puebla	Xicotepec de Juárez - La Union	1200	19°31'14.9" N, 98°59'23.4" W	disturbed forest
kir56	<i>T. kirchoffiana</i>	K.1	G125	na	na	V. Guzman	G. Yardeni	Mexico	Veracruz	Agüita Fría			
lei31	<i>T. leiboldiana</i>	K.2.2	HBV24715	MHJB-B1842	BG Wien (HBV) 0024715 (B82/91)	W. Till	W. & S. Till 7112	Costa Rica	Alajuela	2 km N San Ration	1030		
lei57	<i>T. leiboldiana</i>	K.2.2	B843	MHJB-B843	(field collection)	J.P. Pinzón	J.P. Pinzón & al. 120	Mexico	Chiapas	Tunapaz			
mac78	<i>T. macedougallii</i>	K.1	LSM385A	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	2800	19°33'20.4" N, 104°09'00.8" W	
mac79	<i>T. macedougallii</i>	K.1	LSM385B	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	2800	19°33'20.4" N, 104°09'00.8" W	
mak22	<i>T. makoyana</i>	K.2.1	LSM374B	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1500	19°39'05.2" N, 104°20'06" W	
mak80	<i>T. makoyana</i>	K.2.1	LSM374C	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1500	19°39'05.2" N, 104°20'06" W	

name_in_project	Species	clade	Collection_name	DNA_code in M. Barfuss collection	Garden_code	Det.	Collector	Country	State	Locality	Elevation (m)	Coordinates	notes
mar21	<i>T. marnier-lapostollei</i>	SA	UDI008	MHJB-B1879	BG Wien (HBV) 0024729 (B98887-1)	W. Till, M.H.J. Barfuss	W. Rauh R 35345	Peru	Cajamarca	at Rio Chancay, between km 6 & km 15 at road to Santa Cruz. E Rio Chancay bridge	600-800		
mim11	<i>T. mina</i>	SA	UDI001	MHJB-B1864	Hromadnik HR17078 (1)	L. Hromadnik	H. & L. Hromadnik HR 17078	Ecuador	Azuay	Gualaceo - Remigio Crespo	2500		
mim20417	<i>T. mina</i> var. <i>chiletensis</i>	SA		MHJB-B1438	BG Wien (HBV) 0020417	W. Rauh	W. Rauh R 69103	Peru	Cajamarca	Magdalena District, Huaquillas, valley of Rio Magdalena	1000		
pun2	<i>T. punctulata</i>	K.1	MH016_h	na	na	M. de la Harpe	M. de la Harpe	Mexico	Veracruz	San Andres Tuxtla	1105	18.54803, -95.15098	
pun5	<i>T. punctulata</i>	K.1	LSM_147C	na	na	A. Hernández-Rojas, Dec. 2016	L. Silva-Mijangos	Mexico	Chiapas	Reserva de la Biósfera El Triunfo, Chiapas	1800	15°38'08.1" N, 92°48'14" W	
sch12	<i>T. schiedeana</i>	K.2.3	MH076_B1	na	na	M. de la Harpe	M. de la Harpe	Mexico	Oaxaca	San Juan Bautista Valle Nacional	300	17.824333, -96.243375	
sch41	<i>T. schiedeana</i>	K.2.3	LSM373d	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1500	19°39'05.2" N, 104°20'06" W	
sch83	<i>T. schiedeana</i>	K.2.3	LSM373c	na	na	L. Silva-Mijangos, Jun. 2017	L. Silva-Mijangos	Mexico	Jalisco	Reserva de la Biósfera Sierra de Manantlán, Jalisco	1500	19°39'05.2" N, 104°20'06" W	
spi67	<i>T. spiralliflora</i>	SA	UDI003	MHJB-B1866	Hromadnik HR4203 (1)	L. Hromadnik	H. & L. Hromadnik HR 4203	Peru	Amazonas	km 355, E Balsas	1700		
spi68	<i>T. spiralliflora</i>	SA	UDI004	MHJB-B1867	Hromadnik HR4203 (2)	L. Hromadnik	H. & L. Hromadnik HR 4203	Peru	Amazonas	km 355, E Balsas	1700		
tri42	<i>T. tricolor</i>	K.2.3	OI053C	na	na	O. Loiseau	O. Loiseau	Costa Rica		Estación Biológica Las Cruces	1134	8.7934722, -82.97138	
tri8	<i>T. tricolor</i>	K.2.3	OI053_A	na	na	O. Loiseau	O. Loiseau	Costa Rica		Estación Biológica Las Cruces	1134	8.7934722, -82.97138	
var43	<i>T. variabilis</i>	K.2.3	B1181	MHJB-B1181	(field collection)	J.P. Pinzón	I. M. Ramirez M. 1798						
zoq45	<i>T. zoquensis</i>	K.2.3	LSM188A	na	na	E. Espejo-Serna, Jan. 2017	L. Silva-Mijangos	Mexico	Chiapas	Reserva de la Biósfera El Triunfo, Chiapas	1200	15°36'50" N, 92°49'21" W	
zoq46	<i>T. zoquensis</i>	K.2.3	LSM188C	na	na	E. Espejo-Serna, Jan. 2017	L. Silva-Mijangos	Mexico	Chiapas	Reserva de la Biósfera El Triunfo, Chiapas	1200	15°36'50" N, 92°49'21" W	

Supporting table S3 samples accessions with information about sequence read number, alignment rates and number of sequences after filtering. samples marked grey were removed from analysis

Number in project	Species	clade	read (total)	Coverage sam	read (bowtie reports)	Alignment rate [bowtie.%]	reads_High Q	reads_High Q_%	coverage_High Q	coverage_High Q_%	notes
ach58	<i>T. achyrostachys</i>	K.1	52788378	6.90072	26295156	80.09	29026704	0.5499	4.9	0.12	
ach9	<i>T. achyrostachys</i>	K.1	80291160	9.67788	39292670	89.36	46563014	0.5925	6.54	0.02	
adp50	<i>T. adpressiflora</i>	SA	254468532	26.581	127234266	79.76	114344720	0.4493	14.4	0.14	
alb269	<i>T. albida</i>	K.2.1	111574508	13.3495	55627601	90.09	71520128	0.6428	9.75	0.04	
alb59	<i>T. albida</i>	K.2.1	30988234	4.55497	15346023		19696207	0.6356	3.29	4.55	Removed - kinship
bul16	<i>T. bulbosa</i>	K.2.3	84840706	9.88809	NA	NA	45316108	0.5341	6.44	5719	
bul29	<i>T. bulbosa</i>	K.2.3	38074934	4.77021	18364059	92.08	23750445	0.6467	3.37	0.02	
but3	<i>T. butzii</i>	K.2.3	111336080	6.4598	NA	NA	61116328	0.5489	8.73	0.01	
but87	<i>T. butzii</i>	K.2.3	114263320	17.4267	56988347	90.69	76935744	0.675	13.1	0.18	
cap32	<i>T. caput-medusae</i>	K.2.3	35151386	4.18141	15605180	95.49	20163869	0.5736	2.85	0.01	
cap61	<i>T. caput-medusae</i>	K.2.3	27891184	6.42161	22621489	84.65	27891184	0.6165	4.74	0.15	
car92	<i>T. carnosa</i> var. <i>boliviensis</i>	outgroup	74302808	9.54226	36866422	84.99	36571892	0.496	5.96	0.14	
cat27	<i>T. cf. tricolor</i> (<i>fasciculata</i>)	K.2.3	22654846	2.62	10781334	85.82	4711113	0.2185	1.91	0.01	
com90	<i>T. complanata</i>	outgroup	138144334	14.4118	69072167	78.72	22448508	0.1625	1.32	0.12	
con14	<i>T. concolor</i>	K.2.3	43808582	6.14989	NA	NA	27640577	0.6309	12.9	5013	
cos38	<i>T. cossonii</i>	K.1	50458014	6.06038913	22947865	85.90	29486934	0.6425	4.18	0.02	
cos39	<i>T. cossonii</i>	K.1	65645928	7.92850913	30234425	88.38	38758323	0.641	5.49	0.03	
eca88	<i>T. ecarinata</i>	SA	82002040	11.5218	40297289	90.66	48138176	0.5973	8	0.14	
ext89	<i>T. extensa</i>	SA	98455094	13.7545	48372944	90.52	57308431	0.5924	9.5	0.14	
ext90	<i>T. extensa</i>	SA	67806390	9.54891	33538475	90.37	39655139	0.5912	6.59	0.14	
ext91	<i>T. extensa</i>	SA	94123032	13.0489	46140271	90.05	54138476	0.5867	8.98	0.15	
fas24655	<i>T. fasciculata</i>	K.2.3	111924314	14.7251	55775701	91.60	73296850	0.6571	10.6	0.02	
fas24656	<i>T. fasciculata</i>	K.2.3	119276338	15.9522	59439920	94.33	80802242	0.6797	11.5	0.02	
fas24657	<i>T. fasciculata</i>	K.2.3	99528422	13.5494	49598337	94.39	69184879	0.6975	10	0.02	
fas25334	<i>T. fasciculata</i>	K.2.3	116009950	14.64	57852756	88.98	74013253	0.6397	10.6	0.03	
fas25390	<i>T. fasciculata</i>	K.2.3	118874624	15.9264	59254274	93.95	80811228	0.6819	11.6	0.02	
fas26	<i>T. fasciculata</i>	K.2.3	50581700	10.61466	40331611	96.55	56465112	0.7	8.06	0.01	
fas30	<i>T. fasciculata</i>	K.2.3	113808178	15.266	56724350	94.59	76324527	0.6728	10.9	0.02	
fes15	<i>T. festuoides</i>	K.2.3	137091043	20.5653	NA	NA	24503411	0.1787	12.9	10377	
fes64	<i>T. festuoides</i>	K.2.3	34387048	4.87198	16854289	85.74	21227094	0.6297	3.62	0.12	
fil49	<i>T. filifolia</i>	K.2.1	48655424	5.53092	23756120	73.21	21235876	0.4365	3.56	0.15	
fil65	<i>T. filifolia</i>	K.2.1	41312422	6.00225	20323674	93.29	23202264	0.5708	3.87	0.12	
fla51	<i>T. flabelata</i>	K.2.3	66866856	7.70898	32909030	69.43	33282888	0.4977	5.7	0.14	
fla94	<i>T. flabelata</i>	K.2.3	110010160	17.4969	54861404	94.90	75017296	0.6837	12.8	0.17	
fuc24	<i>T. fuchsii</i>	K.2.1	79838208	9.61418	NA	NA	41364101	0.5181	5.94	12125	
fuc53	<i>T. fuchsii</i>	K.2.1	64787956	7.76007	31729566	77.11	30814205	0.4756	5.13	0.13	
fus4	<i>T. fuchsii</i> var. <i>stephanii</i>	K.2.1	62087942	7.06142	28021332	93.35	34019031	0.5479	4.76	0.02	
gua19	<i>T. guatemalensis</i>	K.2.2	172102484	20.52466	81411367	91.80	113288052	0.6583	15.8	0.05	
gua95	<i>T. guatemalensis</i>	K.2.2	87904454	13.6799	43767379	96.57	64734545	0.7395	10.7	0.17	
gym17	<i>T. gymnobotrya</i>	K.1	93638542	11.7848	NA	NA	57063301	0.6094	8.07	0.03	
gym18	<i>T. gymnobotrya</i>	K.1	101435088	14.9781	NA	NA	70412561	0.6942	9.9	0.01	
hil98	<i>T. hildae</i>	SA	70949240	9.99907	34997243	90.68	41740743	0.5963	6.93	0.14	
iom24697	<i>T. ionantha</i> var. <i>maxima</i>	K.2.3	132065572	15.0519	65855931	82.73	80799847	0.6135	11.3	0.03	
ion219	<i>T. ionantha</i>	K.2.3	124883972	13.1334	57556828	82.48	68984356	0.5993	9.67	0.03	
iov437	<i>T. ionantha</i> var. <i>van-hynigii</i>	K.2.3	109934468	13.5141	54682124	88.95	72554789	0.66	10.2	0.03	
jun6	<i>T. juncea</i>	K.2.3	110743706	16.5723	NA	NA	22571180	0.2038	11.2	6804	
kir28	<i>T. kirchoffiana</i>	K.1	53096276	5.19772	NA	NA	36782664	0.6928	5.23	9556	
kir56	<i>T. kirchoffiana</i>	K.1	81479690	12.4275	39963671	94.07	55756859	0.6843	9.45	0.17	

lei31	<i>T. leiboldiana</i>	K.2.2	150161252	18.8141	NA	NA	118514539	0.7892	15.6	0.04
lei57	<i>T. leiboldiana</i>	K.2.2	42676818	6.45938	20975923	94.77	29498334	0.6912	4.93	0.13
mac78	<i>T. macedougallii</i>	K.1	39408152	6.15202	19643666	96.22	27741194	0.7061	4.64	0.12
mac79	<i>T. macedougallii</i>	K.1	10078600	1.57143	5023495		7091601	0.7058	1.19	0.12 Removed – low coverage
mak22	<i>T. makoyana</i>	K.2.1	106064342	14.4793	NA	NA	21372571	0.2015	9.02	1177
mak80	<i>T. makoyana</i>	K.2.1	72766412	11.0298	36053208	94.89	45989576	0.6378	7.73	0.16
mar21	<i>T. marnier-lapostollei</i>	SA	75255838	10.9557	37299312	92.62	46113479	0.6182	7.67	0.14
mim11	<i>T. mimia</i>	SA	125250164	18.1102	62170380	92.85	78647362	0.6325	12.9	0.11
mim20417	<i>T. mimia var. chiletensis</i>	SA	117572260	13.5333	58627406	87.38	70473252	0.601	9.67	0.04
pun2	<i>T. punctulata</i>	K.1	97971356	9.41656	NA	NA	45184568	0.4612	6.38	17001
pun5	<i>T. punctulata</i>	K.1	63141784	5.65639	NA	NA	26733658	0.4234	3.85	0.13
sch12	<i>T. schiedeana</i>	K.2.3	25957066	3.30669	12437299	94.09	16146387	0.6491	2.29	0.01
sch41	<i>T. schiedeana</i>	K.2.3	80657614	13.208	39910790	97.50	56261863	0.7048	9.64	0.13
sch83	<i>T. schiedeana</i>	K.2.3	71166278	11.2028	35168539		47982660	0.6822	8.23	0.14 Removed – kinship
spi67	<i>T. spiraliflora</i>	SA	91525124	12.9233	45379698	90.50	54596826	0.6016	9.06	0.15
spi68	<i>T. spiraliflora</i>	SA	93990912	13.141	46266543	90.32	55228655	0.5969	9.16	0.16
tri42	<i>T. tricolor</i>	K.2.3	101139726	16.1895	49526526	96.74	68878107	0.6954	11.8	0.2
tri8	<i>T. tricolor</i>	K.2.3	65988912	8.85426	32215167	96.87	42448855	0.6588	6.06	0.01
var43	<i>T. variabilis</i>	K.2.3	127220668	20.2525	NA	NA	88987802	0.6995	15.1	0.15
zoq45	<i>T. zoquensis</i>	K.2.3	12345936	1.90631	6039269	93.11	8049613	0.652	1.38	0.14 removed from analysis
zoq46	<i>T. zuquensis</i>	K.2.3	37574672	5.87874	18465772	93.72	25076729	0.6674	4.31	0.16 removed from analysis

Conclusions and outlook

Recently and rapidly radiating lineages are fascinating to study as they showcase divergence at the intersection of evolution, ecology and adaptation. Their young age provides a glimpse into an early point of the ‘speciation continuum’, when the footprint of genomic processes is still fresh. Nevertheless they remain one of the biggest challenges in the study of the tree of life, posing obstacles that are seldom simply resolved by ‘applying more data’.

The use of NGS provided unprecedented resolution for *Tillandsia* phylogenomics and was especially successful in species delimitation and inference of subgeneric relationships. *Tillandsia* is a textbook example of rapid radiation: their phylogeny is characterized by short internal nodes with high rates of gene tree discordance which coincides with morphological innovation (Kandziora et al., 2022; Parins-Fukuchi et al., 2021). While gene tree discordance most likely resulted from several entangled genomic processes, **chapter 2** of this thesis provided evidence for the extensive role of both ancient and recent hybridization in gene-tree conflict. However, species of *Tillandsia* exhibited genetic cohesiveness, both in assignment to species and to clades, evident in the population genomics analysis section of **chapter 1** and in tests for monophyly in **chapter 2**. We found rather weak evidence for population structure in **chapter 1**, as is typical for species that experience high rates of ongoing gene-flow and species exhibiting ornithophilous pollination – however, our rather limited sampling restricted our ability to draw clear conclusions. Finally, **chapter 2** offered a re-telling of the evolutionary history of the clade. Whole-genome sequences and a phylogenetic network approach surprisingly provided evidence that *Tillandsia* subgenus *Tillandsia* may have expanded into Central America through several independent dispersal events.

Beyond questions aimed at the lineage’s history, this work provided findings on species delimitation and systematics. For example, we noted observations on the clade assignment of *T. zoquensis* and on the relationship between *T. mima* and *T. mima. var. chiletensis* in **chapter 2**. *Tillandsia* phylogeny and systematics will very likely remain a topic for lively scientific discussion, as these intriguing species are not easily delimited by morphological traits nor do they adhere to traditional species concepts. Numerous cryptic species complexes in this lineage await ‘entanglement’ and resolution, which is complicated by recurrent gene-flow.

Abundant data requires an abundance of technical and conceptual considerations. **Chapter 1** explicitly examined the performance of different bait sets for target capture in the context of studying recent, rapid radiations. Contrary to our early expectations, we found that a highly-curated bait-set did not necessarily offer crucial advantages. These results suggest that studies aiming at macro-evolutionary inference can evade the investment required in the design and development of home-made baits without risking satisfactory phylogenetic resolution. These are exciting findings in the context of collaborative and inclusive science: cost-effective and highly efficient genomic tools can provide comparable data-sets that can be readily integrated, encourage collaborative investigations and pose fewer financial barriers to research groups. Conversely, a specifically-designed bait set is beneficial for a number of investigative and technical aspects. As **chapter 1** details, we obtained more information on gene tree incongruence and had greater confidence in testing population genomic structure with Bromeliad1776. During analysis we observed technical advantages that were beyond the scope of the manuscript. Utilizing pipelines based on sequence assembly was difficult for the universal bait-set, especially when we attempted to include regions that were not directly targeted (i.e., ‘flanking’). In the universal set, those sequences exhibited high rates of mapping and assembly error; filtering the data-set proved complex and time-consuming, a property also reported by colleagues (Viruel, J., private correspondence). When using an alternative method of calling variants, we had less confidence in the universal data-set in terms of biologically-meaningful values. For these reasons, we ultimately estimated population genomic parameters using only the taxon-specific set.

The field of genomics advances with a quickly-changing selection of technologies and tools. Two types of data were utilized in this dissertation, each providing different advantages. While whole-genome re-sequencing is less laborious in terms of molecular work and offers a broad view of a species’ genome, it requires a reference genome and may provide data-sets infused with analytical ‘noise’ and challenging to filter. On the contrary, target capture provided the author with ‘simpler’ and ‘cleaner’ data-sets and with valuable knowledge of the regions targeted. A biased choice of loci may amplify error as discussed in the supporting information of **chapter 1**. The scientific questions at hand ultimately drive the choice of genomic tools, however it’s interesting to note that large genomic data-sets are not only costly to produce, but are also costly to store, analyze, document and maintain. To that

end, I believe that classical phylogenetic ‘small-data’ tools should not be underestimated for their reliability and elegant performance.

Tillandsia offers an intriguing model group for studying rapid radiations in general and Neotropical biota in particular. The observed quick accumulation of morphological traits, especially in a correlated manner, invites questions regarding the interplay of genomics and selection in the formation of adaptive syndromes. This dissertation provides early evidence for genomic processes resembling those reported in other fascinating and well-studied evolutionary radiations, like cichlid fish and *Heliconius* butterflies (Brawand et al., 2014; Malinsky et al., 2018; Martin et al., 2013). Key genomic resources are now available to study *Tillandsia*, including a reference genome and a bait set fit for studies spanning population genomics and genomic processes, from novel mutations to copy-number variation. The current work and existing tools will hopefully inspire other evolutionary biologists to do so.

Hybridization in *Tillandsia* had been observed and its role speculated by earlier researchers. In her publication from 1984, Gardner (1984) highlighted the prevalence of interspecific hybridization in *Tillandsia*, writing that “natural hybridization appear [sic] to be responsible for some portion of the phenetic variation expressed by these species”. **Chapter 2** provides evidence for the extent and timing of hybridization in *Tillandsia*, a process that is in no sense unique to this lineage. The role of hybridization in species radiations and the prevalence of speciation with gene-flow have both received growing attention in the last decades. Hybridization may increase adaptive potential by enhancing variation (Grant & Grant, 2019) and adaptive introgression was found to play a role in adaptation to climate (Leroy et al., 2020), disease (Bechsgaard et al., 2017) and photosynthetic shifts (Tefarikis et al., 2022), among others (Suarez-Gonzalez et al., 2018; Taylor & Larson, 2019). Furthermore, the observation of deviations from a bifurcating tree structure and the use of phylogenetic networks to describe species’ diversification are both becoming ubiquitous, if not necessary, to phylogenomic inference.

It is then clear that bifurcating trees do not reliably describe the evolutionary history of many lineages and that numerous assumptions on species behaving like discrete and isolated entities are insufficient to understand speciation (Mallet et al., 2016). These notions seem common within the scientific community (Appendix 1) and are inspiring new analytical models and tools, as well as new forms of thinking. I find interest in observing which modes of speciation and which hypotheses were emphasized by evolutionary biologists of the

modern synthesis, in comparison. Divergence was often described as a process driven by forces of ‘colonization’ and ‘competition’ and was completed in ‘isolation’. The vast evidence for the prevalence of hybridization in species’ radiations suggests a parallel narrative, one in which diversity and survival are driven by subtle ‘sharing’ of genomic resources and allelic ‘cooperation’ between closely-related lineages, rather than competition alone. As a new evolutionary synthesis materializes, so do the manifold stories about evolution. The author and botanist Robin Kimmerer wrote on the topic of sharing knowledge, “[language] ... is a prism through which we see the world.” (Kimmerer, 2013). As scientists, we design many such ‘prisms’ when we communicate ideas and findings. We also have the opportunity to be mindful of their meanings when we tell the story of how species appear, survive and prosper in nature.

References

- Bechsgaard, J., Jorgensen, T. H., & Schierup, M. H. (2017). Evidence for adaptive introgression of disease resistance genes among closely related *Arabidopsis* species. *G3: Genes, Genomes, Genetics*, 7(8), 2677–2683.
- Brawand, D., Wagner, C. E., Li, Y. I., Malinsky, M., Keller, I., Fan, S., Simakov, O., Ng, A. Y., Lim, Z. W., & Bezault, E. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature*, 513(7518), 375–381.
- Gardner, C. S. (1984). Natural hybridization in *Tillandsia* subgenus *Tillandsia*. *Selbyana*, 7(2/4), 380–393.
- Grant, P. R., & Grant, B. R. (2019). Hybridization increases population variation during adaptive radiation. *Proceedings of the National Academy of Sciences*, 116(46), 23216–23224. <https://doi.org/10.1073/pnas.1913534116>
- Kandziora, M., Sklenář, P., Kolář, F., & Schmickl, R. (2022). How to tackle phylogenetic discordance in recent and rapidly radiating groups? developing a workflow using *Loricaria* (Asteraceae) as an example. *Frontiers in Plant Science*, 12, 765719. <https://doi.org/10.3389/fpls.2021.765719>
- Kimmerer, R. (2013). *Braiding sweetgrass: Indigenous wisdom, scientific knowledge and the teachings of plants*. Milkweed editions.
- Leroy, T., Louvet, J.-M., Lalanne, C., Le Provost, G., Labadie, K., Aury, J.-M., Delzon, S., Plomion, C., & Kremer, A. (2020). Adaptive introgression as a driver of local adaptation to climate in European white oaks. *New Phytologist*, 226(4), 1171–1182.
- 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. *Nature Ecology & Evolution*, 2(12), 1940–1955. <https://doi.org/10.1038/s41559-018-0717-x>
- Mallet, J., Besansky, N., & Hahn, M. W. (2016). How reticulated are species? *BioEssays*, 38(2), 140–149. <https://doi.org/10.1002/bies.201500149>
- Martin, S. H., Dasmahapatra, K. K., Nadeau, N. J., Salazar, C., Walters, J. R., Simpson, F., Blaxter, M., Manica, A., Mallet, J., & Jiggins, C. D. (2013). Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Research*, 23(11), 1817–1828.

Parins-Fukuchi, C., Stull, G. W., & Smith, S. A. (2021). Phylogenomic conflict coincides with rapid morphological innovation. *Proceedings of the National Academy of Sciences*, 118(19), e2023058118. <https://doi.org/10.1073/pnas.2023058118>

Suarez-Gonzalez, A., Lexer, C., & Cronk, Q. C. B. (2018). Adaptive introgression: A plant perspective. *Biology Letters*, 14(3), 20170688. <https://doi.org/10.1098/rsbl.2017.0688>

Appendix 1



Twitter meme on the topic of phylogenomic inference (December 2021). The meme received 144 likes and 18 retweets, reaching a total of 11,765 people and indicating certain community interest in the concepts presented

Appendix 2 – conference contributions

Towards understanding the genomic substrate of diversification in *Tillandsia* subgenus *Tillandsia*

Gil Yardeni¹, Jaqueline Hess¹, Michael H. J. Barfuss¹, Margot Paris², Thorsten Krömer³, Michael Kessler⁴, Walter Till¹ & Christian Lexer¹

¹Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria;

²Department of Biology, University of Fribourg, Fribourg, Switzerland; ³Centro de Investigaciones Tropicales, Universidad Veracruzana, Xalapa, Veracruz, Mexico; ⁴Institute of Systematic and Evolutionary Botany, University of Zurich, Zurich, Switzerland

Corresponding author: gil.yardeni@univie.ac.at

Abstract

The species-rich and ecologically diverse Bromeliaceae family provides an excellent system for studying adaptive radiation in neotropical plants. The genus *Tillandsia* is the largest and most diverse within the family with over 600 species, extensive geographical range and a variety of ecological adaptations to different habitats such as epiphytism and various photosynthetic syndromes. Our research goal is to identify the genomic variability that provided the substrate for adaptive radiation within the enigmatic subgenus *Tillandsia* of genus *Tillandsia* using tree-, network, and coalescent-based approaches. To that end, we used whole-genome sequencing data from selected species within a taxonomic clade (Core Group I, also known as ‘The Mexican clade’) to construct phylogenomic trees and networks and infer species delimitation. To shed light on the genomic processes that gave rise to diversity within the clade, we performed tests for reticulation, introgression and post-speciation gene flow. We will present our first results, describe the genomic tools used for inference and outline future prospects towards understanding of genomic processes at micro- and macro-scales. This research project provides a unique case study for emerging research on the evolutionary genomic underpinnings of plant species radiations in the neotropics.

Keywords: *Tillandsia*, adaptive radiation, introgression, reticulation, phylogenomics.

***PopGroup53, The 53rd Population Genetics Group Meeting,
January 2020, Leicester, UK.***

***A target capture approach for exploring the drivers and constraints of adaptive radiation in
tillandsioid bromeliads (*Tillandsia* spp), a highly diverse neotropical plant group***

Gil Yardeni¹, Jaqueline Hess¹, Marylaure de La Harpe¹, Michael H. J. Barfuss¹, Juan Viruel², Margot Paris³, Clara Groot Crego¹, Walter Till¹, Thibault Leroy¹, Ovidiu Paun¹, Christian Lexer¹

¹Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria;
²Royal Botanic Gardens, Kew, Richmond, UK; ³Department of Biology, University of Fribourg, Fribourg, Switzerland

Abstract

The study of adaptive radiation calls for a synthesis between macro-evolutionary and population genetic approaches to uncover the genomic substrate and mechanisms of biological diversification at different time scales. The species-rich and ecologically diverse Bromeliaceae family of flowering plants provides an excellent system for studying the drivers and constraints of adaptive radiation. The genus *Tillandsia* is the largest and most diverse within the family with over 600 species, extensive geographical ranges and a variety of key adaptations such as epiphytism and divergent photosynthetic syndromes. Here, we present a roadmap to the development of a taxon-specific target sequence capture set, specifically created to address hypotheses on neutral processes (tracked via demographic modeling) and the dynamics of adaptive and deleterious variation during a textbook adaptive radiation. To assess the utility of the approach, we compare it to a ‘universal’ capture approach based on a widely used Angiosperm probe set (Angiosperm-353), examining the power of both sets in resolving the phylogenomic relationships within recent radiations and in common population genetics inference of admixture and introgression. We will discuss the benefits and limitations of the approach.

Custom-made or universal? The use of target capture to bridge macro- and microevolutionary processes in tillandsioid bromeliads (*Tillandsia* spp), a Neotropical rapid radiation

Gil Yardeni¹, Jacqueline Hess¹, Marylaure de La Harpe¹, Michael H. J. Barfuss¹, Juan Viruel², Margot Paris³, Clara Groot Crego¹, Walter Till¹, Christian Lexer¹, Thibault Leroy¹, Ovidiu Paun¹

¹Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria;

²Royal Botanic Gardens, Kew, Richmond, UK; ³Department of Biology, University of Fribourg, Fribourg, Switzerland

The study of rapid radiations calls for integrative approaches, bridging phylogenetics and population genetics to uncover the genomic substrate of diversification at different time scales. Target capture approaches emerged as an important tool to study evolutionary radiations in non-model taxa, enabling researchers to retrieve large data sets with few genomic resources. While developing taxon-specific target capture kits requires sustained effort, universal kits are readily available but may offer comparatively shallower insights, especially at short evolutionary timescales. The species-rich and ecologically diverse Bromeliaceae family provides an excellent system for studying the drivers and constraints of rapid, adaptive radiations. We focus on *Tillandsia*, a young subgenus (~6 Mya), yet the largest and most diverse within bromeliads with over 600 species, extensive geographical ranges and a variety of key adaptations such as epiphytism and divergent photosynthetic and pollination syndromes. We present a taxon-specific target sequence capture set for bromeliads, designed to address a wide range of evolutionary hypotheses and to assess the dynamics of adaptive and neutral variation by targeting 1,776 coding regions, including also genes putatively involved in several key traits. We compare our bait set to the ‘universal’ Angiosperms353 probe set, examining their power to resolve phylogenomic relationships using concatenation and species tree methods, estimate population genetic statistics and infer admixture with a focus on *Tillandsia* subgenus *Tillandsia*. The taxon-specific set results in high enrichment success across the entire family, yet both kits offer abundant information. The taxon-specific set outperforms the universal set with respect to gene tree concordance and inferred population structure, and provides reliable data for inference of admixture and population structure. Importantly, the performance of both kits is comparable for phylogenomic analysis, highlighting the vast potential of universal kits in research at different evolutionary scales. We further discuss methodological aspects and limitations of the approach.

Taxon-specific or universal? Using target capture to study the evolutionary history of a rapid radiation

G. Yardeni¹, J. Viruel², M. Paris³, J. Hess^{1,4}, C. Groot Crego¹, M. de La Harpe¹, N. Rivera¹, M. H. J. Barfuss¹, W. Till¹, V. Guzmán-Jacob⁵, T. Krömer⁶, C. Lexer¹, O. Paun¹, T. Leroy¹

¹University of Vienna, Austria, ²Royal Botanic Gardens Kew, UK, ³University of Fribourg, Switzerland, ⁴UFZ Helmholtz, Germany, ⁵University of Goettingen, Germany, ⁶Universidad Veracruzana, Mexico

INTRODUCTION

TARGET CAPTURE EMERGED AS AN IMPORTANT TOOL FOR INVESTIGATING PHYLOGENOMIC RELATIONSHIPS IN NON-MODEL GROUPS. As it allows to retrieve the same loci across wide taxonomic scales. Universal kits are designed for wide ranges of taxa and require reduced efforts and costs compared to taxon-specific kits, however, they may offer limited resolution power.

We designed a taxon-specific kit for Bromeliaceae (Bromeliad1776) and compared it to a universal kit for angiosperms (Angiosperms353), to gain insight into their utility in different evolutionary scales.

BROMELIACEAE

A LARGE, HIGHLY DIVERSE NEOTROPICAL MONOCOT FAMILY include the pineapple (*Ananas comosus*). Characterized by a leaf rosette which impounds rainwater in central tanks, where arthropods, microbes and other species reside. Presents a diversity of repeatedly evolving adaptive traits (=key adaptations¹), like CAM photosynthesis, epiphytism and a diversity of pollination syndromes.



METHODS

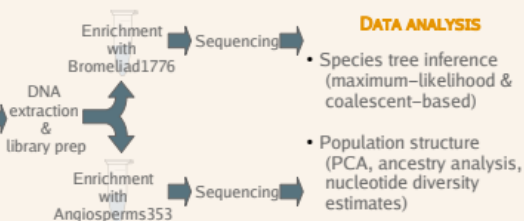
IN-HOUSE DESIGN: BROMELIAD1776 (TAXON-SPECIFIC)

Designed to address a range of evolutionary questions, based on *A. comosus* genome². Targets 1,776 coding regions, including random single-copy genes and genes associated with key innovative traits.

ANGIOSPERMS353³ (UNIVERSAL)

Designed using alignment from > 600 angiosperms, targets 353 single-copy protein-coding genes.

SAMPLING
72 bromeliad samples, pop-level sampling of 5 *Tillandsia* species



RESULTS

HIGHER MAPPING RATES TO TAXON-SPECIFIC KIT BOTH KITS PROVIDED LARGE NUMBERS OF SNPs

The taxon-specific kit provided x5 the number of SNPs and lower rates of missing data.

SIMILAR PERFORMANCE IN SPECIES TREE CONSTRUCTION

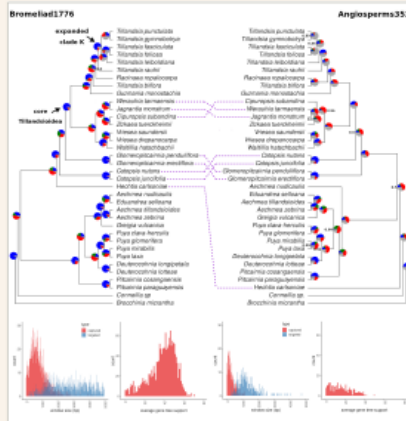
Both kits performed well in recovering the backbone phylogeny of Bromeliaceae with nearly identical topologies.

TAXON-SPECIFIC PERFORMS BETTER FOR INFERENCE OF GENE TREE INCONGRUENCE

In the universal kit, a considerable proportion of gene trees were uninformative. The length and average size of input gene trees was lower for the universal kit.

STRONG INTERSPECIFIC STRUCTURE BUT LITTLE EVIDENCE FOR POPULATION STRUCTURE IN *TILLANDSIA*

The taxon-specific kit provided slightly higher resolution.



Coalescent-based species trees generated with ASTRAL-III for Bromeliad1776 (left) and Angiosperms353 (right, flipped for mirroring) data-sets, respectively. 1600 and 200 genes were included for each set, respectively. Node values represent local posterior probabilities (pp) for the main topology and are equal to 1 unless noted otherwise. Pie charts show levels of gene tree discordance: percentages of concordant gene trees (blue), top alternative bipartition (green), other conflicting topologies (red) and uninformative gene trees (gray). At bottom, length and average bootstrap support for gene trees from each data set. Each gene was considered a single genomic window

DISCUSSION

A TAXON-SPECIFIC SET PERFORMS marginally BETTER FOR PHYLOGENOMICS & POPULATION GENETICS

- Higher number of segregating sites does not necessarily translate to higher resolution power.
- Phylogenomic relationships were resolved even in rapidly radiating clades. Taxon-specific kit provided deeper resolution for examination of gene tree incongruence.
- The taxon-specific kit performed marginally better at recent evolutionary scales.
- Differences in resolution ascribed to length and variability of obtained regions.

BROMELIAD1776 OFFERS FUTURE PROSPECTS FOR BROMELIAD RESEARCH

With highly coverage across taxa and knowledge of targeted regions, the kit offers a broad utility for research of molecular evolution in Bromeliaceae.

WANT TO KNOW MORE?

READ MORE IN THE PREPRINT:

BIORxiv

<https://doi.org/10.1101/2021.05.20.444989>

10.1101/2021.05.20.444989



Taxon-specific or universal? Using target capture to study the evolutionary history of a rapid radiation

Gil Yardeni^{1,*}, Juan Viruel², Margot Paris³, Jaqueline Hess^{1,4}, Clara Groot Crego^{1,5}, Marylaure de La Harpe¹, Norma Rivera¹, Michael H. J. Barfuss¹, Walter Till¹, Valeria Guzmán-Jacob⁶, Thorsten Krömer⁷, Christian Lexer¹, Ovidiu Paun¹, Thibault Leroy¹

¹ Department of Botany and Biodiversity Research, University of Vienna, Vienna, Austria

² Royal Botanic Gardens, Kew, Richmond, UK

³ Unit of Ecology & Evolution, Department of Biology, University of Fribourg, Fribourg, Switzerland

⁴ Department of Soil Ecology, Helmholtz Centre for Environmental Research, UFZ, Halle (Saale), Germany

⁵ Vienna Graduate School of Population Genetics, Vienna, Austria

⁶ Biodiversity, Macroecology and Biogeography, University of Goettingen, Göttingen, Germany

⁷ Centro de Investigaciones Tropicales, Universidad Veracruzana, Xalapa, Mexico

*Corresponding author: gil.c.yardeni@gmail.com

Target capture has emerged as an important tool for phylogenetics and population genetics in nonmodel taxa. We present a newly developed target capture set for Bromeliaceae, a large and ecologically diverse plant family with highly variable diversification rates. The set targets 1776 coding regions, including genes putatively involved in key innovations, with the aim to empower testing of a wide range of evolutionary hypotheses. We compare the relative power of this taxon-specific set to a universal kit designed for all Angiosperms, the Angiosperms353 set, focusing on the rapid radiation of *Tillandsia* subgenus *Tillandsia*. We highlight differences in kit performance for constructing phylogenies, exploring gene tree discordance and investigating nucleotide diversity and population structure. While the overall performance of both kits is comparable, the taxon-specific kit offers benefits in more detailed phylogenetic or population genetic analyses. The comparative study provides potential lessons for future investigations, in particular for rapidly radiating clades.

**Rapid radiation characterized by prevalent hybridization in the Neotropical genus
*Tillandsia***

Yardeni, G.¹, Barfuss, M. J. H.¹, Till, W.¹, Lexer, C.¹, Leroy, T.¹, Paun, O.¹

The recent rapid radiation of *Tillandsia* subgenus *Tillandsia* (family Bromeliaceae) is a highly diverse Neotropical Monocot group which provides an attractive system for studying the drivers and limitations of species diversification. This subgenus, which consists of predominantly epiphytic plants, is thought to have diversified within the last 3 million years during its expansion from South into Central America, following the closing of the Isthmus of Panama. We inferred phylogenomic relationships among 32 species of *Tillandsia* using whole-genome data and employed a tree-based approach to explore the evolutionary history of the clade. We report a lack of monophyly and deviations from a tree-like structure coupled with rampant gene tree discordance. High rates of hybridization within and between clades suggest that the expansion of the subgenus into Central America proceeded in several migration events, followed by episodes of diversification and gene flow. Focusing on hybridization, Patterson's D (ABBA-BABA) and related statistics were used to describe the rates and timing of introgression events and to assign introgression events to tree branches rather than species. Finally, we discuss the possible contribution of intraspecific gene flow to adaptive trait shifts. Our work provides an example of how hybridization and introgression, rather than problems to overcome in phylogenomic inference, can be investigated to deepen our understanding of complex processes that accompany species radiation.