NATURAL HISTORY MUSEUM OF DENMARK

UNIVERSITY OF COPENHAGEN

Sam Bruun-Lund

PhD Thesis

Darwin's abominable mystery

– and the evolution and diversification of *Ficus* L.

Advisor: Prof. & curator Nina Rønsted

Submitted: March 2019

Darwin's abominable mystery

– and evolution and diversification of *Ficus* L.

PhD Thesis

Sam Bruun-Lund

Advisor
Professor & curator Nina Rønsted

Rønsted group
Natural History Museum of Denmark
University of Copenhagen

Submitted: March 2019

Institution	University of Copenhagen
Department	Natural History Museum of Denmark
Author	Sam Bruun-Lund
Title	Darwin's abominable mystery – and the evolution and diversification of *Ficus* L.
Advisor	Professor & curator Nina Rønsted *Natural History museum of Denmark*
Co-advisors	Associate professor Wendy L. Clement, *The College of New Jersey* Professor Finn Kjellberg, *CEFE, CNRS, University of Montpellier* Professor Tom Gilbert, *University of Copenhagen, Denmark*
Submitted on	14/03-2019
Assessment committee	Associate Professor Thomas Pape, *Natural History Museum of Denmark* Associate Professor Anneleen Kool, *Natural History Museum of Oslo* Professor Mark W. Chase, *Royal Botanic Gardens Kew*
Funding	Independent Research Fund Denmark (#7014-00205) & Carlsberg Foundation (#CF16-0883)

Cover photo credits:

1. *Ficus macrophylla* Desf. ex Pers. https://en.wikipedia.org/wiki/Ficus_macrophylla#/media/File:Starr_000501-1307_Ficus_macrophylla.jpg
2. *Ficus* sp. https://pinterest.dk/pin/492722015471450004/
3. *Ficus mollis* Vahl. https://flowersofindia.net/catalog/slides/Soft%20Fig.html
4. *Ficus sumatrana Miq.* http://hungrytrees.com/post/24702095337/national-geographic-april-1997-ficus-sumatrana
5. *Ficus lilleputiana* Dixon http://thefigsofaustralia.com/ficus-lilliputiana-your-new-favourite-fig/
6. *Ficus sp.* https://pinterest.dk/pin/61220876165476861
7. *Ficus destruens* F.Muell. ex C.T.White http://flickriver.com/photos/blackdiamondimages/tags/stranglerfig/
8. *Ficus dammaropsis* Diels https://pinterest.com/pin/112801165643459023/

Backside photo credits:

1. *Ficus sp.* https://mpics.mgronline.com/pics/Images/555000015958012.JPEG
2. *Ficus watkinsiana* F.M.Bailey *https://flickr.com/photos/17674930@N07/39039463314*
3. *Ficus villosa* Blume https:/palomar.edu/users/warmstrong/images2/ficvill1b.jpg
4. *Ficus sp.* with Long-tailed macaque (*Macaca fasicularis*) https://gettyimages.dk/detail/photo/long-tailed-macaque-in-a-strangler-fig-tree-ladened-royalty-free-image/83971378
5. *Ficus religiosa* L. https://en.wikipedia.org/wiki/Ficus_religiosa#/media/File:Ficus_religiosa_Bo.jpg
6. *Ficus septica* Burm.f. and *Ceratosolen jucundus* Grandi – by Finn Kjellberg
7. *Ficus* sp. with parasitic fig wasp (*Apocryta westwoodi* Grandi) https://sciworthy.com/parasitic-fig-wasps-penetrate-figs-with-zinc-hardened-drill-bit-tips-to-lay-eggs/
8. *Ficus sp.* with fig disperser: Rhinoceros Hornbill: https://www.allposters.com/-sp/Young-Rhinoceros-Hornbill-Feeds-on-a-Fig-from-a-Strangler-Fig-Tree-in-Borneo-Indonesia-Posters_i4076206_.htm

Chapter I photo credits:

By George Weiblen ©

Chapter II photo credits:

Modified figure 6 from Machado et al. *2018* (Mol. Phyl. Evol.) ©

Chapter III photo credits:

Ficus sumatrana Miq. http://hungrytrees.com/post/24702095337/national-geographic-april-1997-ficus-sumatrana

Chapter IV photo credits:

http://flower-tales.blogspot.com/2011/11/sacred-bodhi-tree.html

Chapter V photo credits:

Ficus lilliputiana Dixon http://www.thefigsofaustralia.com/wp-content/uploads/2015/07/Ficus-lilliputaina-1.jpg

Plants holds secrets…

By Sam Bruun-Lund

"I know of no pleasure deeper than that which comes from contemplating the natural world and trying to understand it."
— David Attenborough

Preface

The work accumulated in this thesis is the result of my PhD research, supervised and reinforced by professor and curator Nina Rønsted, at the Natural History Museum of Denmark (NHM), University of Copenhagen. The thesis is comprised of an introduction followed by five main chapters that forms the core of the PhD research conducted. Following the main chapters, a series of appendices is included, presenting some of the additional output produced during my PhD. The appendices are not the main body of work, but represent an opportunity to use my knowledge and expertise on other groups of plants too through collaborations and co-supervision of students. Other studies during my PhD years that have not been included in this thesis, but will eventually lead to additional publications, are: 'Origin of *Aloe vera* L. using plastid genomic data' (Grace et al., in prep.), and Sanger sequencing-based studies on 'Phylogeny and biogeography of snowdrops (*Galanthus* L.)' (Bruun-Lund et al., in prep.), 'Phylogeny of South African geophytes (Amaryllidaceae)' (David et al., in prep.), and 'Resolving species concepts in Andean *Elaeagia* Wedd. (Rubiaceae) (Maldonado et al., in prep.)'. These studies are not included, as they are in too early of a stage.

I have had a curiosity about plants and nature ever since I can remember – which has led to taking over my parent's garden and making the lawn into many different biomes/flowerbeds. I also made my room into a forest – starting with succulent plants, then moving into carnivorous plants and later a mix of odd and peculiar plants – at one point in time, you could

count as many as 100 different plants in my room! This interest naturally led to a bachelor and master's degree in biology. Most master-courses were taken at NABiS (Nordic academy of Biodiversity and Systematic Studies) a joint network of Universities in Scandinavia focusing on systematics and related tools.

In 2012 Cruaud, Rønsted and co-authors published a monumental paper on figs and their pollinators from a molecular perspective: *An extreme case of plant-insect codiversification: figs and fig-pollinating wasps* (Cruaud, Rønsted et al. Syst. Biol. 2012). However, the origin and evolution of the fig-wasp mutualism was still not resolved with satisfaction despite the huge effort. This is where I first met the figs! I've never thought that I would come to have such an enormous love of figs or trees. Trees were always something 'a bit boring' to me, compared to other plants such as carnivorous or medicinal plants. Yet, when I started to learn about the figs, the mutualism with their wasp pollinators, and the huge diversity with more than 800 species as well as almost any trait, I could not stop thinking about how cool and interesting they actually were. Besides many unanswered questions about the biology and evolution of figs, the study in Systematic Biology, left out many exciting fig species. Likewise, lack of resolution and sufficient support in the analysis lead to equivocal inferences of the origin and evolution of figs. So, several questions and much more work was left to do on figs! Luckily for me. This was the start of an amazing journey into the exciting world of figs and the many, many layers of interactions and adaptations that have led to their global success.

The amazing thing about figs, is that they are usable for so much. In addition to all the exciting research described in this thesis and the work of many others, figs are also an ideal group of plants for communication and engagement, because most people will actually know at least one fig species (and probably have several as ornamental houseplants at home too). What most do not know, is how much more diversity exists – and how extreme their biology is. Consequently, I have also used my study group for so much more than research, which was made possible by being

based at a Natural History Museum that engage with the public on a regular basis.

Finding a place as NHM and Nina Rønsted's group where I could share the interest and enthusiasm for plants and nature has been absolutely amazing. I find it extremely important to share research with the general public and engage people in natural science to make it less mysterious and more trustworthy, exiting and worthwhile spending time and funding on. It is especially important for the younger generation to see that scientists are not only dusty photos of Charles Darwin or Albert Einstein. Scientists can also be you and me, or anybody else. Luckily, my mission aligns perfectly with the mission for NHM: *"To empower citizens to connect with nature"*.

I have therefore engaged as much as possible in a diversity of communication activities as well as in working groups focusing on creating new visions and strategic plans for visitor experiences in the Botanical Garden collections and the new Natural History Museum to be built in the coming years.

One of the best places for communication activities has been the Botanical Garden and palm houses where students, and the public, can engage with the figs on their own, or, together with me. I have developed content for guided tours in the Botanical Garden, where botanical research at NHM was in focus and importantly, conveyed research in a way that the general public can relate to, instead of focusing on the otherwise difficult and technical aspects. Furthermore, seeing the great potential in the Botanical Garden, and knowing it as well as the back of my hand, I have participated in creating and building up a stronger social media presence at the NHM and Botanical Garden's profiles with exciting and shocking stories from the plant- and research world (@botaniskhavekbh @statensnaturhistoriskemuseum and associated facebook pages).

I have likewise participated in activities outside the museum and garden, such as 'Naturvidenskabsfestivallen' several times, where high-

schools invite scientists to their schools to give talks about their research, methods, fieldwork and findings. I for example presented **chapter IV** with emphasis on murderous strangler figs, and how humanlike this lifeform can be, in an attempt to captivate students' interest in evolution and botany early on in the education system. I also used the strangler figs in video communication for social media (https://vimeo.com/248171446).

I have also been able to contribute to the museum's scientific collections on several occasions. I had the pleasure of conducting fieldwork in Australia in the Summer of 2017 together with Nina. Here we sampled *Ficus* species for **chapter V** in the several habitats and biomes, such as rainforests and the arid outcrops. Thus, having to travel great distances to find the many specimens. We followed the tracks that many great fig scientists have walked, using their references and GPS coordinates. It was a privilege to go back and re-find *Ficus* species that have not been sampled, since they were described many years ago!

In 2017, we acquired 20 selected *Ficus* species from The Arboretum and Botanical Garden, Milde (University of Bergen) from the personal collection of the great fig taxonomist Cornelius C. Berg, who passed away in 2012. The adoption of this collection was made into a press release, several social media posts and signs in the Botanical Garden. The gardeners and I worked together to organize the fig species according to their phylogenetic clades and geographical distribution when possible. See **appendix IV** for press release.

Another great part of the environment at NHM and the University of Copenhagen is teaching responsibilities. The place to encourage the next generation of scientists and producers of knowledge. We need to pass on, how enjoyable it can be in academia. If only I could count how many times I was told that "there's no possible way to get a carrier or take courses in botany in Copenhagen any more" – until I meet Nina and her can-do attitude – and look where I am, ready to defend my PhD thesis in botany! Teachers that can encourage you to keep going and study in a specific

area, are highly important. Therefore, I have found so much satisfaction in teaching university courses and help develop a curriculum that can make students see, how many approaches there can be to botany – and that it is exciting and possible to do botany! The joy of recognizing a plant in nature is big, but then also knowing that it is deadly, or perhaps have been used for treatment of illnesses though centuries, makes you remember that species even more. I think, botany becomes a more attractive subject to study and build a career in, by linking the basic curriculum with all the exciting botanical stories and applications. Courses where I have been involved in teaching include *forensic botany (Msc)*, *plant ecophysiology (Bsc)*, and *plant-animal interactions (Msc)*. I have also participated in the training of the Danish delegation for The International Biology Olympiad in botany a couple of years. Teaching taxonomy, plant ID, phylogenetics, plant anatomy and physiology is even more rewarding as these subjects are often not well-covered in the traditional high-school curriculum.

Near the end of my PhD I was invited to present my latest research results at *The Science Gala*, the annual culmination of the museums popular *Wine and Science* talks in the historical banquet hall of the University of Copenhagen. This had been a goal of mine for a long time, and finally being able to do this was a great privilege and experience; giving the newly produced knowledge back to the citizens.

Sam Bruun-Lund, March 2019

"Standing on the shoulders of Giants"
— **Isaac Newton**, 1675

Acknowledgements

If I was asked the question: "would you do it all over again?" I would without hesitation say yes! I have found so much joy in the work and development of new knowledge in the field of evolution and botany. I have enjoyed all the different aspects, from developing the ideas, to the challenges with new methodologies and bioinformatics analysis, and interpretation and communication of the results, all being so much more enjoyable because I could work on figs. Likewise, I have developed a lot through the years. I've learned to use my strengths and personality, even ones that I did not see as strengths at first. This is all due to the fact that I found Nina at an early state in my scientific career, and at that time, her small group at the NHM. Nina you have been my advisor, officially, but you have been so much more that that! It became very apparent from the first few emails we shared – the energy and lack of patience in both of us, and that we both could use this energy to achieve things was the start of something amazing. I will hope we can continue our collaborative work and friendship for many more years. Thank you for the continuous support, help, guidance, mentorship, opportunities etc. etc. Thank you so much for everything – so far! To my co-supervisors, Wendy Clement, Finn Kjellberg and Tom Gilbert, thanks for being there with specialist advice and general interest and support in my work.

I want to thank the past and present members of the Rønsted group. Especially I would like to thank the past members from when I joined the group, for welcoming me with open arms, and the great times spent together working and laughing – the first insight into a research environment is so important! The present members for the great support

and collaborations. In particular I would like to thank Haris Saslis-Lagoudakis, Mette Wenøe, Olwen M. Grace & Brecht Verstraete. Also big thanks to the present members: Natalie Iwanycki Ahlstrand, Louise Isager Ahl (#instaoffice), Stephen Garrett, Chris Barnes, Nataly Allasi Canales and Kia Dahl Sørensen. Equally, a big thanks to Charlotte Hansen for being one of my first tutors of the basic lab-work – I know I can always come to you, if I need advice in the lab. Thank you all! I also want to thank the EvoGenomics sections past and present people and especially the inspiration from Tom Gilbert and his genomic work and the many people I have got some sort of input and help from: Sama, Fatima, Stine, Liz as well as Mike. Mike is now at NTNU where I spent an exciting research visit learning from him and Vanessa, thanks. The gardeners of the Botanical Garden should also have a big thanks for inspirational chats, and helping me to retain the "real" living plants and the joys that follows being around them. Likewise, the people at the rest of NHM, especially in the 'public outreach department' for letting me be part of your important and fun work. Thank you! And finally, but not least, all the people who provided me with plant material or helped me with fieldwork as detailed in the papers.

In the world outside of DNA, botany, biology and science I also have a lot of people I want to say a big thank to. My martial art community. Being part of this, have made me the person I am today too. I have learned that no obstacle is too big and a 'yes I can' attitude will get you a long way on its own. Thank you! I want to thank two people that I do not have contact to any more, unfortunately. My biology teacher in high-school (gymnasium) Annli Jærger Rasmussen, who saw the 'talent' in me for the natural sciences, similarly, the spark of joy biology gave me, and guided me to pursue my education in biology at KU. Also, my Grandmother, Annlise Lund who I believe has been one of the greatest inspirations for my general interest in botany and plants. The number of hours I have spent in the Botanical Garden, various parks and in her own garden are enormous and irreplaceable. Thank you!

Thank you to my second family, my friends, who have managed to follow me on the side-line for all these years: Magnus, Zandra, Marianne, Alexander, Anne Sofie, Mette and Silvia. Thank you, Marius, for everything along the last, heavy part of the PhD – and for helping me remined myself that sometimes I 'just got to sit still and enjoy the moment'. Also thank you for being so extremely understanding of my craziness and enduring ups and downs in this last few months.

To my parents Ole and Karin Bruun-Lund who have shaped me from birth and given me the best possible conditions for succeeding in life at *all* levels. The amount of support, love and help from you two is without a doubt one of the only reasons I have been able to do what I have done in my life, so far. Having you whenever I needed you and could not cope with work, life or so, has been absolutely irreplaceable. Thank you!!!

Finally, a warm thank you to the committee members taking their time to read through my work and for the funding bodies: Independent Research Fund Denmark and Carlsberg Foundation.

English summary

Figs (*Ficus* spp. L.) constitute one of the largest angiosperm lineages with ~800 species distributed in tropical and sub-tropical regions all over the world. Figs may be mostly known for their obligate mutualism with their pollinating fig-wasps (Agaonidae, Chalcidoidea, Hymenoptera). The two linages have codiversified for the last 75 million years. Additionally, figs present a diversity of habits and habitat usage and are considered keystone species in rainforests due to abundant fruit set. Our understanding of the evolutionary history of figs is based on the milestone monographs of Corner and Berg, primarily based on morphology. Meanwhile, the advent of DNA sequencing has unravelled conflicting relationships and questioned all our understanding of the evolution and diversification of figs. However, decades of Sanger sequencing have not provided a comprehensive phylogenetic hypothesis needed to re-test standing hypotheses about the origin and specificity of the mutualism, nor why the figs have been successful in adapting to many environments. In this PhD project, the figs have entered the era of high-throughput sequencing (HTS) plus undergone investigations to why there are so many fig species and the diversification patterns facilitating their diversity.

In **chapter I**, an updated classification is presented with the current largest dataset available with more than 300 species of figs. This proposed new working classification of major lineages and groups is reflecting their evolutionary relationships. However, many relationships between clades are still hindered by lack of resolution in the DNA data. In **chapter II**, diversification and biogeography in Neotropical figs was investigated. Most diversity is recent, happening within the last 16 million years and diversification during unstable climatic periods seem to have shaped the current diversity. The hemi-epiphytic habit, together with small propagules, seem to lower extinction rates in Neotropical figs. In **chapter III**, the diversification dynamics of the entire genus was explored. The

success and diversity of figs appear to be the product of slow, steady diversification rates with low extinction. No significant bursts of diversification were detected. We hypothesize the success of the genus to be affiliated with hemi-epiphytism, allowing occupation of new niches – and being monoecious with active pollination ensuring reproductive success. In **chapter IV**, figs entered the era of HTS and the first plastid genome of *Ficus religiosa* L. was assembled and annotated. Near complete plastid genomes (plastomes) of 65 taxa were used to gain a first ever insight into the evolutionary history of the plastome in fig phylogenetics. We found that cyto-nuclear discordance was present in the genus, possibly as a result of host-shifts and ancient plastome introgression. Thus, **chapter V** utilizes HTS and targeted sequence capture to focus on nuclear loci for reconstruction the phylogenetic hypothesis. A set of probes that can be used in the entire genus was developed and assessed; our probes can confidently resolve relationships of major lineages. The probes were used on a preliminary taxon set of an Australasian clade of figs and show efficiency to resolve closely related species. The final dataset will be used to test for biomes shifts and linked traits, between wet and dry regions.

Dansk resumé

Figner (*Ficus* spp. L.) er en af de største slægter af dækfrøede planter med ~800 arter fordelt over hele verden i tropiske og subtropiske områder. Figner er mest kendt for deres obligate mutualisme med deres bestøvende figen-hvepse (Agaonidae, Chalcidoidea, Hymenoptera). De to udviklingslinjer har co-diversificeret igennem de sidste 75 millioner år. Dertil har fignerne en stor morfologisk forskellighed, både i vækstform og habitat brug. Derudover er de en *keystone* art i regnskove, da de sætter frugt året rundt. Vores forståelse af fignernes evolutionshistorie er baseret på en milepæl af monografer fra Corner og Berg, primært baseret på morfologi. Nyere DNA-sekventering har afsløret systematiske konflikter og sat spørgsmålstegn ved hele vores forståelse af fignernes evolutionen og diversifikationen. Imidlertid har et årti af Sanger sekventering ikke kunne bidrage med en velunderbygget fylogenetisk hypotese, som er nødvendig for at teste nuværende hypoteser omkring oprindelse og specificitet af mutualismen – samt hvorfor figner har været så succesfulde i deres tilpasning til forskellige miljøer. I denne Ph.d.-afhandling bringes fignerne ind i en ny ære af high-throughput sekventering (HTS) samtidig med at hypoteser for fignernes succes og diversifikationsmønstre undersøges.

I **kapitel I** præsenteres en opdateret klassifikation baseret på det til dato største datasæt med mere end 300 figenarter, således at klassifikationen af de største udviklingslinjer og grupper af figner reflekterer deres evolutionære historie. Dog er mange af relationerne mellem klader stadigvæk udfordret af manglende information i forhåndenværende DNA-data. I **kapitel II** undersøges diversifikation og biogeografi for neotropiske figner. Det meste af diversiteten er udviklet indenfor de sidste 16 millioner år, derudover ser det ud til at diversifikationen af figner i klimatisk ustabile perioder også har været med til at forme den nuværende mangfoldighed. Den hemi-epifytiske form, sammen med små spredningslegemer, ser ud til at have sænket raten for at uddø i neotropiske figner. I **kapitel III** undersøges diversifikations-dynamikker for hele figen-slægten. Succes og diversiteten af figner ser ud til at skyldes en langsom og konstant diversifikations-rate, med meget lidt uddøen. Vores hypotese for succesen er dels tilknyttet hemi-epifytisme, som tillader etablering af nye nicher – og delsværende enbo og med aktivt bestøvende hvepse, som sikrer

reproduktion. I **kapitel IV** ledes fignerne ind i HTS-æraen og det første plastide genom (plastom) af *Ficus religiosa* L. blev samlet og annoteret. Næsten komplette plastomer for 65 taxa blev brugt til at skabe det allerførste første indblik i evolutionshistorien for plastomer i arbejdet med fignernes fylogeni. Cyto-nukleare uoverensstemmelser i figenslægten, kan muligvis skyldes vært-skifte og forhistorisk plastom introgression. I **kapitel V** anvendes HTS-metoden *targeted sequence capture* til at fokusere på kerne-loci til rekonstruktionen af en fylogenetisk hypotese. Et sæt af prober der kan anvendes på hele slægten blev udviklet og testet og kan med statistisk sikkerhed afklare forholdene mellem de største udviklingslinjer hos figner. Proberne blev anvendt på en Australasian klade af figner som et casestudie, og viste effektivitet at kunne afklare forholdet mellem nærtbeslægtede arter. Det endelige datasæt vil blive brugt til at teste for biom-skift mellem våde og tørre miljøer, samt associerede vækstformer og morfologiske karakterer.

Content

Introduction

Plants holds secrets…

Introduction

Plant diversification in the tree of life

Take a walk during summer in the Botanical Garden and Palm Houses at the Natural History Museum of Denmark – you will find yourself surrounded by flowering plants in full bloom. This situation is not just found in a Botanical Garden even though the overall species diversity might be higher, most natural terrestrial ecosystems will be like the Botanical Garden, dominated by flowering plants. The flowering plants, or angiosperms, are an extremely diverse group consisting of ~304.000 named species. An estimated additional 156.000 still remain unnamed (Pimm and Raven, 2017) and all these species originated from 146-100 million years ago (Mya) up till now, more recently than any other clade of vascular plants (Davies et al., 2004; Silvestro et al., 2015).

Flowering plant diversity is astonishing (**Figure 1**). From the size of tiny plants such as the duckweed, *Wolffia globose* (Roxb.) Hartog & Plas – to the largest trees up to 100 meters high as the giant sequoias, *Sequoiadendron giganteum* (Lindl.) J.Buchh., and the world's largest flowers on *Rafflesia arnoldii* R.Br.

Figure 1 | Flowering plant diversity. *Wolffia globose* (top left), *Sequoiadendron giganteum* (top right), *Rafflesia arnoldii* (bottom). See page 50 for photo credits.

When you enter to the Victorian inspired Palm House of the Botanical Gardens you will encounter a magnificent collection of figs (genus *Ficus* L., Moraceae). You will face a small peak view of the heterogeneity of habit, morphology and habitats the genus *Ficus* presents. The genus includes habits as shrubs, root-

climbers, hemi-epiphytes, rheophytes, lithophytes and banyan trees – a wide array of fig shape and color, and morphology. See **Figure 2** and cover of this thesis (Berg and Corner, 2005).

With at least 800 named species in tropical and subtropical areas across the globe, *Ficus* account for more than half of the species diversity of the mulberry family, Moraceae (ca. 1,100 species; Clement and Weiblen, 2009). This diversity is remarkably high even comparing *Ficus* to the rest of the angiosperms where as few as 57 genera together contain more than 500 species each and these genera account for most of the diversity of angiosperms (Frodin, 2004). *Ficus* are not only diverse, but also play an important part in the ecosystems providing edible figs throughout the year making *Ficus* key-stone species in rainforests (Harrison et al., 2012)

The unifying character for members of the genus *Ficus* is the distinctive inverted inflorescence (syconium – **Figure 3**), which is the site of an obligate mutualism with pollinating fig wasps (**Figure 3**) of the Hymenopteran family Agaonidae (Cook and Rasplus, 2003). Generally, each species of fig is only pollinated by its own species of pollinating wasp, although reports of multiple pollinators are common. Figs are pollinated only by female wasps, which lay their eggs exclusively through the style of the ovary where wasp larvae feed on some of the seed's endosperm (Kjellberg et al., 2001; Jousselin et al., 2004).

Figure 2 | **Diversity of the genus *Ficus*.** A wide diversity in habit, roots, leaves, colors and syconium are seen in the many species of *Ficus*. Species: *Ficus drupacea* Thunberg, *Ficus pleurocarpa* F. Muell., *Ficus watkinsiana* F.M. Bailey, *Ficus sycomorus* L., *Ficus benghalensis* L., *Ficus minahassae* Miq., *Ficus dammaropsis* Diels, *Ficus punctata* Thunb. See page 50 for photo credits.

The extreme mutualism has gained much attention by scientists through time because of its high specificity and many trophic layers of interactions (Cook and Rasplus, 2003). In addition to the pollinating wasps, different lineages of parasitic wasps raise their own offspring either on *Ficus* endosperm or other wasps (Kjellberg and Proffit, 2016).

Figure 3 | **The syconium**. The inverted, urn-shaped inflorescence, which is a unifying character for *Ficus*. Here *Ficus ottonifolia* Miq. syconium from the Botanical Garden at NHM Denmark.

Why so species rich and ecologically successful?

Like the devotion the origin and evolution of the mutualism between *Ficus* and their pollinators have received, the origin and subsequent diversification of angiosperms has received high attention and puzzled evolutionary biologists as an "unsolvable mystery". In Charles Darwin's letters to J.D. Hooker in 1879 he famously referred to the origin of angiosperms as "an abominable mystery" (Darwin, 1872; Davies et al., 2004; Crepet and Niklas, 2009). "Darwin's abominable mystery" generally highlight a problem of origin and timing – from what and where

did the angiosperms evolve? However, Darwin was also perplexed about the subsequent diversification of angiosperms, which lead to his famous sentence: "why are the angiosperms so species rich and ecologically successful?", sometimes referred to as Darwin's "second abominable mystery" (Darwin 1871, 1903; Crepet and Niklas, 2009).

In the fossil record there is a point in time in Early to mid-Cretaceous about 125-80 Mya where the angiosperms abruptly began to appear in vast numbers (e.g. Wang et al., 2016). Molecular phylogenetics and dating have shown a similar history and there is evidence of early rapid radiations and many lineages date back to early cretaceous (146-100 Mya). Thus, all major lineages go back to a similar point in time (Soltis et al., 2008; Vamosi et al., 2018; Coiro et al., 2019). Bursts of diversification for other organisms correspond to the rise of angiosperm-dominated forests where several organismal clades also diversified: ferns, ants, herbivores beetles, amphibians, primates and phytophagous insects (Mitter et al., 1988; Wilf et al., 2000; Schneider et al., 2004; Moreau et al., 2006; Wilkinson et al., 2007; Gómez and Verdú, 2012; Ikeda et al., 2012).

Through time, many studies have tried to elucidate the origin of the sudden bursts of diversification focusing on many different aspects such as genomics, phylogenetic inference and paleobotany (Soltis et al., 2008; Chamala et al., 2013). However, what Darwin did not know was that angiosperms represent not just one but multiple deep-level radiations (Soltis et al., 2008). A major unresolved question in angiosperm evolution is what the underlying causes of the subsequent diversification of flowering plants are?

During the cretaceous (145-66 Mya) the Earth did not look like the Earth we know today. The climate was much warmer with a high humidity and the level of CO_2 was much higher as a result of volcanic activity. Accordingly, flowering plants originated and evolved in a much different world than what we see today. Additionally, the break-up of landmasses led to habitat fragmentation and divergence. The most severe change on Earth happened around the Cretaceous-Paleogene or K-P boundary some 66 million years ago when a massive asteroid impact happened affecting the environment. Light was suddenly limited and earth became cooler. Many flowering plants and animals went extinct. Particular linages survived and continued to evolve during the Cenozoic (from 66 Mya), including the rosids which now account for more than one fourth of all angiosperms (Wang et al., 2009). After the major angiosperm lineages had originated and started to diversify, the crown group of *Ficus* also originated around 75 Mya, with a possibility to be even older (Cruaud et al., 2012; Zhang et al., 2018).

Many justifications have been given to explain the diversification ability of the angiosperms, such as the rise of new niches due to the changing of landscapes and environments, plus, the possibility for interactions with simultaneously diversifying insects (Soltis et al., 2008; Wang et al., 2009). Could the presence of all these opportunities have contributed to the diversity we see today?

Rainforest diversification as clues

When studying diversity, it is a great start to look for clues in diversity patterns of rainforests as they are containing most of the biodiversity in terrestrial ecosystems on the planet (Gentry, 1992). The warmer climate in past time show similarities to our present day rainforests (Couvreur et al., 2011a). However, rainforest diversity is not well understood and it has variously been attributed to be a 'museum of diversity', showing constant diversification rates along with low extinction rates leading to a gradual accumulation of lineages in response to a long-lasting and stable tropical ecosystem (Stebbins, 1974; Wallace, 1878), or a 'cradle of diversity', referring to an increase in diversification rates towards the present with rapid radiations in response to climatic, tectonic or biotic changes (Richardson et al., 2001; Pennington et al., 2015). In line with Couvreur et al. (2011b) who suggested a mixed model of steady processes and mixed diversification, Koenen et al. (2015) proposed a concept of highly dynamic diversification processes across ecosystems that are linked to environmental changes (Xing et al., 2014). Explaining the diverse assemblage of species that inhabit the rainforest remains one of the challenges to understand how and which factors contributes to the diversity we see in the present day of angiosperms. To increase our knowledge of speciation processes, potential key innovations, and biogeographical patterns, the use of molecular phylogenetic hypotheses of species-rich plant genera can aid in the understanding of these patterns (**Chapters II-III**).

Ficus L. - evolutionary history and puzzles

Over the past 20 years, there have been important advancements in our understanding of the evolutionary history of figs (Herre et al., 1996; Weiblen, 2000; Jousselin et al., 2003; Rønsted et al., 2005, 2008a; Xu et al., 2011; Cruaud et al., 2012). This work guides the re-evaluation of the current classification of figs primarily based on morphology (Berg and Corner, 2005) which divides *Ficus* into six subgenera (*Ficus*, *Pharmacosycea*, *Sycidium*, *Sycomorus, Synoecia*, and *Urostigma*) and a number of sections – see **Figure 4** and **Chapter I** for further information.

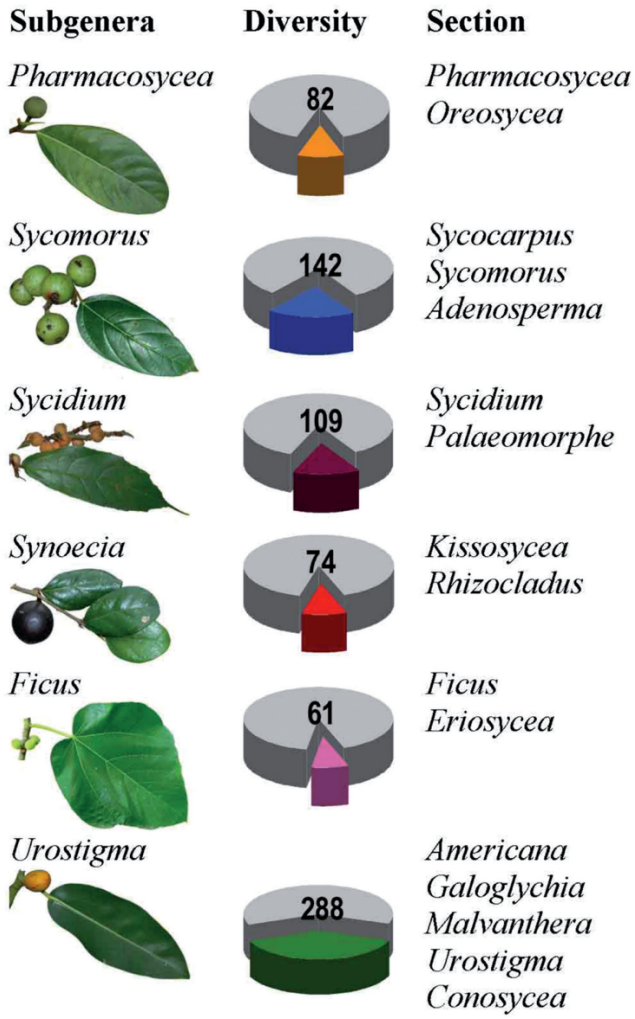


Figure 4 | **Classification of *Ficus***. The numbers of species per subgenus is represented as a proportion of total *Ficus* species richness. Modified from (Cruaud et al., 2012).

The most recent and comprehensive phylogenetic work on figs by Cruaud et al. (2012) sampled ~200 species from five rapidly evolving coding and non-coding nuclear markers (ITS, ETS, *G3pdH*, waxy, *ncpGS*) and did not obtain support for the monophyly of three of the six subgenera – *Ficus*, *Pharmacosycea*, and *Urostigma* (**Figure 5**). Conflict between the morphological

classification and the phylogenetic results based on DNA sequences was also reported in previous studies (Jousselin et al., 2003; Rønsted et al., 2005, 2008a). Furthermore, in the same study by Cruaud et al. (2012) it was found that the pollinators (genus *Tetrapus* Mayr), of section *Pharmacosycea,* was embedded within the phylogenetic tree of the Agaonidae (**Figure 5**). This was unexpected based on previous results showing *Pharmacosycea* to be sister to the remainder of *Ficus* and the wasps' pollination biology with affinity exclusively to section *Pharmacosycea.* Considering the high degree of specificity and co-diversification in the fig-wasp mutualism, this result raises the question of whether *Pharmacosycea* and *Tetrapus* are indeed sisters to the remainder of figs and their pollinating wasps respectively, or whether we simply have not resolved the phylogenetic history of the fig-wasp mutualism yet. These findings further complicate the reevaluation of the classification and evolutionary history for figs.

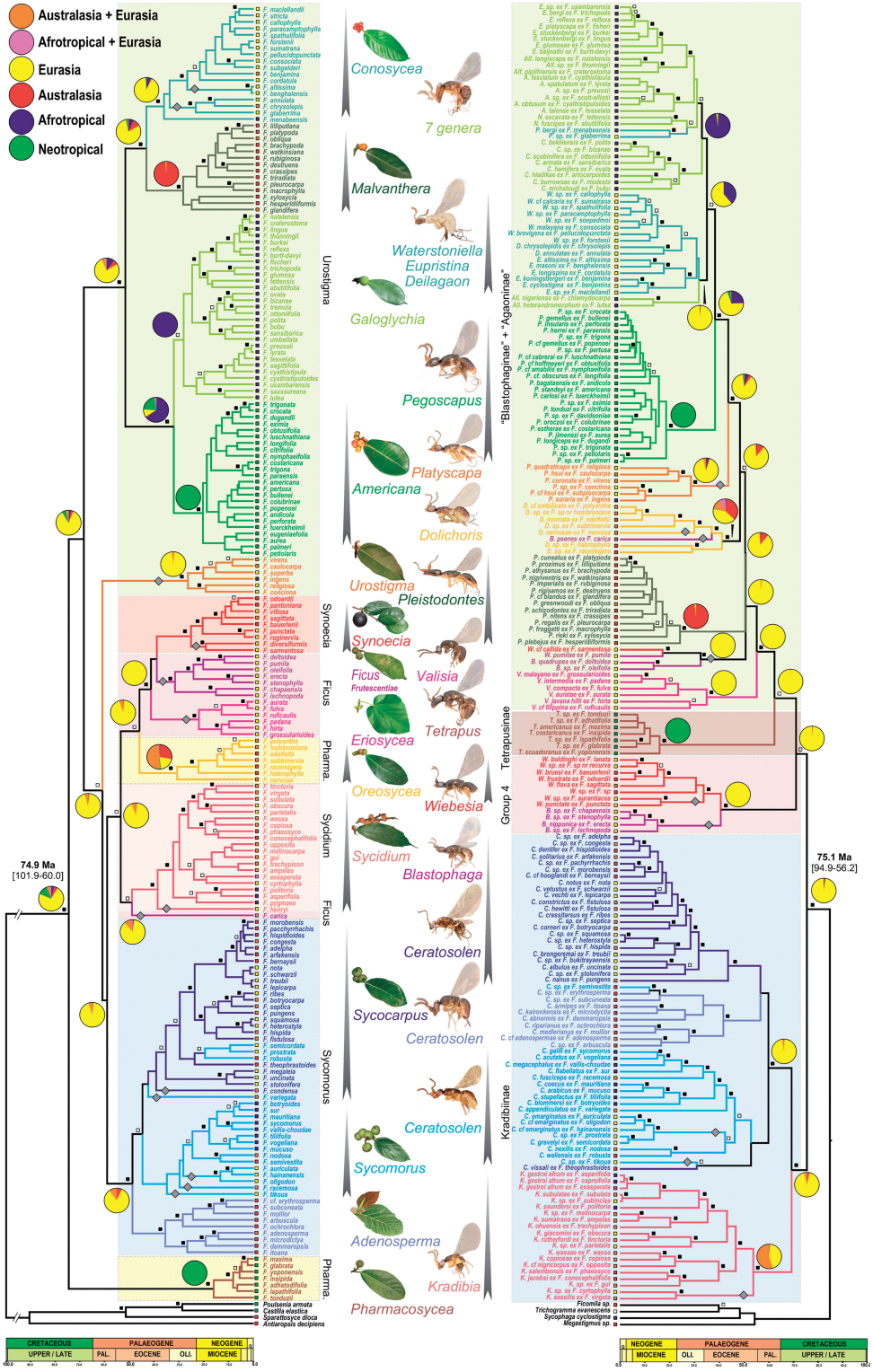

Australasia + Eurasia
Afrotropical + Eurasia
Eurasia
Australasia
Afrotropical
Neotropical
Conosycea
7 genera
Malvanthera
Waterstoniella
Eupristina
Deilagaon
Galoglychia
Pegoscapus
Platyscapa
Americana
Dolichoris
Urostigma
Pleistodontes
Synoecia
Ficus
Frutescentiae
Valisia
Tetrapus
Eriosycea
Oreosycea
Wiebesia
Sycidium
Blastophaga
Ceratosolen
Sycocarpus
Ceratosolen
Ceratosolen
Sycomorus
Adenosperma
Kradibia
Pharmacosycea
Urostigma
Synoecia
Ficus
Pharma.
Sycidium
Ficus
Sycomorus
Pharma.
"Blastophaginae" + "Agaoninae"
Group 4 Tetrapusinae
Kradibiinae
74.9 Ma
[101.9-60.0]
75.1 Ma
[94.9-56.2]
Ficomila sp.
Trichogramma evanescens
Sycophaga cyclostigma
Megastigmus sp.
Pouteria armata
Castilla elastica
Sparattosyce dioica
Antiaropsis decipiens
CRETACEOUS
PALAEOGENE
NEOGENE
UPPER / LATE
PAL.
EOCENE
OLI.
MIOCENE

Figure 5 (previous page) | Phylogenetic inference of *Ficus* and pollination wasps showing co-diversification. *Ficus* phylogenetic tree based on 5 nuclear markers and dated using secondary calibrations and a single fossil. From Cruaud et al. (2012).

Despite a consistent increase in genetic data and species sampling of *Ficus* through more than a decade, the current status of the phylogenetic tree of *Ficus* does not provide sufficient resolution or clade support to unambiguously resolve relationships deep within *Ficus*, thus leaving the backbone of the phylogenetic tree uncertain (**Figure 5**). Lack of a well-supported and densely sampled phylogenetic hypothesis for *Ficus* has hindered progress on key research questions regarding diversification, biogeography and species interactions (Herre et al., 2008; Cruaud et al., 2012). Lack of resolution has also hindered our understanding of the evolution and adaptations of *Ficus* within clades. For example, the Australasian *Ficus* section *Malvanthera* (Rønsted et al., 2008b; **Chapter V**) contains both rainforest hemi-epiphytes, lithophytes adapted to drier environments and transitional species. Again, lack of resolution obtained by Rønsted et al. (2008b) did not allow for testing of a hypothesis of radiation from rainforest to savanna though transitional species.

Most previous evolutionary studies of figs have focused on using rapidly evolving nuclear markers. Chloroplast markers are commonly employed in plant phylogenetic studies but often lack variability to confidently resolve relationships at infrageneric levels, especially within large genera (Rønsted et al., 2007b; Roy et al., 2010). To date, only few studies have employed chloroplast markers and mostly on a small or taxonomically narrow sample of *Ficus*. Herre et al. (1996) produced the first

published molecular phylogenetic hypotheses of figs including only 15 species based on *trnL-F* and *rbcL* chloroplast markers. In 2009 Renoult and coworkers sampled five non-coding plastid markers for 38 species of African figs in the section *Galoglychia.* They discovered significant conflicts when the plastid phylogenetic inference they produced was compared to the phylogenetic inference of *Galoglychia* produced by Rønsted and coworkers (2007a) based on nuclear ribosomal DNA markers ETS (nuclear ribosomal external transcribed spacer) and ITS (nuclear ribosomal internal transcribed spacer). As such, the evolutionary history of *Ficus* reconstructed from plastid markers has yet to be fully explored leading to **Chapter IV** in this thesis where near-complete plastid genomes (plastomes) were used to infer a phylogenetic hypothesis.

The position of section *Pharmacosycea* as sister to all other *Ficus* has also been questioned as introduced above. Prior studies including an outgroup of other genera in the Moraceae have all recovered this relationship (Herre et al., 1996; Rønsted et al., 2005, 2008a; Xu et al., 2011; Cruaud et al., 2012), but not with significant support and it has been suggested that the recovered placement of section *Pharmacosycea* may be the result of long-branch attraction with the outgroup taxa based on lack of characters in common with the ingroup (Cruaud et al., 2012) although tests have not been able to confirm this. The implication of and interest in the placement of *Pharmacosycea* come from a long-standing debate about where *Ficus* originated. As *Pharmacosycea* is a Neotropical lineage, intuitive interpretation based on *Pharmacosycea* being sister to the

remainder of *Ficus*, leads to a hypothesis of a Southern Gondwanan origin (Machado et al., 2001; Rønsted et al., 2005) rather than a Eurasian origin suggested by ancestral area analysis (Zerega et al., 2005; Cruaud et al., 2012). Additionally, in the phylogenetic hypothesis of Agaonidae, *Tetrapus*, the genus of pollinators associated with section *Pharmacosycea*, was embedded in Agaonidae rather than recovered as sister to all other agaonids (Cruaud et al., 2012). The mismatch between *Ficus* and the Agaonidae phylogenetic topologies contradicts the theory of strict cospeciation among *Ficus* and their pollinating wasps (Machado et al., 2005). The closest relatives of *Ficus* in the subtribe Castilleae (Zerega et al., 2005) contain a mixture of Neotropical and Australasian genera. A better understanding of the evolutionary history of *Ficus* seems to be the only way forward to resolve the debate of the origin of *Ficus*.

The purpose of this thesis work has therefore been to use both traditional Sanger sequencing and high through-put sequencing techniques to help resolve the evolutionary history and infrageneric classification of *Ficus* as well as to shed new light on Darwin's abominable mystery of angiosperm diversification.

Methods, techniques and technology

Many old and profound questions in evolutionary biology can now be re-addressed with the improvement of molecular techniques over the last 10-25 years or so (Savolainen and Chase, 2003; Harrison and Kidner, 2011; Soltis et al., 2013). A set of different molecular techniques from the classical Polymerase

Chain Reaction (PCR) to innovative high-throughput hybridization captures techniques have been used in this thesis work. In the following, a brief overview is provided of the methods applied to generate data in this thesis.

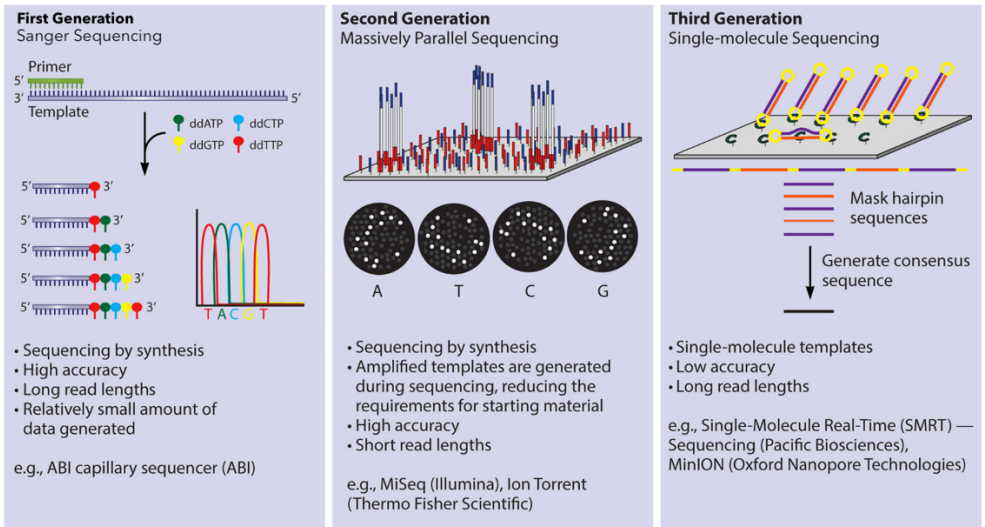


Figure 6 | Overview of first to third generation technology of sequencing. Highlighted is examples of platforms and pros/cons of each. Modified from Ronholm et al. (2016).

For over 40 years, sequencing was performed using the so-called Sanger sequencing method that requires a single stranded DNA template and will use modified dideoxynucleoside triphosphates (ddNTP's) to terminate the elongation process at different lengths together covering the entire DNA template. Subsequently, ddNTP's labeled with fluorescent dyes can be detected using chromatographs to decode the template DNA sequence. Plant molecular systematics has undergone a revolution over the past decades with the breakthrough of DNA

sequences to infer evolutionary relationships among species (Chase et al., 1993; Soltis et al., 1998, 2000, 2011; Ruhfel et al., 2014). Sanger sequencing is still a tremendously powerful method widely used for decoding DNA sequences due to its simplicity and well-known workflow. However, the high-throughput sequencing (HTS) methods revolutionized the way sequencing works. Several technologies have been developed in a very short amount of time and costs have gone down but are still substantial, hindering their general use in taxonomy and across larger genera (Delseny et al., 2010; Metzker, 2010). The continuous aim to reduce costs and time for sequencing while increasing data output led to the invention of second generation technology (**Figure 6**), e.g. Ion Torrent's PGM and the Illumina MiSeq/HiSeq platforms, which applies sequencing by synthesis using reversible terminator technology and thereby reaching a very high output compared to earlier platforms (Bentley et al., 2008; Quail et al., 2012). Other so-called long-read sequencing techniques (**Figure 6**) like Pacific Biosciences and the Oxford Nanopore platforms work by reading single molecules rather than strands, but are still having high error rates making them primarily useful for complementing the standard Illumina and other techniques (Bleidorn, 2016).

DNA sequencing methods are constantly under development trying to find alternative options to make it faster, more precise, easier and less costly. Tunneling current sequencing and sequencing by hybridization using microarrays is under development and an alternative to second and third generations technologies (Qin et al., 2012; Di Ventra, 2013).

HTS approaches greatly improves resolution for closely related species and are particularly valuable for resolving relationships between species with low genetic variation, and have been found to outperform previously widely used microsatellite data (Baird et al., 2008). In addition to direct sequencing techniques, new HTS based fingerprinting methods such as Genome by Sequencing (GBS) and Restriction site associated DNA (RADseq) have also been developed for population level studies (Miller et al., 2007; Elshire et al., 2011).

Most HTS technologies have a couple of steps that are universal. They require DNA to be sheared, which can be done enzymatically or with sonication depending on preference, need and availability. In the library preparation steps the goal is to repair the ends of the DNA, plus, ligate adaptors to the double DNA strands making them sequencing able. After library preparation the sample can in principle be sequenced using so-called shotgun sequencing, however, as the names suggests this is a very unspecific method. Although this principally makes it possible to sequence the whole genome of any organism in a single run, this is often neither required nor desired as a lot of the sequencing capacity is spent on unused data. However, in the case of **Chapter IV**, multiple samples where indexed and pooled into a single lane. The resulting coverage was still high making it possible to extract further data from each sample besides the plastome, which could be used to verify the samples.

Differences in copy number and size of nuclear, plastome, and mitochondrial genomes and factors such as coverage depth needs, experimental focus, sample number, and cost can make target enrichment more suitable. The two basic approaches to

target enrichment is PCR-based amplification and hybrid capture. These methods make it possible to focus on specific regions of interest and increasing depth of coverage of targeted sequences (Ávila-Arcos et al., 2011; Ekblom and Galindo, 2011). Currently the most widely used method for targeted enrichment is based on in-solution target enrichment where biotinylated baits are manufactured on known sequences (see **Chapter V**).

In this thesis work I have applied Sanger sequencing of PCR products and products from molecular cloning, together with different platforms of Illumina sequencing using and further optimizing different library preparation kits and protocols.

High-throughput sequencing of plants

The low cost of HTS is making it achievable for non-model plants, but as highlighted by Hirsch and Buell (2013), at least four major factors hinder the process of obtaining sequence data in a standardized pipeline: the extent of genome duplication (segmental, tandem, and whole-genome), the heterozygosity, the ploidy level, and repetitive sequence composition. Genome duplication is assumed to be a factor in the evolution and diversification of plants as it enables the evolution of new gene functions, rearrangements and is likely driving speciation events. The origin of angiosperms was probably accompanied by a genome duplication event indicating that even the ancestral angiosperm was a polyploid with an assembly of both unique and ancient genes that survived to play a role in angiosperm biology and evolution (Unamba et al., 2015). Several other duplication

events in separate lineages of angiosperms have likewise been identified (Cui et al., 2006; Soltis and Soltis, 2009).

The consequence of the above for using DNA sequencing means an extensive phase of trial and error before a working pipeline can be applied without, or with limited, verification. For example, even the otherwise routinely used ITS and nearby regions such as ETS in plants are still criticized for the lack of knowledge about their conservation mechanism when using Sanger sequencing, which may lead to multiple copies being sequenced and compromising interpretation of results (Soltis et al., 1998; Calonje et al., 2009) as we also explored for ETS in **Chapter I**. The explorative phase of HTS of plants is still on-going – as is the development of pipelines to analyze the data. The go-to pipelines that work well in the animal Kingdom, with much less potential genomic variability, does not apply to the plant kingdom due to the reasons above.

The downstream analysis and inference of evolutionary history through phylogenetics is often based on the assumption of dichotomous branching, which is rarely met due to hybridization, introgression and polyploidy (Lemey et al., 2009). Therefore, we need more examples and many more attempts to use the many genomic tools available today. We might continue to run into issues along the way, but as long as they are pointed out and considered in the most appropriate way, the promises of exploring plant genomics in a phylogenetic framework are very worthwhile.

Ficus *in the high-throughput sequencing era*

The huge interest in the evolutionary history of *Ficus* has resulted in a plethora of studies using many different approaches (see **Chapter I**). Especially PCR and Sanger sequencing have been applied, but only a few studies have taken the figs into the HTS era besides the work in this thesis. The latest applied use of RADseq data (Rasplus et al., 2018) shows the possibility to use the RADseq in *Ficus*, however, their findings contradicts many previous findings with regards to the first divergent lineage and sister to the rest of *Ficus* not being section *Pharmacosycea* (see **Chapter IV** for further information). In **Chapter IV**, plastome data provides full statistical support for *Pharmacosycea* as the first divergent lineage.

The huge phylogenetic work on the global phylogenetic hypothesis of *Ficus* produced over the last two decades all experience lack of resolution in the short internal branches that should resolve clade relationships (Weiblen, 2000; Jousselin et al., 2003; Rønsted et al., 2005, 2008a; Xu et al., 2011; Cruaud et al., 2012). Even using rapidly evolving markers such as single-copy coding and non-coding nuclear markers to infer the evolutionary history has not provided enough resolution of these short internal branches (Rønsted et al., 2008a; Cruaud et al., 2012). This indicates that *Ficus* have experienced several rapid diversification events or possibly isolations (e.g. reproductive or geographical) leading to very little time to accumulate differences in their genetic makeup, resulting in short branches separating clades. Moreover, as only few differences have had time to accumulate in the DNA, the statistical support inferred using e.g. bootstrap support or posterior probabilities, will be

low as a consequence of short branches – i.e. little resolution is observed (Lemey et al., 2009).

In this thesis work, HTS genome skimming techniques were used to explore the evolution of plastomes across *Ficus* (**Chapter IV**) and a targeted sequence capture approach was developed to capture genomic information across the genomes and provide a powerful tool for resolving the evolutionary history of *Ficus* at various taxonomic scales (**Chapter V**).

Targeted sequence capture

Working with phylogenetic inferences the power and scale of HTS provides the necessary sequencing depth for phylogenetic analysis and pairing with target capture technology can greatly increase discovery (**Figure 7**). Isolating specific genomic loci, results in enriched pools of target sequences, thus reducing wasted coverage on regions with little/too much information. By targeting specific regions or loci, costs can also be reduced compared to sequencing of whole genomes. In the design of probes that target specific traits of the (phylogenetic) markers is very useful in evolutionary biology where clade specific issues, such as short internal branches, can appear (Grover et al., 2012; Tennessen et al., 2013).

The development of techniques and pipelines that allow for processing samples for species that currently lack reference genomes have advanced evolutionary genomic studies on many

plants that have not been studied with HTS before. However, lack of a reference genome complicates downstream analysis and could potentially result in subsequent phylogenetic inference achieving low levels of support (Grover et al., 2012; Straub et al., 2012; Jones and Good, 2015).

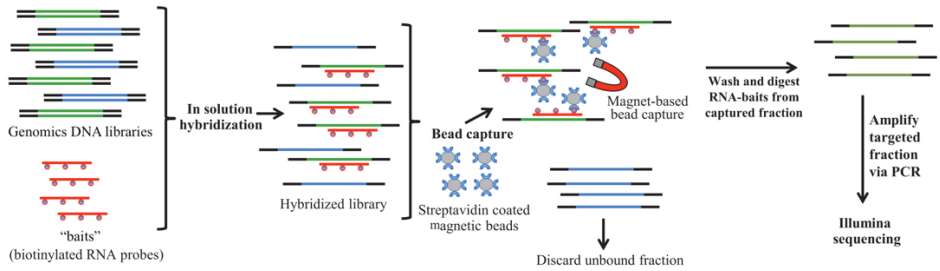


Figure 7 | Schematic overview of sequence capture. Probes/baits bind to specific sequences, which already are indexed and with adaptors attached. The retained sequences are PCR amplified, pooled and used for sequencing. Modified from Soltis et al. (2013)

Bioinformatic pipelines for HTS data

The processing of HTS data present a number of challenges and advantages. Projects involving HTS approaches result in vast volumes of DNA sequence data (reads) and require high-performance computer systems for data processing and storage – a laptop computer will often have too little power for these tasks. Additionally, many of the analysis steps are greatly memory- and CPU-intensive. There has been a huge progression in bioinformatic programs being published, to match the analysis associated with data handling (see e.g. Godden et al., 2012). To discuss the toolbox available for HTS analysis would be a mission

in itself. However, generally there are certain steps, which are universal for the analysis; such as quality check (QC) and cleaning of the data to get past issues from low quality reads and adaptor leftovers in the reads. Most times mapping to a reference genome will give information useable in phylogenetics, sometimes this is not possible and *a de novo* assembly of for example the plastome will follow (**Figure 8**). Subsequently, phylogenetic inference is deducted (see **Chapters IV+V)**. The range of possibilities expands impressively depending on data and goal (Godden et al., 2012; Soltis et al., 2013).

Most programs for analysis of HTS data are Unix/Linux command-line based, and often require custom scripts to be developed, manipulated or parse the input and output datafiles. Thus, the handling of HTS data requires ample bioinformatics skills, which is mainly attained by trained bioinformaticians. Without the proper training or assistance, a number of graphical and web-based interfaces exists that can be applied for the data processing and analysis of more simple tasks. The graphical interfaced software available, such as Geneious (www.geneious.com, Biomatters Limited, New Zealand) are practical, but flexibility can be missing for some analysis, and they often require extensive analysis times on a desktop computer. Several limitations exist, for example, absence of possibilities to adjust several settings, which would be possible in command-line based programs. However, using these graphical based programs, instead of command-line workflows, will make the barrier of missing skill-sets be easier to overcome in some cases and allow for a broad application of HTS techniques.

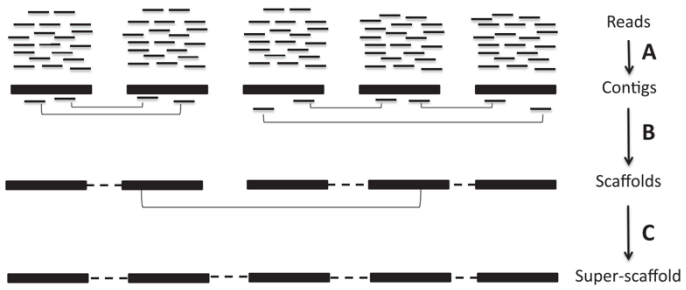


Figure 8 | General strategy for genome assembly. (A) DNA reads are assembled into contigs based on overlaps. (B) contigs are linked together into scaffolds based on paired read information. (C) super-scaffolds are made from e.g. PCR based information or further information in the DNA data itself. Modified from Soltis et al. (2013).

Conclusions and perspectives

Ficus and their pollinators offer a unique model system for comparative biology. Therefore, substantial research into the system and especially their evolutionary history, will aid the understanding of mutualisms and diversification patterns on a larger scale. The mutualism is well studied, which makes the system perfect for explaining some of the challenging questions in evolutionary biology on both *Ficus* and their pollinators.

Ficus have been studied using the traditional Sanger method for more than a decade and several well supported clades have been documented with strong support, however their relationships are still uncertain (**Chapter I**). It has proven challenging to infer a robust phylogenetic hypothesis due to

limitations in the Sanger method, consequently not providing sufficient resolution in large genera like *Ficus*. This limits the interpretation of species relationships, diversification patterns, codiversification and other studies (**Chapter I-III**). The phylogenetic hypothesis at genus level has shown issues along the backbone of the phylogeny due to short, internal branches, causing low statistical support in e.g. bootstrap percentages. Using HTS methods (such as plastome data) have shown to be a powerful tool for resolving the short internal branches with strong support in the phylogenetic inferences (**Chapter IV**) and provides substantial resolution and information in internal branches of the phylogenetic tree (**Chapter IV-V**).

The capacity of HTS methods will make it possible to re-visit the global phylogenetic hypothesis for *Ficus* and hopefully resolve the relationships of clades recovered and discussed in **Chapter I.** The use of plastome data to resolve clades have proven to be powerful (**Chapter IV**), but due to the complicated evolutionary history of the plastome in *Ficus*, the plastome data is not suitable for systematic purposes on a global scale. Nonetheless, the results of **Chapter IV** have provided an important insight into the evolutionary dynamics in past of *Ficus* revealing cyto-nuclear discordance with potential ancient introgression of the plastome. Thus, care should be taken in using plastid markers for phylogenetic inferences.

The use of targeted sequence capture methods in **Chapter V** have proven to be extremely useful in the Australasian section *Malvanthera*. The bait set developed for this chapter has also been applied on global scale (not included in this thesis), where preliminary analysis has shown similar promising results in

recovering strongly supported phylogenetic trees with clades corresponding to known relationships based on morphological characters, biogeography and pollinators. Hence, the next step will be to construct a phylogenetic hypothesis for the genus using targeted sequence capture. Furthermore, we have tested the method on tissues from historical herbarium specimens with promising results, making it possible to endeavor in retrieving DNA from the historically important collections in herbaria. Due to the huge work undertaken by previous and present day *Ficus* taxonomists, many important species of *Ficus* can be found in herbaria with trustworthy species identification, which naturally otherwise constitute a problem in such large and complicated genera as *Ficus*.

The application of *Ficus* research, as mentioned, will make it possible to dive into questions such as how a strict and well-functioning mutualism originates, and continues to existent with vast number of trophic layers, and internal competition between *Ficus* and pollinators. Our research has so far shown the need for seeing *Ficus* in a large perspective and in principle not as a single genus (as shortly mentioned above). *Ficus* have a strong and important tie to their pollinators. Every assumption should include their perspective too, to ensure the best possible conclusions. Furthermore, phylogenetic inference is one thing for systematics and diversification analysis; when dealing with classifications of such complicated genera, it is important to remember classifications must be useful in understanding key aspects and characters of the taxon in question. Classifications hereby become a synergistic result, where information from

morphology, biogeography, pollination biology etc. will be included. Still, we will never know the true underlying phylogeny, making it even more essential to evaluate classifications in the light of all available evidence.

Using an old, pan-tropical genus as *Ficus* to explain and understand diversification patterns has also proven to be beneficial for understanding global patterns, rather than focusing on local patterns that might not apply elsewhere. However, the lack of a robust phylogenetic hypothesis is a hindrance for any applied use (**Chapters II-III**). *Ficus*, and especially their hemi-epiphytic habit, exhibits an extraordinary ability to adapt to niches that other plants cannot exploit due to the physiological adaptions of hemi-epiphytes and the ability to ensure high pollination success. The ability to adapt to new niches is also seen in other successful pioneer plants that are generally opportunistic species with low competitive abilities such as orchids, *Piper* L. species and Bromeliads (Harrison, 2005; Silvestro et al., 2014; Frenzke et al., 2016). Additionally, there has been a shift in our understanding of how plant groups, with high species numbers, might be accumulating. Previously it was believed that one specific trait (a key innovation) was responsible for the success of a group, however, in recent years it has become apparent that several (key) traits might constitute a synergistic syndrome (Donoghue, 2005). Thus, the paradigm is changing – and with an abundance of diverse traits, *Ficus* continues to be an extraordinary model system for addressing questions in evolutionary biology.

Returning to Darwin's abdominal mystery, of why some lineages are more diverse in species numbers and morphology than others, *Ficus* have proven useful for casting new light on the mystery. The right conditions and circumstances have been present for some *Ficus* lineages to be evolutionary successful, a success that does not appear to be related to one key trait. Moreover, the huge diversification of flowering plants could be seen as a synergistic result of geological changes/opportunities, together with new possibilities for niche-growth and exploration of insects and mammals for pollination and dispersal (Harrison, 2005; Soltis et al., 2008; Cardelu et al., 2012; Wang et al., 2016). Darwin's mystery still remains but with a better understanding of potential explanations from the perspective of *Ficus*.

Objectives of the thesis

Present an improved phylogenetic framework for *Ficus* and Castilleae using 307 *Ficus* species and 43 Castilleae species - the most robust species sampling to date. Using the phylogenetic tree reconstructed here as a framework to suggest revisions to the classification of Involucraoideae to reflect evolutionary relationships (**Chapter I)**

Investigate diversification dynamics in genus *Ficus* and drivers of diversification (**Chapters II & III).**

Apply high-throughput sequencing technology to the phylogenetic reconstruction of *Ficus* on global and local scale. Using both plastid genome sequences (**Chapter IV**) and targeted sequence capture (**Chapter V**) and thereafter assess the usefulness of these techniques to understand the evolutionary history of *Ficus*.

Photo credits for introduction

Figure 1

Wolffia globose: https://goo.gl/images/hdYUek

Sequoiadendron giganteum: https://goo.gl/images/QCZQFu

Rafflesia arnoldii: https://goo.gl/images/d6bimS

Figure 2

Ficus pleurocarpa (sample NR772) by Sam Bruun-Lund

F. benghalensis (sample NR767) *& F. watkinsiana* (NR752) taken by Sam Bruun-Lund.

F. drupaceae:
https://upload.wikimedia.org/wikipedia/commons/a/a4/Ficus_drupacea_%28Mysore_Fig%29_in_Bhongir_fort%2C_AP_W_IMG_2955.jpg,

F. Sycomorus:
https://upload.wikimedia.org/wikipedia/commons/4/4f/Ficus_sycomorus_near_Segeneyti_Eritrea.jpg,

F. minahassae: https://alamendah.files.wordpress.com/2011/03/langusei-ficus-minahassae.jpg,

F. dammaropsis: https://davesgarden.com/guides/pf/showimage/149486/#b,

F. punctate: https://floraofsingapore.files.wordpress.com/2011/05/img_3608.jpg

Figure 3

Ficus ottonifolia by Sam Bruun-Lund in the Botanical Garden Natural History Museum DK

References

Ávila-Arcos, M.C., Cappellini, E., Romero-Navarro, J.A., Wales, N., Moreno-Mayar, J.V., Rasmussen, M., Fordyce, S.L., Montiel, R., Vielle-Calzada, J.-P., Willerslev, E., Gilbert, M.T.P., 2011. Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. *Scientific Reports*, ***1***. https://doi.org/10.1038/srep00074

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Zachary, A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD Markers ***3***, 1–7. https://doi.org/10.1371/journal.pone.0003376

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Helen, R., Boutell, J.M., Bryant, J., Carter, R.J., Cheetham, R.K., Cox, A.J., Ellis, D.J., Flatbush, M.R., Gormley, N.A., Sean, J., Irving, L.J., Karbelashvili, M.S., Kirk, S.M., Li, H., Maisinger, K.S., Murray, L.J., Obradovic, B., Ost, T., Michael, L., Pratt, M.R., Rasolonjatovo, I.M.J., Reed, M.T., Rigatti, R., Rodighiero, C., Ross, M.T., Sabot, A., Sankar, S. V, Schroth, G.P., Smith, M.E., Smith, V.P., Spiridou, A., Peta, E., Tzonev, S.S., Vermaas, E.H., Walter, K., Wu, X., Zhang, L., Banerjee, S., Barbour, S.G., Baybayan, P.A., Benoit, V.A., Bridgham, A., Brown, R.C., Brown, A.A., Buermann, D.H., Bundu, A.A., Cooley, R.N., Crake, N.R., Dada, O.O., Konstantinos, D., 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, ***456***, 53–59. https://doi.org/10.1038/nature07517.Accurate

Berg, C.C., Corner, E.J.H., 2005. Flora Malesiana, Series I. Volume 17 part 2., in: Nooteboom, H. (Ed.), Flora Malesiana Series I - Seed Plants Vol. 17 Part 2. Nationaal Herbarium Nederland, Leiden, pp. 1–730.

Bleidorn, C., 2016. Third generation sequencing: Technology and its potential impact on evolutionary biodiversity research. *Systematics and Biodiversity*, ***14***, 1–8. https://doi.org/10.1080/14772000.2015.1099575

Calonje, M., Martín-Bravo, S., Dobeš, C., Gong, W., Jordon-Thaden, I., Kiefer, C., Kiefer, M., Paule, J., Schmickl, R., Koch, M. a., 2009. Non-coding nuclear DNA markers in phylogenetic reconstruction. *Plant Systematics and Evolution*, ***282***, 257–280. https://doi.org/10.1007/s00606-008-0031-1

Cardelu, C.L., Fern, T., Cycle, L., Farrar, R., 2012. Ferns in an angopsperm world: Cretaceous radiation into the epiphytic niche and diversification on the forest floor. *International Journal of Plant Sciences*, ***173***, 695–710.

https://doi.org/10.1086/665974

Chamala, S., Chanderbali, a. S., Der, J.P., Lan, T., Walts, B., Albert, V. a., DePamphilis, C.W., Leebens-Mack, J., Rounsley, S., Schuster, S.C., Wing, R. a., Xiao, N., Moore, R., Soltis, P.S., Soltis, D.E., Barbazuk, W.B., 2013. Assembly and validation of the genome of the nonmodel basal angiosperm Amborella. *Science*, ***342***, 1516–1517. https://doi.org/10.1126/science.1241130

Chase, M.W., Soltis, D.E., Olmstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., Duvall, M.R., Price, R.A., Hills, H.G., Qiu, Y.-L., Kron, K.A., Rettig, J.H., Conti, E., Palmer, J.D., Manhart, J.R., Sytsma, K.J., Michaels, H.J., Kress, W.J., Karol, K.G., Clark, W.D., Hedren, M., Gaut, B.S., Jansen, R.K., Kim, K.-J., Wimpee, C.F., Smith, J.F., Furnier, G.R., Strauss, S.H., Xiang, Q.-Y., Plunkett, G.M., Soltis, P.S., Swensen, S.M., Williams, S.E., Gadek, P.A., Quinn, C.J., Eguiarte, L.E., Golenberg, E., Learn Jr., G.H., Graham, S.W., Barrett, S.C.H., Dayanandan, S., Albert, V.A., 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbcL*. *Annals of the Missouri Botanical Garden*, ***80***, 528–580. https://doi.org/10.2307/2399846

Clement, W.L., Weiblen, G.D., 2009. Morphological evolution in the mulberry family (Moraceae). *Systematic Botany*, ***34***, 530–552. https://doi.org/10.1600/036364409789271155

Coiro, M., Doyle, J.A., Hilton, J., 2019. How deep is the conflict between molecular and fossil evidence on the age of angiosperms?, New Phytologist. https://doi.org/10.1111/nph.15708

Cook, J.M., Rasplus, J.Y., 2003. Mutualists with attitude: Coevolving fig wasps and figs. *Trends in Ecology and Evolution*, ***18***, 241–248. https://doi.org/10.1016/S0169-5347(03)00062-4

Couvreur, T.L.P., Forest, F., Baker, W.J., 2011a. Origin and global diversification patterns of tropical rain forests: Inferences from a complete genus-level phylogeny of palms. *BMC Biology*, ***9***, 44. https://doi.org/10.1186/1741-7007-9-44

Couvreur, T.L.P., Pirie, M.D., Chatrou, L.W., Saunders, R.M.K., Su, Y.C.F., Richardson, J.E., Erkens, R.H.J., 2011b. Early evolutionary history of the flowering plant family Annonaceae: Steady diversification and boreotropical geodispersal. *Journal of Biogeography*, ***38***, 664–680. https://doi.org/10.1111/j.1365-2699.2010.02434.x

Crepet, W.L., Niklas, K.J., 2009. Darwin's second "abominable mystery": Why

are there so many angiosperm species? *American Journal of Botany*, ***96***, 366–381. https://doi.org/10.3732/ajb.0800126

Cruaud, A., Rønsted, N., Chantarasuwan, B., Chou, L.S., Clement, W.L., Couloux, A., Cousins, B., Genson, G., Harrison, R.D., Hanson, P.E., Hossaert-McKey, M., Jabbour-Zahab, R., Jousselin, E., Kerdelhué, C., Kjellberg, F., Lopez-Vaamonde, C., Peebles, J., Peng, Y.-Q., Pereira, R.A.S., Schramm, T., Ubaidillah, R., van Noort, S., Weiblen, G.D., Yang, D.-R., Yodpinyanee, A., Libeskind-Hadas, R., Cook, J.M., Rasplus, J.-Y., Savolainen, V., 2012. An extreme case of plant-insect codiversification: Figs and fig-pollinating wasps. *Systematic Biology*, ***61***, 1029–47. https://doi.org/10.1093/sysbio/sys068

Cui, L., Wall, P.K., Leebens-mack, J.H., Lindsay, B.G., Soltis, D.E., Doyle, J.J., Soltis, P.S., Carlson, J.E., Arumuganathan, K., Barakat, A., Albert, V.A., Ma, H., Claude, W., 2006. Widespread genome duplications throughout the history of flowering plants 738–749. https://doi.org/10.1101/gr.4825606.738

Darwin , C. 1871 . On the origin of species by means of natural selection. Appleton, New York, New York, USA.

Darwin , F. [ed.]. 1903 . More letters of Charles Darwin, a record of his work in hitherto unpublished letters, vol. 2. John Murray, London, UK.

Davies, T.J., Barraclough, T.G., Chase, M.W., Soltis, P.S., Soltis, D.E., Savolainen, V., 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences*, ***101***, 1904–1909. https://doi.org/10.1073/pnas.0308127100

Delseny, M., Han, B., Hsing, Y.I., 2010. High throughput DNA sequencing: The new sequencing revolution. *Plant Science*, ***179***, 407–422. https://doi.org/10.1016/j.plantsci.2010.07.019

Di Ventra, M., 2013. Fast DNA sequencing by electrical means inches closer. *Nanotechnology*, ***24***, 2–4. https://doi.org/10.1088/0957-4484/24/34/342501

Donoghue, M.J., 2005. Key innovations, convergence, and success: Macroevolutionary lessons from plant phylogeny. *Paleobiology*, ***31***, 77–93. https://doi.org/10.1666/0094-8373(2005)031[0077:KICASM]2.0.CO;2

Ekblom, R., Galindo, J., 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, ***107***, 1–15. https://doi.org/10.1038/hdy.2010.152

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, ***6***, 1–10. https://doi.org/10.1371/journal.pone.0019379

Frenzke, L., Goetghebeur, P., Neinhuis, C., Samain, M.-S., Wanke, S., 2016. Evolution of epiphytism and fruit traits act unevenly on the diversification of the species-rich genus *Peperomia* (Piperaceae). *Frontiers in plant science*, ***7***, 1145. https://doi.org/10.3389/fpls.2016.01145

Frodin, D.G., 2004. History and concepts of big plant genera. *Taxon*, ***53***, 753–776.

Gentry, A.H., 1992. Tropical forest biodiversity: Distributional patterns and their conservational significance. *Nordic Society Oikos*, ***63***, 19–28.

Godden, G.T., Jordon-Thaden, I.E., Chamala, S., Crowl, A. a., García, N., Germain-Aubrey, C.C., Heaney, J.M., Latvis, M., Qi, X., Gitzendanner, M. a., 2012. Making next-generation sequencing work for you: Approaches and practical considerations for marker development and phylogenetics. *Plant Ecology & Diversity*, ***5***, 427–450. https://doi.org/10.1080/17550874.2012.745909

Gómez, J.M., Verdú, M., 2012. Mutualism with plants drives primate diversification. *Systematic Biology*, ***61***, 567–577. https://doi.org/10.1093/sysbio/syr127

Grover, C.E., Salmon, A., Wendel, J.F., 2012. Targeted sequence capture as a powerful tool for evolutionary analysis. *American Journal of Botany*, ***99***, 312–319. https://doi.org/10.3732/ajb.1100323

Harrison, N., Kidner, C., 2011. Next-generation sequencing and systematics. What can a billion base pairs of DNA sequence data do for you? *Taxon*, ***60***, 1552–1566.

Harrison, R.D., 2005. Figs and the diversity of tropical rainforests. *BioScience*, ***55***, 1053. https://doi.org/10.1641/0006-3568(2005)055[1053:FATDOT]2.0.CO;2

Harrison, R.D., Rønsted, N., Xu, L., Rasplus, J.-Y., Cruaud, A., 2012. Evolution of fruit traits in *Ficus* subgenus *Sycomorus* (Moraceae): To what extent do frugivores determine seed dispersal mode? *PloS one*, ***7***, e38432. https://doi.org/10.1371/journal.pone.0038432

Herre, E.A., Machado, C.A., Bermingham, E., Nason, J.D., Windsor, D.M., McCafferty, S.S., VanHouten, W., Bachmann, K., 1996. Molecular

phylogenies of figs and their pollinator wasps. *Journal of Biogeography*, ***23***, 521–530. https://doi.org/10.1111/j.1365-2699.1996.tb00014.x

Herre, E.A., Jandér, K.C., Machado, C.A., 2008. Evolutionary ecology of figs and their associates: recent progress and outstanding puzzles. *Annual Review of Ecology, Evolution, and Systematics*, ***39***, 439–458. https://doi.org/10.1146/annurev.ecolsys.37.091305.110232

Hirsch, C.N., Buell, C.R., 2013. Tapping the promise of genomics in species with complex, nonmodel genomes. *Annual Review of Plant Biology*, ***64***, 89–110. https://doi.org/10.1146/annurev-arplant-050312-120237

Ikeda, H., Nishikawa, M., Sota, T., 2012. Loss of flight promotes beetle diversification. *Nature Communications*, ***3***. https://doi.org/10.1038/ncomms1659

Jones, M.R., Good, J.M., 2015. Targeted capture in evolutionary and ecological genomics. *Molecular Ecology*, n/a-n/a. https://doi.org/10.1111/mec.13304

Jousselin, E., Rasplus, J., Kjellberg, F., 2003. Convergence and coevolution in a mutualism: Evidence from a molecular phylogeny of *Ficus*. *Evolution*, ***57***, 1255–1269. https://doi.org/10.2307/3448849

Jousselin, E., Kjellberg, F., Herre, E.A., 2004. Flower specialization in a passively pollinated monoecious fig: A question of style and stigma? *International Journal of Plant Sciences*, ***165***, 587–593. https://doi.org/10.1086/386558

Kjellberg, F., Jousselin, E., Bronstein, J.L., Patel, A., Yokoyama, J., Rasplus, J., 2001. Pollination mode in fig wasps: The predictive power of correlated traits. *Proceedings of the Royal Society B: Biological Sciences*, ***6***. https://doi.org/10.1098/rspb.2001.1633

Kjellberg, F., Proffit, M., 2016. Tracking the elusive history of diversification in plant – herbivorous insect – parasitoid food webs: Insights from figs and fig wasps. *Molecular Ecology*, ***25***, 843–845.

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

Lemey, P., Salemi, M., Vandamme, A.-M., 2009. The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing, 2nd ed. Cambridge University Press. https://doi.org/10.1002/ajhb.20017

Machado, C.A., Jousselin, E., Kjellberg, F., Compton, S.G., Herre, E.A., 2001. Phylogenetic relationships, historical biogeography and character evolution of fig-pollinating wasps. *Proceedings. Biological sciences / The Royal Society*, ***268***, 685–694. https://doi.org/10.1098/rspb.2000.1418

Machado, C.A., Robbins, N., Gilbert, M.T.P., Herre, E.A., 2005. Critical review of host specificity and its coevolutionary implications in the fig fig-wasp mutualism. *Proceedings of the National Academy of Sciences of the United States of America*, ***102***.

Metzker, M.L., 2010. Sequencing technologies - the next generation. *Nature Reviews. Genetics*, ***11***, 31–46. https://doi.org/10.1038/nrg2626

Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A., Johnson, E.A., 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, ***17***, 240–248. https://doi.org/10.1101/gr.5681207.high-throughput

Mitter, C., Farrell, B., Wiegmann, B., 1988. The phylogenetic study of adaptive zones: Has phytophagy promoted insect diversification? *The American Naturalist*, ***132***, 107–128. https://doi.org/10.1086/284840

Moreau, C.S., Bell, C.D., Vila, R., Archibald, S.B., Pierce, N.E., 2006. Phylogeny of the Ants: Diversification in the age of angiosperms. *Science*, ***312***, 101–104.

Pennington, R.T., Hughes, M., Moonlight, P.W., 2015. The origins of tropical rainforest hyperdiversity. *Trends in Plant Science*, ***20***, 693–695. https://doi.org/10.1016/j.tplants.2015.10.005

Pimm, S.L., Raven, P.H., 2017. The fate of the world's plants. *Trends in Ecology and Evolution*, ***32***, 317–320. https://doi.org/10.1016/j.tree.2017.02.014

Qin, Y., Schneider, T.M., Brenner, M.P., 2012. Sequencing by hybridization of long targets. *PLoS ONE*, ***7***, 5–10. https://doi.org/10.1371/journal.pone.0035819

Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., Gu, Y., 2012. A tale of 3 NGS sequencing platforms. https://doi.org/10.1186/1471-2164-13-341

Rasplus, J., Rodriguez, L.J., Tollon-cordet, C., 2018. Revisiting the phylogeny of *Ficus* (Moraceae): When next generation sequencing corroborates past generation botanists. *Submitted paper.*

Renoult, J.P., Kjellberg, F., Grout, C., Santoni, S., Khadari, B., 2009. Cyto-

nuclear discordance in the phylogeny of *Ficus* section *Galoglychia* and host shifts in plant-pollinator associations. *BMC Evolutionary Biology*, ***9***, 248. https://doi.org/10.1186/1471-2148-9-248

Richardson, E.J., 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. https://doi.org/10.1126/science.1061421

Ronholm, J., Nasheri, N., Petronella, N., Pagotto, F., 2016. Navigating microbiological food safety in the era of whole-genome sequencing. *Clinical Microbiology Reviews*, ***29***, 837–857. https://doi.org/10.1128/CMR.00056-16

Rønsted, N., Weiblen, G.D., Cook, J.M., Salamin, N., Machado, C.A., Savolainen, V., 2005. 60 million years of co-divergence in the fig-wasp symbiosis. *Proceedings of the Royal Society B: Biological Sciences*, ***272***, 2593–2599. https://doi.org/10.1098/rspb.2005.3249

Rønsted, N., Salvo, G., Savolainen, V., 2007a. Biogeographical and phylogenetic origins of African fig species (*Ficus* section *Galoglychia*). *Molecular Phylogenetics and Evolution*, ***43***, 190–201. https://doi.org/10.1016/j.ympev.2006.12.010

Rønsted, N., Yektaei-Karin, E., Truk, K., Clarkson, J.J., Chase, M.W., 2007b. Species-level phylogenetics of large genera: prospects of studying coevolution and polyploidy, in: Reconstructing the Tree of Life: Taxonomy and systematics of species rich taxa. Systematic Association Series, pp. 129–147.

Rønsted, N., Weiblen, G.D., Clement, W.L., Zerega, N.J.C., Savolainen, V., 2008a. Reconstructing the phylogeny of figs (*Ficus*, Moraceae) to reveal the history of the fig pollination mutualism. *Symbiosis*, ***45***, 1–12.

Rønsted, N., Weiblen, G.D., Savolainen, V., Cook, J.M., 2008b. Phylogeny, biogeography, and ecology of *Ficus* section *Malvanthera* (Moraceae). *Molecular phylogenetics and evolution*, ***48***, 12–22. https://doi.org/10.1016/j.ympev.2008.04.005

Roy, S., Tyagi, A., Shukla, V., Kumar, A., Singh, U.M., Chaudhary, L.B., Datt, B., Bag, S.K., Singh, P.K., Nair, N.K., Husain, T., Tuli, R., 2010. Universal plant DNA barcode loci may not work in complex groups: a case study with Indian Berberis species. *PloS one*, ***5***, e13674. https://doi.org/10.1371/journal.pone.0013674

Ruhfel, B.R., Gitzendanner, M. a, Soltis, P.S., Soltis, D.E., Burleigh, J.G., 2014. From algae to angiosperms-inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC evolutionary biology*, ***14***, 23. https://doi.org/10.1186/1471-2148-14-23

Savolainen, V., Chase, M.W., 2003. A decade of progress in plant molecular phylogenetics. *Trends in Genetics*, ***19***, 717–724. https://doi.org/10.1016/j.tig.2003.10.003

Schneider, H., Schuettpelz, E., Pryer, K.M., Cranfill, R., Magallo, S., Lupia, R., 2004. Ferns diversified in the shadow of angiosperms. *Global Biogeochemical Cycles*, ***428***, 553–557. https://doi.org/10.1029/2001GB001442

Silvestro, D., Zizka, G., Schulte, K., 2014. Disentangling the effects of key innovations on the diversification of Bromelioideae (Bromeliaceae). *Evolution*, ***68***, 163–175. https://doi.org/10.1111/evo.12236

Silvestro, D., Cascales-Miñana, B., Bacon, C.D., Antonelli, A., 2015. Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. *New Phytologist*, ***207***, 425–436. https://doi.org/10.1111/nph.13247

Soltis, D.E., Soltis, P.S., Doyle, J.J., 1998. Molecular systematics of plants, II: DNA sequencing, 2nd ed, Molecular systematics of plants, II: DNA sequencing. Kluwer Academic Publishers.

Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, W.J., Nixon, K.C., Farris, J.S., 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. *Botanical Journal of the Linnean Society*, ***133***, 381–461.

Soltis, D.E., Bell, C.D., Kim, S., Soltis, P.S., 2008. Origin and early evolution of angiosperms. *Annals of the New York Academy of Sciences*, ***1133***, 3–25. https://doi.org/10.1196/annals.1438.005

Soltis, P.S., Soltis, D.E., 2009. The role of hybridization in plant speciation. https://doi.org/10.1146/annurev.arplant.043008.092039

Soltis, D.E., Smith, S.A., Cellinese, N., Wurdack, K.J., Tank, D.C., Brockington, S.F., Refulio-Rodriguez, N.F., Walker, J.B., Moore, M.J., Carlsward, B.S., Bell, C.D., Latvis, M., Crawley, S., Black, C., Diouf, D., Xi, Z., Rushworth, C.A., Gitzendanner, M.A., Sytsma, K.J., Qiu, Y.L., Hilu, K.W., Davis, C.C., Sanderson, M.J., Beaman, R.S., Olmstead, R.G., Judd, W.S., Donoghue,

M.J., Soltis, P.S., 2011. Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany*, ***98***, 704–730. https://doi.org/10.3732/ajb.1000404

Soltis, D.E., Gitzendanner, M.A., Stull, G., Chester, M., Chanderbali, A., Jordon-Thaden, I., Chamala, S., Jordon-Thaden, I., Soltis, P.S., Schnable, P.S., Brad Barbazuk, W., 2013. The potential of genomics in plant systematics. *Taxon*, ***62***, 886–898. https://doi.org/10.12705/625.13

Stebbins, G.L., 1974. Flowering Plants: Evolution above the species level. Harvard University Press, Cambridge, MA, USA.

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***, 349–64. https://doi.org/10.3732/ajb.1100335

Tennessen, J.A., Govindarajulu, R., Liston, A., Ashman, T., 2013. Targeted sequence capture provides insight into genome structure and genetics of male sterility in a gynodioecious diploid strawberry, *Fragaria vesca* ssp. *bracteata* (Rosaceae) ***3***, 1341–1351. https://doi.org/10.1534/g3.113.006288

Unamba, C.I.N., Nag, A., Sharma, R.K., 2015. Next generation sequencing technologies: The doorway to the unexplored genomics of non-model plants. *Frontiers in plant science*, ***6***. https://doi.org/10.3389/fpls.2015.01074

Vamosi, J.C., Magallón, S., Mayrose, I., Otto, S.P., Sauquet, H., 2018. Macroevolutionary patterns of flowering plant speciation and extinction. *Annual Review of Plant Biology*, ***69***, 685–706. https://doi.org/10.1146/annurev-arplant-042817-040348

Wallace, A.R., 1878. Tropical Nature, and Other Essays. Macmillian, London, UK. Wang,

Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., Davis, C.C., Latvis, M., Manchester, S.R., Soltis, D.E., 2009. Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proceedings of the National Academy of Sciences*, ***106***, 3853–3858. https://doi.org/10.1073/pnas.0813376106

Wang, W., Lin, L., Xiang, X.G., Ortiz, R.D.C., Liu, Y., Xiang, K.L., Yu, S.X., Xing, Y.W., Chen, Z.D., 2016. The rise of angiosperm-dominated herbaceous floras: Insights from Ranunculaceae. *Scientific Reports*, ***6***, 6–13.

https://doi.org/10.1038/srep27259

Weiblen, G.D., 2000. Phylogenetic relationship of functionally dioecious *Ficus* (Moraceae) based on ribosomal DNA sequences and morphology. *American Journal of Botany*, ***87***, 1342–1357. https://doi.org/10.2307/2656726

Wilf, P., Labandeira, C.C., Kress, W.J., Staines, C.L., Windsor, D.M., Allen, A.L., Johnson, K.R., 2000. Latest cretaceous to cecent timing the radiations of Leaf Beetles: Hispines on gingers from latest cretaceous to recent. *Science*, ***289***, 291–295. https://doi.org/10.1126/science.289.5477.291

Wilkinson, M., Roelants, K., Gower, D.J., Wilkinson, M., Loader, S.P., Biju, S.D., Guillaume, K., Moriau, L., Bossuyt, F., 2007. Global patterns of diversification in the history of modern amphibians. *Proceedings of the National Academy of Sciences*, ***104***, 887–892. https://doi.org/10.1073/pnas.0608378104

Xing, Y., Onstein, R.E., Carter, R.J., Stadler, T., Linder, H.P., 2014. Fossils and a large molecular phylogeny show that the evolution of species richness, generic diversity, and turnover rates are disconnected. *Evolution*, ***68***, 2821–2832. https://doi.org/10.1111/evo.12489

Xu, L., Harrison, R.D., Yang, P., Yang, D.R., 2011. New insight into the phylogenetic and biogeographic history of genus *Ficus*: Vicariance played a relatively minor role compared with ecological opportunity and dispersal. *Journal of Systematics and Evolution*, ***49***, 546–557. https://doi.org/10.1111/j.1759-6831.2011.00155.x

Zerega, N.J.C., Clement, W.L., Datwyler, S.L., Weiblen, G.D., 2005. Biogeography and divergence times in the mulberry family (Moraceae). *Molecular Phylogenetics and Evolution*, ***37***, 402–416. https://doi.org/10.1016/j.ympev.2005.07.004

Zhang, Q., Onstein, R.E., Little, S.A., Sauquet, H., 2018. Estimating divergence times and ancestral breeding systems in *Ficus* and Moraceae. *Annals of Botany*, ***123***, 191–204. https://doi.org/10.1093/aob/mcy159

Chapter I

Evolution and classification of figs (*Ficus*) and their close relatives (Castilleae) united by involucral bracts

(Submitted)

Examples of inflorescences partially enclosed in Castilleae (right) and cross section of enclosed *Ficus* (left)

The Linnean Society of London
BOTANICAL Journal of the Linnean Society

Evolution and classification of figs (*Ficus*) and their close relatives (Castilleae) united by involucral bracts

Journal:	*Botanical Journal of the Linnean Society*
Manuscript ID	Draft
Manuscript Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Clement, Wendy ; College of New Jersey, Biology Bruun Lund, Sam; University of Copenhagen, The Natural History Museum of Denmark Cohen, Alanna; Rutgers University, Department of Biochemistry and Microbiology Kjellberg, Finn; University of Montpellier Weiblen, George; University of Minnesota Twin Cities, Plant Biology Rønsted, Nina ; University of Copenhagen, The Natural History Museum of Denmark
Keywords:	classifications < Systematics, phylogenetic relationships < Systematics, Morphology, ITS < Genetics, molecular phylogenetics < Genetics, lineage history < Systematics, infrageneric classification < Taxonomy

Evolution and classification of figs (*Ficus*) and their close relatives (Castilleae) united by involucral bracts

Short running title: Evolution of figs and Castilleae

WENDY L. CLEMENT[1], SAM BRUUN-LUND[2], ALANNA COHEN[1,3]; FINN KJELLBERG[4]; GEORGE D. WEIBLEN[5], NINA RØNSTED[2]*

[1]*Department of Biology, The College of New Jersey, Ewing, NJ USA.*

[2]*The Natural History Museum of Denmark, University of Copenhagen, Denmark.*

[3]*Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ USA.*

[4]*CEFE, CNRS, University of Montpellier, University Paul Valéry Montpellier, EPHE, IRD, Montpellier, France.*

[5]*Department of Plant Biology, University of Minnesota, MN USA.*

*Corresponding author: nronsted@snm.ku.dk

ABSTRACT

Figs and fig wasps are a classic example of an obligate pollination mutualism. Decades of work untangling the ecology and evolution of these organisms has simultaneously contributed to development of the fields of mutualism, coevolution, and plant-insect interactions at large. With more than 800 species, figs (*Ficus*, Moraceae) are among some of the larger genera of angiosperms. Phylogenetic studies of Moraceae have supported the clade Castilleae as the sister lineage of figs. Compared to figs, Castilleae has many fewer species (60 species and 11 genera), suggesting changes in rates of diversification along these two branches. Relatively little is known about Castilleae compared to *Ficus*, and we argue that defining the clade comprising figs and Castilleae, hereafter Involucraoideae, focuses attention on opportunities for comparative studies of pollination mutualisms and diversification rates. In this study, we define Involucraoideae and propose a revised classification scheme that accounts for the phylogenetic reconstruction based on the most comprehensive sampling of this group to date. Moving forward, this classification will better guide and support evolutionary, ecological, and comparative pollination biology studies of this most notable group.

ADDITIONAL KEYWORDS: Castilleae – classification – external transcribed spacer – parology – *Ficus* – involucral bracts – Involucraoideae – morphology – phylogenetic reconstruction

INTRODUCTION

With at least 800 named species, figs account for more than half of the species diversity of the mulberry family, Moraceae (ca. 1,100 species; Clement & Weiblen, 2009). Phylogenetic analyses of Moraceae have strongly supported Castilleae C.C. Berg as sister to the figs (*Ficus* L.) based on plastid (Datwyler & Weiblen, 2004), nuclear (Zerega *et al.*, 2005), and morphological data (Clement & Weiblen, 2009). Figs have been central to advancing study of pollination mutualisms, coevolution, and cospeciation (Bronstein, 1988; Herre, 1989; Herre & West, 1997; Lopez-Vaamonde *et al.*, 2001; Weiblen, 2001; Weiblen *et al.*, 2001; Weiblen & Bush, 2002; Cook &Rasplus, 2003; Jousselin *et al.*, 2003; Weiblen, 2004; Machado *et al.*, 2005; Rønsted *et al.*, 2005; Marussich & Machado, 2007; Silveus *et al.*, 2007; Jackson *et al.*, 2008; Jousselin *et al.*, 2008; Herre *et al.*, 2008; Cruaud *et al.*, 2012a; Cruaud *et al.*, 2012b; McLeish & van Noort, 2012; Conchou *et al.*, 2014; Bain *et al.*, 2016; Rodriguez *et al.*, 2017). Figs occur in tropical and subtropical regions worldwide and include trees, hemi-epiphytes, epiphytes, shrubs, climbers, rheophytes, and lithophytes. By comparison, Castilleae are a group of 11 genera and 60 species of trees and shrubs with four species distributed in the Paleotropics and 56 species in the Neotropics. Figs and Castilleae diverged from one another at least 65 Ma (Zerega *et al.*, 2005) and the striking difference in contemporary species richness suggests differing rates of diversification.

Together, figs and Castilleae differ from other Moraceae in having involucral bracts that subtend the inflorescences on a disc or urn shaped receptacle. In the case of Castilleae, the involucral bracts do not completely enclose the inflorescence as in figs. The positioning of these bracts has profound implications for their reproductive ecology. In the case of figs, the involucral bracts form a tight pore, or ostiole, at the apex of the receptacle. Mated pollinating wasps force themselves through this opening into the cavity of the fig

(syconium) where they pollinate flowers, lay eggs, and usually die. Pollinator offspring emerge from galls inside the fig to mate and collect pollen from staminate flowers before exiting in search of other receptive figs. In contrast to the 'cradle to grave' relationship between figs and their pollinating wasps, Castilleae inflorescences are only partially enclosed by involucral bracts thereby allowing pollinators to come and go. From the limited study of Castilleae pollination, both wind (Osmaston, 1965; Croat, 1978) and insect (Sakai *et al.* 2000; Zerega *et al.*, 2004) pollination syndromes are present. As in figs, insect pollinated Castilleae are also involved in brood site pollination mutualisms where pollinators mate and lays eggs in the inflorescences. Pollination by thrips has been documented for two Castilleae species, *Antiaropsis decipiens* K. Schum., endemic to New Guinea (Zerega *et al.*, 2004) and *Castilla elastica* Sess., widespread in the Neotropics (Sakai *et al.*, 2000).

Comparative study of figs and Castilleae can offer insight on the evolution of morphological and molecular diversity, pollination ecology, diversification rates, and historical dispersal patterns. However, aside from family level phylogenetic studies (Datwyler & Weiblen, 2004; Zerega *et al.*, 2005; Clement & Weiblen, 2009), figs and Castilleae have seldom been the subject of comparative work (Clement, 2008; Moe *et al.*, 2012). Comparing Castilleae and fig pollination syndromes, Moe *et al.* (2012) hypothesized that the nature of the pollinator reward and the number of floral visits by a pollinator may account for the difference in diversification in these two lineages. For instance, fig wasp offspring develop in galled or fertilized fig ovules. When wasp offspring fare better in pollinated flowers, pollination can increase wasp fitness, and the fig can furthermore reduce pollen production to the benefit of pollinator production. Thrips pollinated Castilleae do not depend on successful pollination as thrips eat pollen and mate on male inflorescences.

Selective pressure on host choice also differs among fig and Castilleae pollination syndromes. In many species, foundress fig wasps lose their wings and antennae on entering a fig so that they cannot reach another tree which likely results in intense selection to discern host quality before host selection. Castilleae pollinators visit multiple inflorescences per generation with little consequence for visiting a non-rewarding inflorescence. Differing selective pressures resulting from the nature of these pollination interactions may have impacted the evolutionary trajectory of both lineages. Further testing of this hypothesis requires additional study of Castilleae pollination biology and an improved phylogenetic framework for figs and Castilleae.

Our current understanding of *Ficus* classification is largely based on a massive Malesian revision of *Ficus* initiated by E. J. H. Corner completed by C. C. Berg after Corner's death (Berg 2003a-e; 2004a-b; Berg & Corner, 2005) building on earlier work (summarized in Corner, 1965). Berg's classification based on morphological and anatomical characters added emphasis on vegetative characters compared to Corner's treatments that focused on flower and fruit characters (Corner, 1965). Ultimately, Berg and Corner (2005) subdivided *Ficus* into six subgenera: (1) *Pharmacosycea* (Miq.) Miq. (monoecious), (2) *Urostigma* (Gasp.) Miq. (monoecious) (3) *Ficus* Corner (gyno-dioecious), (4) *Sycidium* (Miq.) Mildbr. & Burret (gyno-dioecious), (5) *Synoecia* (Miq.) Miq. (gyno-dioecious), and (6) *Sycomorus* (Gasp.) Miq. (gyno-dioecious and monoecious). Subgenera *Pharmacosycea*, *Sycidium, Sycomorus* and *Urostigma* are distributed from the Pacific to West Africa and subgenera *Pharmacosycea* and *Urostigma* additionally include a distinct Neotropical section. Subgenera *Ficus* and *Synoecia* are almost exclusively restricted to the Malesian region and Mainland Asia (Berg, 2003a).

The most recent comprehensive molecular phylogenetic analysis of 200 species of figs supported the monophyly of subgenera *Sycidium, Sycomorus,* and *Synoecia*, while subgenera *Ficus, Pharmacosycea*, and *Urostigma* were paraphyletic (Cruaud *et al.*, 2012b) concurring with prior work on fig phylogenetic trees (Weiblen, 2000; Jousselin *et al.*, 2003; Rønsted *et al.*, 2005; Rønsted *et al.*, 2008a; Xu *et al.*, 2011). While many sections and subsections within these subgenera were not monophyletic, several supported clades do broadly correspond to published sections (*Adenosperma* Corner*, Americanae* Miq.*, Eriosycea* Miq.*, Galoglychia* Gasp.*, Oreosycea* (Miq.) Miq., *Pharmacosycea* (Miq.) Benth.& Hook.f*, Sycocarpus* Miq.*, Sycomorus* (Gasp.) Miq) and subsections (*Conosycea* (Miq.) C.C. Berg*, Ficus* Corner*, Frutescentiae* Sata*, Malvanthera* (Corner) C.C. Berg*, Urostigma* (Gasp.) C.C. Berg) (Berg & Corner, 2005; Rønsted *et al.*, 2008a). Given that phylogenetic evidence only partly supports previous taxonomic treatments based on morphology, there is much potential for confusion.

Relationships along the backbone of the fig phylogenetic tree remain unsupported and conflicts between ribosomal DNA and low copy nuclear gene trees for *Ficus* are not resolved (Cruaud *et al.*, 2012b; Harrison *et al.*, 2012). Further, a recent phylogenetic reconstruction from whole plastids representing 59 species of *Ficus* (Bruun-Lund *et al.*, 2016) provided strong support for relationships deep in the *Ficus* phylogenetic tree. However, a number of conflicts were identified and await increased resolution and clade support of phylogenies reconstructed from nuclear gene regions for further investigation.

Similar to *Ficus*, the current classification of Castilleae is primarily based on morphology. Castilleae are trees, generally diagnosed by unisexual inflorescences with discoid to cup-shaped receptacles, bracts subtending the inflorescence (involucre), large seeds, septate wood fibers, and the lack of cystoliths. Molecular phylogenetic analysis of

plastid (*ndhF*; Datwyler & Weiblen, 2004) and nuclear (*26S*; Zerega *et al.*, 2005) sequence data in addition to morphology (Clement & Weiblen, 2009) supported the unity of Castilleae, including *Antiaropsis* K. Schum*, Poulsenia* Eggers*,* and *Sparattosyce* Bureau (formerly part of tribe Artocarpeae, breadfruit and relatives) plus all eight genera of neotropical Castilleae (Datwyler & Weiblen, 2004). Morphological analysis of the tribe further supported two subtribes, Antiaropsineae, comprising *Antiaropsis* and *Sparattosyce*, and Castillineae, including the remaining nine genera (Clement & Weiblen, 2009). As Castilleae has only been treated in the context of Moraceae, revision of Castilleae classification awaits molecular phylogenetic study.

To facilitate further comparative work among figs and Castilleae, we present an improved phylogenetic framework for both clades. First, we propose the name Involucraoideae to recognize the well-supported clade including Castilleae and *Ficus*. This name reflects a key morphological feature shared between the two lineages – involucral bracts. Next, we present a molecular phylogenetic tree of 307 *Ficus* species and 43 Castilleae – the most robust species sampling of the group to date. Finally, using the current classification of *Ficus* and Castilleae based on morphology (Berg, 1977; Berg & Corner, 2005; Berg *et al.*, 2006), we use the phylogenetic tree reconstructed here as a framework to suggest revisions to the classification of Involucraoideae that now reflect evolutionary relationships. The subfamily Involucraoideae is circumscribed by the tribe Castilleae and the monotypic tribe Ficeae.

MATERIALS AND METHODS

TAXON SAMPLING

To assess the current classification and describe the evolutionary relationships of *Ficus* and Castilleae, we assembled the most comprehensive data matrix to date, sampling representatives of all 11 Castilleae genera and more than 40% of 800 named *Ficus* species. Data were assembled in two matrices. The first data matrix focused on phylogenetic reconstruction of Involucraoideae and included 133 taxa. Taxon sampling included 94 fig species representing 2-3 species per major clade (Cruaud *et al.*, 2012b), 39 species of Castilleae representing all 11 genera, and *Prainea* King ex Hook.f. (Artocarpeae, Moraceae) as an outgroup. This data set included three gene regions: the internal transcribed spacer region of nuclear ribosomal DNA (ITS), glyceraldehyde 3-phosphate dehydrogenase (*G3pdh*), and granule bound starch synthase (*GBSSI*; Suppl. Table S1). The second matrix focused on *Ficus* and included 307 fig species adding more than 100 species to the most recent comprehensive phylogenetic sample (Cruaud *et al.*, 2012b). We designated *Antiaropsis decipiens, Castilla elastica, Poulsenia armata* (Miq.) Standl. and *Sparattosyce dioica* Bureau as outgroups to root the phylogenetic tree. This data set included six gene regions: ITS, external transcribed spacer region (ETS), and four low copy nuclear gene regions: *G3pdH*, *GBSSI*, glutamine synthase (*ncpGS*), and for the first time in *Ficus*, Mg-protoporphyrim monomethyl ester cyclase (*At103*; Suppl. Table S1).

Leaf material for sequencing newly added species was obtained from herbaria (A, AAU, F, HON, HUH, K, LAE, MIN, MO, PUH, UNAM), living collections (BG, BR, C, HITBC, K, NBG, REU), and recent field collections (Suppl. Table S1). New data (>400 = 34 % of analysed sequences) were combined with data from prior phylogenetic work on Moraceae (Azuma *et al.*, 2010; Chantarasuwan *et al.*, 2015; Cruaud *et al.*, 2012b; Harrison *et al.*, 2012; Jackson *et al.*, 2008; Jousselin *et al.*, 2003; Kusumi *et al.*, 2012; Machado *et al.*, 2005; Mcleish *et al.*, 2011; Renoult *et al.*, 2009; Rønsted *et al.*, 2005, 2008a, 2008b;

Silvieus *et al.*, 2007; Weiblen, 2000; Xu *et al.*, 2011). Genbank accessions for all taxa are available in Supplementary Table S1.

DNA Extraction, Amplification, and Sequencing

Total genomic DNA was extracted from 15-30 mg of dried leaf-fragments or herbarium material following Rønsted *et al.* (2008a). Amplification of ITS, ETS, G3pdh, ncpGS and GBSSI for all *Ficus* species were performed following Cruaud *et al.* (2012b) and references therein. Amplification of At103 followed protocols by Li *et al.* (2008). Amplification primers are listed in Supplementary Table S2.

ITS, *G3pdh* and *GBSSI* for Castilleae genera were amplified in a 25 μL reaction using 1x *TaKaRa Ex Taq* buffer (2mM $MgCl_2$; Otsu, Shiga, Japan), 0.2 mM each dNTP, 10 μM Bovine Serum Albumin (BSA), 12-25 μM forward and reverse primers (Suppl. Table S2), 1.25 U *TaKaRa Ex Taq* DNA polymerase, and ~20 ng of genomic DNA. In instances when ITS amplification was not successful, a nested PCR approach was used by first amplifying a larger region encompassing ITS with 25 μM of external primers 17SE and 26SE (Sun *et al.*, 1994), followed by a second PCR using 1 μL of the previous PCR product, and 25 μM of both ITS4 and ITS5. Thermal cycler conditions for all ITS amplifications were: 94ºC for 2 min, 25 cycles of 94ºC for 1 min, 50ºC for 1 min, 70ºC for 2 min, followed by 72ºC for 7 min. Thermal cycler conditions for *G3pdh* were: 95ºC for 3 min 30 sec, 35 cycles of 95ºC for 1 min, 49ºC for 1 min, 70ºC for 2 min, followed by 72ºC for 7 min. Thermal cycler conditions for *GBSSI* followed a "stepdown" protocol modified from Evans *et al.* (2000) as follows: 94ºC for 3 min, 2 cycles of 94ºC for 1 min, 58ºC for 1 min, 72ºC for 2 min, 2 cycles of 94ºC for 1 min, 54ºC for 2 min, 72ºC for 2 min, 2 cycles of 94ºC for 1 min, 50ºC for 1 min, 72ºC for 2 min, and 24 cycles of 94ºC for 1 min, 48ºC for 2 min, 72ºC for 2 min,

followed by 72ºC for 20 min. PCR products were column purified using a Qiagen PCR cleanup kit (Qiagen, Valencia, California, USA) and quantified using a Turner Quantech Fluorometer (Barnstead-Thermolyne, Dubuque, Iowa, USA) using Hoecsht 33258 dye prior to sequencing.

All ITS, ETS, *G3pdh, ncpGS* and *At103* PCR products were directly sequenced. *GBSSI* and ITS amplicons showing signs of divergent alleles in direct sequencing were cloned prior to sequencing using either a TOPO-TA (Invitrogen, Carlsbad, CA) or Stratagene PCR cloning kit (Agilent Technologies, Santa Clara, California, USA) following manufacturer protocols. Transformed bacteria were grown overnight on LB + ampicillin agar plates at 37ºC. Eight to ten colonies per PCR product were screened using PCR for insert size. Three positive clones per accession were grown in LB + ampicillin broth overnight at 37ºC and plasmids were isolated using Qiagen Plasmid Isolation kit (Qiagen, Valencia, California, USA). In other cases, the gene region of interest was cleaned directly from the clone screen PCR using a Qiagen PCR cleanup kit.

Previously published ETS gene trees in *Ficus* have been in conflict with other nuclear genes, as the ETS tree failed to recover a monophyletic subgenus *Sycomorus* because section *Sycocarpus* formed a separate clade sister to subgenus *Urostigma* (excluding subsection *Urostigma*) (e.g. Rønsted *et al.*, 2008a). Multiple copies of ETS within *Ficus* have been suspected (Cruaud, personal communication and NR personal observations) and potential problems with ETS parology have been reported (Calonje *et al.*, 2009). We explored the problem in *Ficus* by resampling species from clades in conflict and not in conflict among the ETS and other gene trees. Our sampling included: section *Sycocarpus* (*F. condensa* King, *F. fistulosa* Reinw. ex Blume*, F. hispida* Blanco*,* and *F. scortechinii* King), section *Adenosperma* (*F. ochrochlora* Ridl., *F. pseudopalma* Blanco,

and *F. itoana* Diels), and section *Sycomorus* (*F. sur* Forssk., *F. sycomorus* L., and *F. vallis-choudae* Delile) covering subgenus *Sycomorus* as well as subsection *Conosycea* (*F. drupacea* Thunb.), and subsection *Urostigma* (*F. lacor* Buch.-Ham). In an effort to capture a greater proportion of ETS paralogs potentially present, we relaxed PCR conditions by lowering the annealing temperature from 49°C to 45°C, increasing the number of cycles from 25 to 40, and extending the duration of the premelt from 2 min 30 sec to 4 min. We also designed and used a *Ficus* specific primer (ETS-Fic1, Suppl. Table S2), and cloned all PCR products. We column purified and sequenced 6-9 clones per accession (except for *F. hispida* in which only three amplicons were recovered).

Sequencing for all cleaned PCR products was performed using Big Dye 3.1 sequencing reagents and protocols (Applied Biosystems, Foster City, CA, USA). Sequencing reactions were performed in 10 μL reactions with 20 ng of PCR product or 200 ng of isolated plasmids. Sequencing primers for each gene region are listed in Supplementary Table S2. Products were visualized and data were collected on an ABI 377 automated DNA sequencer (Applied Biosystems). Sequences were assembled using Sequencher 4.6 (Gene Codes Corp., Ann Arbor, Michigan, USA) or Geneious v. R6-7 (www.biomatters.com). Individual gene regions within each data set were first aligned using MAFFT (Katoh & Standly, 2013) and manually inspected.

Phylogenetic Analyses

Trees for each gene region were reconstructed using maximum likelihood and Bayesian inference for Involucraoideae and *Ficus*. Prior to analysis, the best fitting model of sequence evolution was determined using jModeltest v. 2.1.4. (Darriba *et al.*, 2012) following the AIC criterion (Posada & Buckley, 2004). In the Involucraoideae dataset,

TIM3+G, TVM+I+G, and TIM2+I+G was selected for *G3pdh*, ITS, and *GBSSI* respectively. For *Ficus*, a GTR+G model of sequence evolution was selected for ITS, ETS and *G3pdh*, and TIM2+G, TPM2uf+G, TPM3uf+I+G were selected for *ncpGS, GBSSI*, and *At103* respectively. Maximum likelihood analyses were performed in Garli v 2.01.167 (Zwickl, 2006) and repeated five times, each time using a random starting tree and allowing model parameters to be estimated. Support was assessed using 500 bootstrap replicates in Garli (Zwickl, 2006). As these models are nested within the general time reversible model, all matrices were analysed with a GTR+G model for Bayesian analyses. Bayesian analyses were run with MrBayes v. 3.2.1 (Huelsenbeck & Ronquist, 2001) for 30 million generations. Stationarity was assessed using the Trace option in Geneious v R7 (Biomatters, Ltd) and with Tracer v 1.5 (Rambaut, 2007), and the first 25% of trees sampled in the posterior distribution were removed as burnin.

Before concatenation in a combined analysis, gene trees were visually inspected and compared for supported (using bootstrap and posterior probabilities) topological congruence. Using PartitionFinder (Lanfear *et al.*, 2012), we determined the best partitioning strategy and models of sequence evolution for the combined datasets. The combined analyses of the *Ficus* and *Involucraoideae* datasets were conducted using the same analysis protocols as described for individual gene regions.

RESULTS

CONGRUENCE OF INVOLUCRAOIDEAE GENE TREES

ML and Bayesian analyses recovered similar topologies but with different levels of clade support. Bayesian analyses often had higher support for relationships as compared to ML bootstrap analyses (Fig. 1, TreeBase accession S24008). Here, we recovered congruent

relationships among the gene trees with one exception. Subsection *Urostigma* was recovered as monophyletic in the ITS gene tree (bootstrap [BS]=97, posterior probability [PP]=1) but not the *G3pdh* gene tree (subsection *Urostigma* was not fully sampled in GBSSI gene tree; Fig. 1, TreeBase accession S24008). As the dedicated analysis of *Ficus* offered an expanded sampling of this clade, a detailed description of relationships recovered in trees resulting from that analysis will be described in the *Ficus* gene tree section below.

With respect to the Castilleae clade with the Involucraoideae analyses, *Castilla*, *Helicostylis* Trécul, and *Maquira* Aubl. were recovered as monophyletic (Fig.1). *Antiaris* Lesch. and *Poulsenia* are monotypic and *Antiaropsis* and *Sparattosyce* were represented by one of the two species in the respective genera. *Naucleopsis* Miq. was recovered as monophyletic in *G3pdh* and ITS gene trees (Fig. 1, TreeBase accession S24008). However, two clades of *Naucleopsis* species were consistently recovered in all gene trees with one clade containing *N. glabra* Spruce*, N. krukovii* (Standl.) C.C. Berg*, N. ulei* (Warb.) Ducke*,* and *N. imitans* (Ducke) C.C. Berg and a second clade containing *N. caloneura* Ducke*, N. guianensis* (Mildbr.) C.C. Berg*, and N. ternstroemiiflora* (Mildbr.) C.C. Berg. *Perebea* Aubl. and *Pseudolmedia* Trécul were not consistently recovered among the gene trees. The paraphyly of *Perebea* was due to the exclusion of *Perebea mollis* (Poepp. & Endl.) Huber and *P. rubra* (Trécul) C.C. Berg*,* which formed a clade independent of other *Perebea* species (Fig. 1). The core *Perebea* clade often did not include *P. guianensis* Aubl. but there was little support for excluding it. *Pseudolmedia* was recovered as monophyletic in the GBSSI gene tree and two well-supported *Pseudolmedia* clades were recovered by ITS. These relationships differ as ITS suggested *P. laevis* (Ruiz & Pav.) J.F. Macbr. and *P. macrophylla* Trécul are sister taxa (BS=100, PP=1) while GBSSI placed *P. laevis* as sister to all of *Psuedolmedia* including *P. macrophylla* (BS=90, PP=1; Fig. 1, TreeBase accession

S24008). *G3pdh* did not recover a clade containing *Psedolmedia* as *P. laevigata* Trécul and *P. rigida* (Klotzsch & H.Karst.) Cuatrec. (which are well-supported sister taxa in all three gene trees) were more closely related to *Perebea mollis* and *P. rubra* (BS=73, PP=1; Fig. 1, TreeBase accession S24008).

Few well-supported relationships among Castilleae genera were recovered in the gene tree analyses. Neotropical taxa were supported as a clade only by ITS (BS=86, PP=1; Fig. 1, TreeBase accession S24008), and none of the gene trees recovered the relationship of the paleotropical to neotropical genera due to lack of resolution. ITS strongly supported a clade containing *Pseudolmedia*, *Perebea*, *Helicostylis* and *Maquira* (BS=91, PP=1; Fig. 1, TreeBase accession S24008) and GBSSI was unresolved for these nodes. The *G3pdh* gene tree conflicted with this clade; this gene tree recovered a clade of *Pseudolmedia*, *Perebea* and *Helicostylis* (BS=88, PP=1; Fig. 1, TreeBase accession S24008) to the exclusion of *Maquira*. Instead, *Maquira* was recovered as sister to *Naucleopsis* with moderate to strong support (BS=71, PP=0.98). Further, within the clade containing *Psedolmedia*, *Perebea*, and *Helicostylis*, the placement of *Pseudolmedia rigida* and *P. laevigata* (as described above) conflicted with both the ITS and *G3pdh* gene trees.

INVOLUCRAOIDEAE COMBINED ANALYSIS

Although there were supported conflicts when comparing the gene trees, many of these supported conflicts were only supported by the results of the Bayesian analysis and have low to moderate support in the ML bootstrap analysis. As such, we chose to combine our gene trees in a total evidence analysis knowing that more data will be needed in the future to resolve deeper relationships of the group.

Combining the ITS, *G3pdh*, and GBSSI data improved the resolution and clade support of the Involucraoideae phylogenetic tree (Fig. 1). All Castilleae genera, except *Perebea*, were strongly supported with high bootstrap support and high posterior probabilities (Fig. 1). *Perebea* was recovered paraphyletic as *P. mollis* and *P. rubra* formed a well-supported clade outside of *Perebea* (BS=82, PP=1; Fig. 1) similar to results of the gene tree analyses. *Antiaropsis decipiens* and *Sparattosyce dioica* were sister taxa (BS=99, PP=1; Fig. 1) and formed a clade sister to all other Castilleae (BS=86, PP=0.99; Fig. 1). *Antiaris toxicaria* was recovered as sister to *Mesogyne insignis* BS=96, PP=1; Fig. 1), and this clade was recovered as sister to the well-supported clade of Neotropical Castilleae (BS=94, PP=1; Fig. 1). Within the Neotropical Castilleae, *Poulsenia* was recovered as sister to all other Neotropical genera (BS=91, PP=1; Fig. 1). Here, *Maquira* was well supported as sister to *Helicostylis*, *Perebea* and *Pseudolmedia* (BS=78, PP=1; Fig. 1) similar to the ITS and GBSSI gene trees. Also, *Pseudolmedia laevigata* and *P. rigida* were recovered within a larger clade of *Pseudolmedia* as opposed to *Perebea rubra* and *P. mollis* as observed in the *G3pdh* gene tree.

Ficus Gene Tree Congruence

The final data set included 307 species of *Ficus*. Numbers of species sampled for each gene region were as follows: *At103* - 140, ETS - 244, ITS - 311, *G3pdh* - 209, *GBSSI* - 60, and *ncpGS* - 79. No strongly supported conflicts between individual dataset were recovered. Individual analysis of the *At103* region provided limited resolution and support but did not conflict with previous findings (phylogenetic reconstruction not shown).

Amplification success of the ETS region was improved considerably using the new *Ficus* specific primer ETS-Fic1 (Supplementary Table S2) resulting in the addition of 39 new sequences of the ETS region (Supplementary Table S1). The targeted sampling of ETS

using relaxed PCR conditions recovered two copies of the ETS region for several accessions from section *Sycocarpus* (*F. condensa, F. fistulosa, F. hispida* and *F. scortechinii*) and sect. *Adenosperma* (*F. adenosperma*). We found that the Hel1 primer used in previous studies preferentially amplified a paralogous copy of ETS for some taxa, which resulted in the polyphyly of subgenus *Sycomorus* recovered in previous studies. Using the new *Ficus* specific primer ETS-Fic1, we successfully amplified the presumably correct copy resulting in new sequences placing section *Sycocarpus* and all members of section *Adenosperma* with the remainder of subgenus *Sycomorus* as supported by ITS and other genes and morphology. Using the ETS-Fic1 primer (Supplementary Table S2), the new ETS data recovered a monophyletic subgenus *Sycomorus*. All ETS sequences of section *Sycocarpus* and *F. adenosperma* produced prior to this study that represent a paralogous copy were excluded from the data matrix prior to the final analysis.

Ficus Combined Analysis

The emerging picture of the phylogenetic tree of *Ficus* (Figures 2 and 3A-F) was largely consistent with sections or subsections proposed by morphology and provided a coherent global framework, although infrageneric relationships remain uncertain and many relationships were not well-supported. The extensive sampling in the present study allowed for interpretation of relationships of several taxa that have been difficult to place using morphology.

Three of the six subgenera (Berg & Corner, 2005), namely *Sycidium* (80% BP/PP=0.99), *Sycomorus* (97% BP/PP=1.00) and *Synoecia* (100% BP/PP=1.00), were monophyletic, whereas subgenera *Ficus, Pharmacosycea* and *Urostigma* were polyphyletic. The American section *Pharmacosycea* (100% BP/PP=1.00) was sister to the remainder of

Ficus (68% BS/PP=0.90) although this was not strongly supported. Relationships within the remainder of *Ficus* were not well resolved, but a number of clades were well supported. Section *Oreosycea* (Miq.) Miq. is divided between two clades consisting of the *Albipilae* species group (100 % BS/PP=1.00) and the remainder of section *Oreosycea* (77% BP/PP=1.00). Subgenus *Urostigma* is also split into a clade with subsection *Urostigma* (100% BP/PP=1.00) and a larger clade (100% BS/PP=1.00) including the remainder of the former subgenus *Urostigma.* Sections *Urostigma* (Gasp.) Endl. and *Stilpnophyllum* Endl. are polyphyletic. Subgenus *Ficus* is split into three clades corresponding to the *Ficus carica* group (100 % BS/PP=1.00), which is unplaced, and sections *Frutescentiae* Sata (92% BP/PP=1.00) and *Eriosycea* Miq. (100% BP/PP=1.00), which form a clade (98% BP/PP=1.00) together with subgenus *Synoecia* (Miq.) Miq. (100% BP/PP=1.00).

DISCUSSION

PHYLOGENETIC TREE OF INVOLUCRAOIDEAE

Here we present the name, Involucracoideae, to represent the clade containing *Ficus* and Castilleae. We present the first integrated systematic study of Involucraoideae as well as the first focused phylogenetic study of Castilleae. With striking variation in numbers of species, genetic diversity, and morphology, we discuss differences in historical biogeography, molecular evolution, and pollination ecology between figs and Castilleae to propose future research on evolutionary mechanisms driving the diversification of these two lineages.

The center of diversity of Castilleae is in the Neotropics, whereas the center of diversity of figs resides in the Paleotropics, specifically Borneo and New Guinea (Berg 2005b; Berg *et al.*, 2006). Our study of the phylogenetic tree of Castilleae strongly supports the

monophyly of Neotropical Castilleae, suggesting a single colonization event to the New World tropics. In contrast, figs likely colonized the Neotropics twice as phylogenetic studies of *Ficus* have recovered two well-supported clades of Neotropical figs which diversified at different points in evolutionary history (Jousselin *et al.*, 2003; Rønsted *et al.*, 2005; Rønsted *et al.*, 2008a; Cruaud et al., 2012b). Molecular phylogenetic analysis of figs has tentatively identified the Neotropical section *Pharmacosycea* as sister to all other lineages of *Ficus* (Herre *et al.*, 1996; Rønsted *et al.*, 2005; Rønsted *et al.*, 2008a; Cruaud *et al.*, 2012b, Bruun-Lund *et al.*, 2016; Zhang *et al.*, 2018) although the crown group of section *Pharmacosycea* diversified only 16 Ma ago and long after the origin of *Ficus* at least 75-48.5 Ma ago (Rønsted *et al.*, 2005; Zhang *et al.*, 2018). Estimates of the crown age of Castilleae (50-31.2 Ma) predate the diversification of the Neotropical figs (Rønsted *et al.*, 2005; Zerega *et al.*, 2005; Xu *et al.*, 2011; Cruaud *et al.*, 2012b; Zhang *et al.*, 2018). Differences in the number of colonization events and in the timing of diversification, seen in light of differences in historical climate and biogeographical events (e.g. the Andean uplift; Machado *et al.*, 2018) should inform our comparison of diversification rates between the two lineages.

Highly specific pollination mutualisms, like the fig-fig wasp interaction, have been hypothesized to increase rates of speciation (Stebbins, 1981), although studies in yuccas and yucca moths have shown the opposite (Smith *et al.*, 2008). Pollination syndromes of the sister group (Castillae) are worthy of consideration in terms of how they might influence speciation and extinction (Moe *et al.*, 2012; Sakai *et al.*, 2000; Zerega *et al.*, 2004). It remains unknown if thrips and Castilleae depend on each other for survival, as thrips may be able to breed elsewhere and Castilleae could receive pollen from other insects. Research dedicated to assessing the probability of extinction in the two lineages given their

pollination syndromes ought to examine the degree to which speciation and extinction rates are associated with diversification (Moe *et al.*, 2012).

If we consider the morphological evolution of figs and Castilleae as it relates to pollination biology, some of the traits associated with the fig-fig wasp pollination mutualism evolved in the ancestor of both figs and Castilleae (Clement & Weiblen, 2009). For instance, the appearance of an involucre, which is correlated with a shift from wind to insect pollination, occurred prior to the split between figs and Castilleae (Datwyler & Weiblen, 2004; Clement & Weiblen, 2009). Although the involucre is not exclusive to figs, tracking subsequent modifications of this trait are important to understanding the evolution of fig pollination where pollinators, born from the functional male figs, are part of the male investment of the plant (Anstett *et al.*, 1997). Comparisons of molecular evolutionary rates, morphologies, and pollination syndrome are needed to identify factors affecting rates of diversification.

PHYLOGENETIC TREE AND TAXONOMY OF CASTILLEAE

Strong support was recovered for the monophyly of the Neotropical taxa (Fig. 1) also recovered in prior phylogenetic studies of the family (Zerega *et al.*, 2005). Within this group, monotypic *Poulsenia* was recovered as sister to all other Neotropical Castilleae. *Poulsenia* has several unique characters that separate it from the rest of Castilleae including prickles and the loss of septate wood fibers (Berg, 2001).

Perebea was consistently recovered as paraphyletic in the individual and combined analysis (Fig. 1, TreeBase accession S24008). *Perebea* section *Noyera* (Trécul) Engler including *P. rubra* and *P. mollis* did not group with the rest of the genus. *Noyera* Trécul (1847) was first designated as a genus with the description of *Noyera rubra*. The genus was

later reduced to a section of *Perebea* (Engler, 1889) and also included *P. mollis*. Ducke (1922) reinstated *Noyera* including *N. mollis, N. rubra,* and later a third species, *N. glabrifolia* (Ducke, 1932). In 1972, *Noyera* was again reduced to a section of *Perebea* (Berg, 1972), and *P. rubra* was reduced to a subspecies of *P. mollis.* Later, *P. mollis* ssp. *rubra* was reinstated as *P. rubra*, and *P. glabrifolia* was reduced to *P. rubra* ssp. *glabrifolia* (Berg, 2001). Section *Noyera* differs from the rest of *Perebea* in having pluricellular globose capitate hairs on the lower leaf surface, filiform stigmas, and inner involucral bracts that are long and incurved prior to anthesis (Berg, 1972; Berg, 2001). Based on molecular evidence and these diagnostic features, we recommend reinstating the genus *Noyera* with *N. mollis* and *N. rubra* as the sole members.

An alternative taxonomic proposal would be to expand the circumscription of *Perebea* to encompass *Pseudolmedia*. However, *Pseudolmedia,* has recognizably distinct morphology that supports maintaining it as a genus for practical reasons. All species of *Pseudolmedia* are dioecious with uniflorous pistillate inflorescences (Berg, 1972; 1977a; 2001). Further, ITS and GBSSI gene trees support the monophyly of *Pseudolmedia*, but the *G3pdh* gene tree recovered a paraphyletic *Pseudolmedia*. While more data are needed to investigate this conflict among gene trees, the relationships recovered by the ITS and GBSSI gene trees, not *G3pdh*, are corroborated by morphology.

Our analysis supported the monophyly of *Helicostylis* and confirms the position of the morphologically distinct *H. tovarensis* (Klotzsch & H.Karst) C.C. Berg as sister to all other *Helicostylis* (Fig. 1). *Helicostylis tovarensis* differs from the rest of the genus on account of free rather than basally connate tepals in pistillate flowers, which are uniflorous rather than multiflorous, and one or two staminate inflorescences per leaf axil (Berg, 1972).

Although a combined analysis strongly supported the monophyly of all Castilleae genera except *Perebea* (and apart from the three monotypic genera – *Poulsenia*, *Antiaris*, *Mesogyne* Engl.), gene tree analysis of the Involucraoideae data set shed light on a number of conflicts. As the analysis was based on just two low copy nuclear genes and the internal transcribed spacer region of ribosomal DNA, there is much room for conflict among diverging gene trees. We speculate that the *G3pdh* gene tree is discordant with a Castilleae species tree based on nuclear ITS, GBSSI, 26S (Zerega *et al.*, 2005; 2010), plastid gene region *ndhF* (Datywler & Weiblen, 2004), and morphology. Specifically, the placement of *Maquira* and the monophyly of *Pseudolmedia* were called to question by *G3pdh* (Fig. 1). Other conflicts were observed but supported only by Bayesian posterior probabilities that seem to consistently over estimate branch support (Huelsenbeck *et al.*, 2002; Erixon *et al.*, 2003).

PHYLOGENETIC TREE AND TAXONOMY OF *FICUS*

Compared to the most recent comprehensive phylogenetic studies (Xu *et al.*, 2011; Cruaud *et al.*, 2012b), the present study increased taxon sampling by 42 species that were not included in any of the previous studies, introduced data from a gene region, AT103, new to the study of the fig phylogenetic tree, and reduced the amount of missing data in the matrix adding ca. 140 new *Ficus* sequences. The topology obtained from the At103 region was consistent with prior phylogenetic studies of figs (e.g., Cruaud *et al.*, 2012). Of the *Ficus* species included for the first time here (highlighted in bold font on figures 3A-F), most are placed in the same clades as their closest relatives predicted from their current classification *sensu* Berg & Corner (2005). The inclusion and verification of the placement

of these taxa in a comprehensive phylogenetic framework provides stronger evidence for the current circumscription of clades and infrageneric relationships of *Ficus*.

A number of taxa that have been difficult to classify based on their morphology were also included in this phylogenetic analysis of *Ficus* for the first time. For example, inclusion of additional taxa from subgenus *Sycidum* including *F. tsiangii* Corner as a second representative of the *Sinosycidium* group (section *Sinosycidium* Corner) helped to confidently identify four major subclades of subgenus *Sycidium*, namely *Palaeomorphe*, *Phaeopilosae*, *Sinosycidium*, and *Sycidium* (Fig. 3D). On the other hand, additional sampling of the *Oreosycea* and *Synoecia* clades highlighted the need for further revision of these groups as emerging subclades do not reflect the current morphological classification (Fig. 3A, 3B). Taxonomic implications of this most comprehensive phylogenetic framework are discussed below.

CURRENT CLADES TO GUIDE THE CLASSIFICATION OF *FICUS*

The comparison of morphology-based classification to phylogenetic reconstruction of evolutionary relationships among *Ficus* identifies taxonomic revisions that are needed to guide future evolutionary studies of the clade. We propose the recognition of a number of clades within *Ficus* that in some cases reinforce the classification of Berg & Corner (2005) and in other cases depart from it to provide clarity and precision when communicating about *Ficus* diversity. Each clade name is presented in conjunction with the closest Linnaean name and rank when possible for comparison to prior publications on *Ficus* classification.

Synoecia. This clade (Fig. 3A; 100% BS/PP=1.00) corresponds to the subgenus *Synoecia* (Miq.) Miq. including about 72 species of dioecious root climbers in Asia and Austral-Asia (Berg 2003d; Berg and Corner 2005). Berg & Corner (2005) subdivided

Synoecia into two sections *Rhizocladus* Endl. (primarily in New Guinea) and *Kissosycea* Miq. (primarily in Borneo), which are not clear-cut based on morphology. The sections are not resolved by the present molecular study. Notably there is a clade consisting of *F. sarmentosa* Buch.-Ham. Ex Sm. and *F. diversiformis* Miq. *Ficus sarmentosa* is traditionally considered a member of section *Rhizocladus*, but is a very variable species with affinities to the *Punctata* group of section *Kissosycea* (Berg & Corner, 2005). *Ficus diversiformis* is traditionally considered a member of the Malesian section *Kissosycea*, but it is one of only two species confined to the Asian Mainland (Berg & Corner, 2005). The other species, *F. hederacea* Roxb., has not been sequenced for this study. *Ficus pumila* L. is also a root climber traditionally included in section *Rhizocladus*, but previous studies (e.g. Rønsted, 2008a) have shown that *F. pumila* is more closely related to traditional subgenus *Ficus* species of section *Frutescentiae*, showing that the rootclimbing habit has evolved at least twice. A couple of other root climbers such as the essentially Sino-Himalayan *F. laevis* Desf. and *F. pubigera* (Wall. ex Miq.) Miq. also show affinities to subgenus *Ficus* members (Berg & Corner, 2005). *Ficus laevis* was not sequenced for this study, but *F. pubigera* is imbedded in section *Rhizocladus*.

Frutescentiae. This clade (Fig. 3A; 92% BS/PP=0.87) corresponds to section *Ficus* subsection *Frutescentiae* Sata and consists of 25-30 species including *F. pumila* and *F. iidaiana* Wilson, mostly from the Sino-Himalayan region and eight species from western Malesia. The *Frutescentiae* clade is closely related to the *Eriosycea* and *Synoecia* clades.

Eriosycea. This clade (Fig. 3A; 100% BS/PP=1.00) corresponds to section *Eriosycea* Miq. with about 34 species ranging from Sino-Himalaya to New Guinea. The *Eriosycea* and *Frutescentiae* clades are closely related to the *Synoecia* clade and together this group forms

a well-supported clade (Fig. 3A; 98% BS/PP=0.98), which has also been resolved in previous studies.

Sycidium. This clade (Figs. 2, 3B; 80% BS/PP=0.81) corresponds to subgenus *Sycidium* (Miq.) Berg & Corner and includes about 110 dioecious species primarily in Asia and Australasia both with about 10 species in Africa and Madagascar (Berg, 2003e; Berg & Corner, 2005). The *Sycidium* clade also largely corresponds to section *Sycidium* sensu Corner 1965, but excluding series *Pungentes* Corner (*F. minnahassae* (Teifjsm. & de Vriese) Miq. and *F. pungens* Reinw. ex Blume), which Berg transferred to subgenus *Sycomorus,* and including section *Sinosycidium* and series *Sinosyceae* (Berg, 2003e). Berg (2003e) subdivided subgenus *Sycidium* into two sections based primarily on differences in growth habit and the flowers, section *Palaeomorphe* King with aerial adventitious roots and hermaphroditic flowers with ovules galled by pollinators, and section *Sycidium* without aerial adventitious roots. In the present study, four major clades are recognised, which may be ranked as sections if stronger support is obtained in the future: *Palaeomorphe*, *Phaeopilosae*, *Sinosycidium*, and *Asperae* clades. Three Asian mainland species constituting section *Sinosycidium* are sister to the remaining subclades.

Asperae. This clade (Fig. 3B; 55% BS/PP=0.56) corresponds to section *Sycidium* (Miq.) Berg & Corner, but excluding *Phaeopilosae* (King) Corner and *Sinosycidium*.

Phaeopilosae. This clade constitutes a well-supported clade (Fig. 3B; 92% BS/PP=0.91) of species endemic to New Guinea largely corresponding to the *Conocephalifolia* group *sensu* Berg including *F. wassa* Roxb. and *F. copiosa* Steud. but excluding *Ficus gul* Lauterb. & K. Schum. As a result, the *Phaeopilosae* clade is confined to Eastern New Guinea. *Ficus complexa* Corner, the type species for Corner's series *Phaeopilosae*, as well as a number of other species included in Corner's series *Phaeopilosae* or in Bergs

Conocephalifolia group are not included in this study so that the circumscription and name of the *Phaeopilosae* clade is uncertain at present.

Palaeomorphe. This clade (Fig. 3B; 60% BS/PP=0.65) corresponds to section *Palaeomorphe* (King) Berg & Corner and includes about 30 species of climbers or hemi-epiphytes with aerial adventitious roots. The name refers to the frequent presence of hermaphroditic flowers instead of male ones, with an ovule capable of becoming a gal.

Sinosycidium. This clade (Fig. 3B; 100% BS/PP=1.00) corresponds to the monotypic Chinese section *Sinosycidium* Corner (*F. tsiangii*) and subsection *Ficus* series *Sinosycea* Corner comprising *F. henryi* Diels and *F. subincisa* Sm. from the Asian mainland. *Ficus subincisa* was not included in this study. The species of section *Sinosycidium* are atypical within *Sycidium* in that they present elongate stigmas in female figs and two anthers per male flower in male figs, two traits probably linked to being passively pollinated. Passive pollination has not been reported for any other species of subgenus *Sycidium*.

Sycomorus. This clade (Fig. 3C; 97% BS/PP=1.00) corresponds to subgenus *Sycomorus* (Gasp.) Miq., which includes members of sections *Sycomorus s.l.* (18 spp. including former section *Neomorphe*), *Sycocarpus* (86 spp.), and *Adenosperma* (20 spp.). In addition, this group includes a number of smaller sections (sensu Berg & Corner, 2005) with difficult affinities, namely *Dammaropsis* (Warb.) C.C. Berg (5 spp.), *Hemicardia* C.C. Berg (3 spp.), *Papuasyce* (Corner) C.C. Berg (3 spp.), and *Bosscheria* (Teijsm. & de Vriese) C.C. Berg (2 spp.). Corner (1965) only included the monoecious section *Sycomorus* in subgenus *Sycomorus*. However, based on early molecular studies (Weiblen 2000; Jousselin *et al.,* 2003), morphological evidence (Corner, 1967; Berg, 1989; Weiblen, 2000) and a shared genus of pollinating wasps (*Ceratosolen*), Berg & Corner (2005) transferred a number of dioecious sections from Corner's (1965) subgenus *Ficus* into an enlarged subgenus

Sycomorus, which we here refer to as the Sycomorus clade. Two preceding molecular studies including more taxa (Rønsted *et al.*, 2005; 2008a) did not find support for such an expanded subgenus *Sycomorus*, but this was attributed to lack of resolution and informative characters using limited DNA sequence information. Undiscovered paralogous copies were problematic in Rønsted *et al.* (2005; 2008a). In the present study, it was discovered that the primers used in previous studies preferably amplified a paralogous copy of ETS explaining previously unconfirmed incongruence between the ETS topology and other DNA regions. The design of a new *Ficus* specific ETS primer resulted in preferential amplification of the correct copy. After removal of the erroneous copies of ETS, Sycomorus was recovered as monophyletic. However, relationships within the *Sycomorus* clade are not well supported and are likely to change with future analyses, but we expect to recover clades largely corresponding to sections *Sycomorus s.l., Sycocarpus* and *Adenosperma* once the many difficult taxa in the subgenus *Sycomorus* clade are placed. Section *Sycocarpus*, and sect. *Adenoesperma*, are both resolved with low support. Section *Sycomorus s.l.* is not resolved (Fig. 3C) and we therefore refrain from naming a corresponding clade.

Berg and Corner's (2005) section *Papuasyce* includes three species, *F. itoana* Diels and *F. microdictya* Diels endemic to New Guinea and New Britain, and *F. pritchardii* Seem. endemic to Fiji (Berg & Corner, 2005). Section *Papuasyce* was listed as subsection *Papuasyce* in section *Sycocarpus* by Corner (1965). Section *Papuasyce* and section *Adenosperma* lack the nodal glands typical of section *Sycocarpus* Berg & Corner (2005). The dioecious *F. itoana* and the monoecious *F. microdictya* are sisters in the present study, whereas *F. pritchardii* was not included.

Section *Dammaropsis* includes five species, *F. dammaropsis* Diels*, F. pseudopalma* Blanco*, F. rivularis* Merr.*, F. solomonensis* Rech. and *F. theophrastoides* Seem. ranging

from the Philippines to the Solomon Islands. Corner (1965) placed *F. dammaropsis* as subsection *Dammaropsis* and *F. solomonensis* and *F. theophrastoides* in subsection *Auriculisperma*, as series *Theophrastoides* in section *Sycocarpus*. *Ficus pseudopalma* and *F. rivularis* was included as series *Pseudopalmae* and *Rivulares* respectively in subsection *Ficus* by Corner (1965). In the present analysis, all of the above species except *F. solomonensis* are included and their relationship is unresolved amongst members of Berg & Corner's (2005b) section *Adenosperma*, with which they share spirally and terminally arranged and more or less conspicuously tufted leaves (Berg, 2004a; Berg & Corner 2005).

Berg & Corners (2005) section *Hemicardia* was originally described as series *Prostratae* in section *Sycidium* (subgenus *Sycidium*; Corner, 1965). Section *Hemicardia* is supported by free tepals, and 1-2 anthers per male flower, is primarily Sino-Himalayan and includes *F. koutumensis* Corner, *F. prostrata* (Wall. ex. Miq.) Miq. and *F. semicordata* Buch.-Ham. ex Sm., the latter extending to Malesia.

Berg (2004a) remarked closer morphological affinities of section *Hemicardia* to section *Sycomorus* than to any of the other sections of the subgenus. In the present analysis, *F. koutumensis* is not included, but *F. prostrata* and *F. semicordata* form a clade (Fig. 3C; 98% BS/PP=1.00) with uncertain affinity.

Berg & Corner's (2005) section *Bosscheria* includes *F. minnahassae* and *F. pungens* ranging from the Philippines to New Guinea. Berg & Corner's (2005) section *Bosscheria* forms a clade, which is embedded in the *Sycocarpus* group in the present analysis. They are atypical within the subgenus because of their very small figs and very small flowers.

Sycocarpus. This clade (Fig. 3C; 68% BS/PP=0.71) corresponds to section *Sycocarpus* Miq.

Adenosperma. This clade (Fig. 3C; 68% BS/PP=0.51) largely corresponds to section *Adenosperma* Corner.

***Oreosycea*.** This clade (Fig. 3D; 77% BS/PP=0.62) corresponds to the paleotropic section *Oreosycea* (Miq.). Miq. tentatively including most of subsections *Glandulosae* C.C. Berg and *Pedunculatae* Sata sensu Berg and Corner (2005), but excluding subseries *Albipilae* Corner (Berg, 2003a; Berg & Corner, 2005). Corner (1960) placed section *Oreosycea* in subgenus *Pharmacosycea* (Miq.) Miq, but molecular phylogenetic evidence has suggested section *Oreosycea* is more closely related to subgenus *Sycomorus* although this is not well-supported (54 % BS/PP<0.50 in this study) nor consistent. Berg and Corner (Berg, 2003b; Berg & Corner, 2005) divided section *Oreosycea* into subsections *Glandulosa*e C.C. Berg (including series *Austrocaledonicae* Corner and series *Nervosae* Corner and *Pedunculatae* (including subseries *Vasculosae* Corner and subseries *Albipilae* Corner).

***Urostigma*.** This clade (Fig. 3D; 100% BS/PP=0.99) corresponds to section *Urostigma* sensu Corner 1960. Due to the placement of section *Urostigma* in this phylogenetic analysis and prior studies of *Ficus* (Jousselin *et al.* 2003; Rønsted *et al.* 2005; 2008a), subgenus *Urostigma* is polyphyletic. Additionally, Berg & Corner (2005) expanded section *Urostigma* uniting Corner's sections *Urostigma, Leucogyne* and *Conosycea* which is not supported by this study. The Sino-Himalayan *F. orthoneura* H. Lév. & Vanoit appears to be sister to the rest of (sub)section *Urostigma* (100 % BP/PP=1.00). *Ficus orthoneura*, *F. hookeriana* Corner (also Sino-Himalayan, but not included in this study), and *F. cornelisiana* Chantaras & Y.Q. Peng (Chanterasuwan *et al.* 2014) present a mixture of section *Urostigma* and section *Conosycea* characters and were placed in their own series within section *Urostigma* by Corner (1965). In a recent study of (sub)section *Urostigma*

(Chantarasuwan *et al.*, 2015), *F. madagascariensis* C.C. Berg (not included here) was found to be sister to the remainder of the (sub)section and the next diverging clade consisted of *F. orthoneura* and *F. hookeriana*.

***Albipilae*.** This clade (Fig. 3A; 100% BS/PP=1.00) corresponds to subseries *Albipilae* Corner and comprised two African species *F. variifolia* Warb. and *F. dicranostyla* Mildbr., as well as *F. albipila* (Miq.) King which occurs from Thailand to Australia. Morphological study of subseries *Albipilae* also assigns *F. capillipes* Gagnep. from mainland Asia and the Madagascan *F. assimilis* Baker and *F. ampana* C.C. Berg to this group but remain to be included in phylogenetic studies. The *Albipilae* clade can be distinguished from the *Oreosycea* clade primarily by the presence of hairs on the inner surface of the fig receptacle. The exact circumscription of the *Albipilae* clade awaits comprehensive species sampling.

Cariceae. This clade (Fig. 3D; 100% BS/PP=1.00) includes only the domesticated Mediterranean *F. carica* L.*,* and *F. palmata* Roxb. extending from north-eastern Africa to Pakistan. Together with *F. iidaiana* Wilson from Bonin Island (Japan), these three species formerly constituted *Ficus* section *Ficus* subsection *Ficus* Berg & Corner, but *F. iidaiana* is a member of *Frutescentiae* in the present study. The traditional subgenus *Ficus* is polyphyletic consisting of three strongly supported major clades: *Cariceae*, *Eriosycea* and *Frutescentiae* clades corresponding to clear-cut subdivisions by Berg & Corner (2005; Berg, 2003c). The relationship of the *Cariceae* clade is uncertain. *Ficus carica* is the type of genus *Ficus*.

Mixtiflores. This clade (Fig. 3D; 100% BS/PP=1.00) corresponds to subgenus *Urostigma* (Gasp.) Miq. excluding section *Urostigma* (Gasp.) Miq and includes about 265 monoecious species in two subclades, one (100% BP/PP=1.00) consisting of sect. *Conosycea* Corner (98% BP/PP=0.99) and (sub)sect. *Malvanthera* Corner (100%

BP/PP=0.99), and the other (100% BP/PP=1.00) including sect. *Galoglychia* (Gasp.) Endl. (66% BP/PP=0.68) and sect. *Americanae* Miq. (100% BP/PP=1.00). In all the species, the staminate flowers are scattered among the pistilate flowers in the fig cavity.

Galoglychia. This clade (Fig. 3E; 66% BS/PP=0.68) corresponds to the African section *Galoglychia* (Gasp.) Endl. Early studies (Rønsted *et al.*, 2005; 2007; 2008a) suggested that *Galoglychia* is paraphyletic to *Americanae*, but monophyly of *Galoglychia* has been confirmed by later studies (Renoult *et al.*, 2009; Cruaud *et al.*, 2012b). Detailed phylogenetic studies of section *Galoglychia* were published by Rønsted *et al.* (2008b) and Renoult *et al.* (2009). Based on nuclear sequences, Rønsted *et al.* (2007) found that *Galoglychia* consists of two major clades within Africa possibly corresponding to two main centers of diversity. One clade comprises members of subsections *Platyphylla*e (Mildbraed & Burret) C.C. Berg and *Chlamydodorae* (Mildbraed & Burret) C.C. Berg, which are more concentrated in Eastern Africa, and extend to Madagascar and neighbouring archipelagos (Comoros, Mascarenes, Aldabra Islands and Seychelles) and is sister to *Americanae* in the study by Rønsted *et al.* (2007). The other main clade (includes members of subsections *Caulocarpae* (Mildbraed & Burret) C.C. Berg, *Cyathistipulae* (Mildbraed & Burret) C.C. Berg, *Crassicostae* (Mildbraed & Burret) C.C. Berg and *Galoglychia,* which are concentrated in West and Central Africa (Berg, 1986). Renoult *et al.* (2009) found discordance of highly variable plastid data with the nuclear data, possibly caused by introgressive hybridization. In the present study, the six subclades are evident, but their relationships are not well-supported.

Americanae. This clade (Fig. 3E; 100% BS/PP=1.00) corresponds to Neotropical section *Americanae* Miq. including about 110 species of hemi-epiphytes with low sequence

variation possibly representing a rapid radiation. A detailed study of the Americanae clade has been published by Machado *et al.* (2018).

Conosycea. This clade (Fig. 3F; 99% BS/PP=0.99) corresponds to Corners (1965) section *Conosycea* (Miq.) Corner plus Corner's acceptation of section *Stilpnophyllum* Endl. (*Ficus elastica* Roxb.) and section *Leucogyne* (*F. amplissima* Sm. and *F. rumphii* Bl.), which Berg and Corner (2005) considered members of section *Urostigma* s.s. (= subsection *Urostigma*).

A number of clades are resolved within section *Conosycea*, some of which correspond to traditional series and subseries but the subdivisions proposed by Corner (1965) and Berg and Corner (2005) are not reflected.

Malvanthera. This clade (Fig. 3F; 98% BS/PP=0.99) corresponds to section *Malvanthera* Corner, which was reduced to subsection rank by Berg and Corner (2005). The *Malvanthera* clade includes 23 Australasian species with centers of diversity in New Guinea and in Australia. The section was included in section *Stilpnophyllum* Endl. by Berg & Corner (2005) together with *F. elastica*, but phylogenetic evidence shows that *F. elastica* is a member of the *Conosycea* clade and section *Stilpnophyllum* sensu Berg & Corner (2005) is therefore polyphyletic. A detailed phylogenetic tree of the *Malvanthera* clade was published by Rønsted *et al.* (2008b) and relationships in that study are mirrored in the present study including the same sampling for the section. Rønsted *et al.* (2008b) also highlighted problems with the species concept of Berg & Corner (2005) for *Malvanthera*. In particular Berg and Corner united the majority of the New Guinea species under *F. hesperidiiformis* King, which is not supported by phylogenetic evidence (Rønsted *et al.*, 2008b), and at the same time Berg & Corner (2005) kept a narrow species concept for the Australian species.

Pharmacosycea. This clade (Fig. 3D; 100% BS/PP=1.00) corresponds to section *Pharmacosycea* (Miq.) Bent. & Hook, includes ca. 25 species restricted to the Neotropics and was recovered as sister to all other species of *Ficus*. Polyphyly of subg. *Pharmacosycea* has been firmly established by molecular phylogenetic tree (e.g. Weiblen, 2000; Rønsted *et al.*, 2005; 2008a; Cruaud *et al.,* 2012b). Morphologically, the *Pharmacosycea* clade is very similar to the Old World section *Oreosycea* s.s, the remaining section of subg. *Pharmacosycea* (sensu Berg and Corner, 2005). Relationships within section *Pharmacosycea* were recently evaluated by Pederneiras (2015), although species names were not fully clarified.

CONCLUSIONS

Despite the extensive study of figs for their striking diversity and brood-site pollination mutualism, the deep evolutionary history of the group cannot be understood without attention to and comparison with its closest relatives, Castilleae. We introduce the subtribe Involucraoideae to recognize that figs and Castilleae comprise a group united by a trait that is central to their inflorescence morphology and pollination syndromes– involucral bracts. Here with the first intensive sampling of Castilleae and the most comprehensive phylogenetic reconstruction of figs to date, we delineate and name clades that are well supported to guide sampling in future studies of Involucraoideae and highlight those aspects of phylogenetic tree that warrant further investigation.

ACKNOWLEDGEMENTS

The authors would like to thank B. Chantarasuwan, J.M. Cook, D. Dixon, F. Forest, R. Harrison, B. Isua, E. Jousselin, R. de Kok, O. Maurin, C. Michenaud, H. Navarrete, D. Neill, K. Oyama, Y. Qiong-Peng, S. Swensen, R. Valencia, and J. Wright for providing material or assistance in obtaining material for this study. This work was supported by Garden Club of America, Explorer's Club, and ASPT Graduate awards to WLC, an International Network Grant from the Danish Ministry of Science, Innovation & Education, 2012 to NR, WLC and GDW, the Danish Research Councils grant (7014-00205) to NR and SBL, an NSF Doctoral Dissertation Improvement Grant (0603768) to WLC and GDW, and NSF DEB grant (0128833) to GDW.

REFERENCES

Anstett M-C, Hossaert-McKey M, Kjellberg F. 1997. Figs and fig pollinators: evolutionary conflicts in a coevolved mutualism. *Trends of Research in Ecology and Evolution* **12**: 94-99.

Azuma H, Harrison RD, Nakamura K, Su Z-H. 2010. Molecular phylogenies of figs and fig-pollinating wasps in the Ryukyu and Bonin (Ogasawara) islands, Japan. *Genes and Genetic Systems* **85**: 177–192.

Bain A, Borges RM, Chevallier MH, Vignes H, Kobmoo N, Peng YQ, Cruaud A, Rasplus JY, Kjellberg F, Hossaert-McKey M. 2016. Geographic structuring into vicariant species-pairs in wide-ranging, high dispersal plan-insect mutualism: the case of *Ficus racemosa* and its pollinating wasps. *Evolutionary Ecology* **30**: 663-684.

Berg CC. 2001. Moreae, Artocarpeae, and *Dorstenia* (Moraceae). New York Botanical

Garden, New York. 346 pp.

Berg CC. 2004a. Flora Malesiana precursor for the treatment of Moraceae 6: *Ficus* subgenus *Sycomorus*. *Blumea* **49**: 155–200.

Berg CC. 2004b. Flora Malesiana precursor for the treatment of Moraceae 7: *Ficus* subgenus *Urostigma*. *Blumea* **49**: 463–480.

Berg CC. 2003a. Flora Malesiana precursor for the treatment of Moraceae 1: The main subdivision of *Ficus*: The subgenera. *Blumea* **48**: 166–177.

Berg CC. 2003b. Flora Malesiana precursor for the treatment of Moraceae 2: *Ficus* subgenus *Pharmacosycea* section *Oreosycea*. *Blumea* **48**: 289–301.

Berg CC. 2003c. Flora Malesiana precursor for the treatment of Moraceae 3: *Ficus* subgenus *Ficus*. *Blumea* **48**, 529–550.

Berg CC. 2003d. Flora Malesiana precursor for the treatment of Moraceae 4: *Ficus* subgenus *Synoecia*. *Blumea* **48**: 551–571.

Berg CC. 2003e. Flora Malesiana precursor for the treatment of Moraceae 5: *Ficus* subgenus *Sycidium*. *Blumea* **48**: 573–597.

Berg CC. 1989. Classification and distribution of *Ficus*. *Experientia* 45: 605–611.

Berg CC. 1977. The Castilleae, a tribe of the Moraceae, renamed and redefined due to the exclusion of the type genus *Olmedia* from the "Olmedieae." *Acta Botanica Neerlandica* **26**: 73-82.

Berg CC. 1972. Olmedieae Brosimeae (Moraceae), in: Flora Neotropica. Hafner Publishing Company, New York, pp. 1–228.

Berg CC. 1986. Subdivision of *Ficus* subgenus *Urostigma* section *Galoglychia* (Moraceae).

Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series C **89**: 121-127.

Berg CC, Corner EJH. 2005. Flora Malesiana, Series I. Volume 17 part 2., in: Nooteboom, H. (Ed.), Flora Malesiana Series I - Seed Plants Vol. 17 Part 2. Nationaal Herbarium Nederland, Leiden, pp. 1–730.

Berg CC, Corner EJH, Jarrett FM. 2006. Flora Malesiana, Series 1. Volume 17, Part 1: Moraceae-genera other than *Ficus*., in: Nootboom, H.P. (Ed.), Flora Malesiana Series I - Seed Plants Vol. 17 Part 1. Nationaal Herbarium Nederland, Leiden, pp. 1–154.

Bronstein JL. 1988. Mutualism, Antagonism, and the fig-pollinator interaction. *Ecology* **69**: 1298–1302.

Bruun-Lund S, Clement WL, Kjellberg F, Rønsted N. 2016. First plastid phylogenomic study reveals cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). *Molecular Phylogenetics and Evolution* **109**: 93-104.

Calonje M, Martín-Bravo S, Dobeš C, Gong W, Jordon-Thaden I, Kiefer C, Kiefer M, Paule J, Schmickl R, Koch MA. 2009. Non-coding nuclear DNA markers in phylogenetic reconstruction. *Plant Systematics and Evolution* **282**: 257–280.

Chantarasuwan B, Peng Y-Q, Baas P, Rasplus J-Y, van Heuven B-J, van Welzen PC. 2014. *Ficus cornelisiana*, a new species of *Ficus* subsection *Urostigma* (Moraceawe) from the Sino-himalayan region. ***Blumea* 59:** 6-9.

Chantarasuwan B, Berg CC, Kjellberg F, Rønsted N, Garcia M, Baider C, van Welzen PC. 2015. A New Classification of *Ficus* Subsection *Urostigma* (Moraceae) Based on four nuclear DNA markers (ITS, ETS, *G3pdh*, and ncpGS), morphology and leaf anatomy. *PLoS One* 10, e0128289.

Clement WL. 2008. Phylogeny and pollination ecology of Castilleae (Moraceae). Ph.D. Diss. University of Minnesota, St. Paul, MN.

Clement WL, Weiblen GD, 2009. Morphological evolution in the mulberry family (Moraceae). *Systematic Botany* **34**: 530–552.

Conchou L, Cabioch L, Rodriguez LFV, Kjellberg F. 2014. Daily rhythm of mutualistic pollinator activity and scent emission in *Ficus septica*: ecological differentiation between co-occuring pollinators and potential consequences for chemical communication and facilitation of host speciation. *PlosONE* **9**: e103581.

Cook JM, Rasplus J-Y. 2003. Mutualists with attitude: coevolving fig wasps and figs. *Trends in Ecology and Evolution* **18**: 241–248.

Corner EJH. 1959. Taxonomic notes on *Ficus* Linn., Asia and Australasia III. Subgen. *Ficus* and sect. *Ficus*. *Gardens Bulletin of Singapore* **17**: 416-441.

Corner EJH. 1967. *Ficus* in the Solomon islands and its bearing on the post-jurassic history of Melanesia. *Philosophical Transactions of the Royal Society of London B. Biological Sciences* **253**: 23–159.

Corner EJH. 1965. Check-list of *Ficus* in Asia and Australasia with keys to identification. Botanic Gardens Singapore, 1965. *Gardens Bulletin of Singapore* **21**: 1–186.

Croat TB. 1978. Flora of Barro Colorado Island. Stanford University Press, California, USA.

Cruaud A., Jabbour-Zahab R, Genson G, Ungricht S, Rasplus JY. 2012a. Testing the emergence of new caledonia: Fig wasp mutualism as a case study and a review of evidence. *PLoS One* **7**: e30941.

Cruaud A, Rønsted N, Chantarasuwan B, Chou LS, Clement WL, Couloux A, Cousins B, Genson G, Harrison RD, Hanson PE, Hossaert-McKey M, Jabbour-Zahab R, Jousselin E, Kerdelhué C, Kjellberg F, Lopez-Vaamonde C, Peebles J, Peng Y-Q, Pereira RAS, Schramm T, Ubaidillah R, van Noort S, Weiblen GD, Yang D-R, Yodpinyanee A, Libeskind-Hadas R, Cook JM, Rasplus J-Y, Savolainen V. 2012b. An extreme case of plant-insect codiversification: figs and fig-pollinating wasps. *Systematic Biology* **61**: 1029–47.

Darriba D, Taboada GL, Doallo R, Posada D, 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772–772.

Datwyler SL, Weiblen GD. 2004. On the origin of the fig: phylogenetic relationships of Moraceae from ndhF sequences. *American Journal of Boatny* **91**: 767–777.

Erixon P, Svennblad B, Britton T, Oxelman B. 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Systematic Biology* **52**: 665–673.

Evans RC, Alice LA, Campbell CS, Kellogg EA, Dickinson TA. 2000. The granule-bound starch synthase (GBSSI) gene in the Rosaceae: multiple loci and phylogenetic utility. *Molecular Phylogenetics and Evolution* **17**: 388–400.

Harrison RD, Rønsted N, Xu L, Rasplus J-Y, Cruaud A. 2012. Evolution of fruit traits in *Ficus* subgenus *Sycomorus* (Moraceae): to what extent do frugivores determine seed dispersal mode? *PLoS One* **7**: e38432.

Herre EA. 1989. Coevolution of reproductive characteristics in 12 species of New World figs and their pollinator wasps. *Experientia* **45**, 637–647.

Herre EA, Jandér KC, Machado CA. 2008. Evolutionary ecology of figs and their

associates: recent progress and outstanding puzzles. *Annual Reviews of Ecology, Evolution, and Systematics* **39**: 439–458.

Herre EA, Machado CA, Bermingham E, Nason JD, Windsor DM, McCafferty SS, VanHouten W, Bachmann K. 1996. Molecular phylogenies of figs and their pollinator wasps. *Journal of Biogeography* **23**: 521–530.

Herre EA, West SA. 1997. Conflict of interest in a mutualism: documenting the elusive fig wasp-seed trade-off. *Proceedings of The Royal Society of London B. Biological Sciences* **264**: 1501–1507.

Huelsenbeck JP, Larget B, Miller RE, Ronquist F. 2002. Potential applications and pitfalls of Bayesian inference of phylogeny. *Systematic Biology* **51**: 673–688.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17**: 754–755.

Jackson AP, Machado CA, Robbins N, Herre EA. 2008. Multi-locus phylogenetic analysis of neotropical figs does not support co-speciation with the pollinators : The importance of systematic scale in fig/wasp cophylogenetic studies. *Symbiosis* **45**: 57-72.

Jousselin E, Van Noort S, Berry V, Rasplus J-Y, Rønsted N, Erasmus JC, Greeff JM. 2008. One fig to bind them all: Host conservatism in a fig wasp community unraveled by cospeciation analyses among pollinating and nonpollinating fig wasps. *Evolution* **62**: 1777–1797.

Jousselin EE, Rasplus J-Y, Kjellberg F. 2003. Convergence and coevolution in a mutualism: Evidence from a molecular phylogeny of *Ficus*. *Evolution* **57**: 1255–1269.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7:

Improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.

Kusumi J, Azuma H, Tzeng H-Y, Chou L-S, Peng Y-Q, Nakamura K, Su Z-H., 2012. Phylogenetic analyses suggest a hybrid origin of the figs (Moraceae: *Ficus*) that are endemic to the Ogasawara (Bonin) Islands, Japan. *Molecular Phylogenetics and Evolution* **63**: 168–179.

Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* **29**: 1695–1701.

Li HQ, Chen JY, Wang S, Xiong SZ. 2012. Evaluation of six candidate DNA barcoding loci in *Ficus* (Moraceae) of China. *Molecular Ecology Resources* **12**: 783–790.

Li M, Wunder J, Bissoli G, Scarponi E, Gazzani S, Barbaro E, Saedler H, Varotto C. 2008. Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. *Cladistics* **24**: 727–745.

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

Lopez-Vaamonde C, Rasplus J-Y, Weiblen GD, Cook JM. 2001. Molecular phylogenies of fig wasps: partial cocladogenesis of pollinators and parasites. *Molecular Phylogenetics and Evolution* **21**: 55–71.

Machado A, Rønsted N, Bruun-Lund S, Pereira R, Queiroz L. 2018. From Atlantic forest to the all Americas: Biogeographical history and divergence times of Neotropical *Ficus* (Moraceae). *Molecular Phylogenetics and Evolution* **122**: 46-58.

Machado CA, Robbins N, Gilbert MTP, Herre EA. 2005. Critical review of host

specificity and its coevolutionary implications in the fig fig-wasp mutualism. *Proceedings of the National Academy of Sciences of the Unites States of America* **102** Suppl 1: 6558-65.

Marussich WA, Machado CA. 2007. Host-specificity and coevolution among pollinating and nonpollinating New World fig wasps. *Molecular Ecology* **16**: 1925–1946.

Mcleish M, Guo D, van Noort S, Midgley G. 2011. Life on the edge: Rare and restricted episodes of a pan-tropical mutualism adapting to drier climates. *New Phytologist* **191**: 210–222.

McLeish MJ, van Noort S. 2012. Codivergence and multiple host species use by fig wasp populations of the *Ficus* pollination mutualism. *BMC Evolutionary Biology* **12**: 1.

Moe AM, Clement W, Weiblen GD. 2012. Rapid evolution of pollinator-mediated plant reproductive isolation, in: Singh RS, Xu J, Kulathinal R. (Eds.), Evolution in the Fast Lane: Rapidly Evolving Genes and Genetic Systems. Oxford University Press, Oxford.

Osmaston HA. 1965. Pollen and seed dispersal in *Chlorophora excelsa* and other Moraceae, and in *Parkia filicoidea* (Mimosaceae), with special reference to the role of the fruit bat, *Eidolon helvum*. *Commonwealth Forestry Review* **44**: 96–104.

Pederneiras LC, Romaniuc-neto S, Mansano VDF. 2015. Molecular phylogenetics of *Ficus* section *Pharmacosycea* and the description of *Ficus* subsection *Carautaea* (Moraceae). *Systematic Botany* **40:** 504–509.

Posada D, Buckley TR. 2004. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Systematic Biology* **53**: 793–808.

Rambaut A. 2007. Tracers v1.5. Available from http://tree.bio.ed.ac.uk/software/tracer/.

Renoult JP, Kjellberg F, Grout C, Santoni S, Khadari B. 2009. Cyto-nuclear discordance in the phylogeny of *Ficus* section *Galoglychia* and host shifts in plant-pollinator associations. *BMC Evolutionary Biology* **9**: 248.

Rodriguez LJ, Bain A, Chou LS, Conchou L, Cruaud A, Gonzales R, Hossaert-McKey M, Rasplus JY, Tzeng HY, Kjellberg F, 2017. Diversification and spatial strcuturing in the mutualism between *Ficus septica* and its pollinating wasps in insular South East Asia. *BMC Evolutionary Biology* **17**: 207.

Rønsted N, Salvo G, Savolainen V, 2007. Biogeographical and phylogenetic origins of African fig species (*Ficus* section *Galoglychia*). *Molecular Phylogenetics and Evolution* **43**: 190–201.

Rønsted N, Weiblen GD, Clement WL, Zerega NJC, Savolainen V. 2008a. Reconstructing the phylogeny of figs (*Ficus*, Moraceae) to reveal the history of the fig pollination mutualism. *Symbiosis* **45**: 1–12.

Rønsted N, Weiblen GD, Cook JM, Salamin N, Machado CA, Savolainen V. 2005. 60 million years of co-divergence in the fig-wasp symbiosis. *Proceedings of the Royal Society of London B. Biological Sciences* **272**: 2593–2599.

Rønsted N, Weiblen GD, Savolainen V, Cook JM. 2008b. Phylogeny, biogeography, and ecology of *Ficus* section *Malvanthera* (Moraceae). *Molecular Phylogenetics and Evolution* **48**: 12–22.

Sakai S, Kato M, Nagamasu H. 2000. *Artocarpus* (Moraceae) - Gall midge pollination mutualism mediated by a male-flower parasitic fungus. *American Journal of Botany* **87**: 440–445.

Silvieus SI, Clement WL, Weiblen GD. 2007. Cophylogeny of Figs, Pollinators, Gallers,

and Parasitoids, in: Tilmon K. (Ed.), Specialization, Speciation and Radiation: The Evolutionary Biology of Herbivorous Insects. University of California Press, California, pp. 225–239.

Smith CI, Pellmyr O, Althoff DM, Balcázar-Lara M, Leebens-Mack J, Segraves KA. 2008. Pattern and timing of diversification in *Yucca* (Agavaceae): specialized pollination does not escalate rates of diversification. *Proceedings of the Royal Academy of Biological Sciences* **275**: 249–258.

Stebbins GL. 1981. Coevolution of grasses and herbivores. *Annals of The Missouri Botanical Garden* **68**: 75-86.

Sun Y, Skinner DZ, Liang GH, Hulbert SH. 1994. Phylogenetic analysis of *Sorghum* and related taxa using internal transcribed spacers of nuclear ribosomal DNA. *Theoretical and Applied Genetics* **89**: 26-32.

Weiblen G. 2004. Correlated evolution in fig pollination. *Systematic Biology* **53**: 128–139.

Weiblen GD. 2001. Phylogenetic relationships of fig wasps pollinating functionally dioecious *Ficus* based on mitochondrial DNA sequences and morphology. *Systematic Biology* **50**: 243–267.

Weiblen GD. 2000. Phylogenetic relationship of functionally dioecious *Ficus* (Moraceae) based on ribosomal DNA sequences and morphology. *American Journal of Botany* **87**: 1342–1357.

Weiblen GD, Bush GL. 2002. Speciation in fig pollinators and parasites. *Molecular Ecology* **11**: 1573–1578.

Weiblen GD, Yu DW, West SA. 2001. Pollination and parasitism in functionally dioecious figs. *Proceedings of the Royal Society of London B. Biological Sciences* **268**: 651–659.

Xu L, Harrison RD, Yang P, Yang DR. 2011. New insight into the phylogenetic and biogeographic history of genus *Ficus*: Vicariance played a relatively minor role compared with ecological opportunity and dispersal. J*ournal of Systematics and Evolution* **49**: 546–557.

Zerega N, Mound L, Weiblen GD. 2004. Pollination in the New Guinea endemic *Antiaropsis decipiens* (Moraceae) is mediated by a new species of thrips, *Thrips antiaropsidis* sp. nov. (Thysanoptera: Thripidae). *International Journal of Plant Sciences* **165**: 1017–1026.

Zerega NJC, Clement WL, Datwyler SL, Weiblen GD. 2005. Biogeography and divergence times in the mulberry family (Moraceae). *Molecular Phylogenetics and Evolution* **37**: 402–416.

Zerega NJC, Nur Supardi MN, Motley TJ. 2010. Phylogeny and recircumscription of Artocarpeae (Moraceae) with a focus on *Artocarpus*. *Systematic Botany* **35**: 766–783.

Zhang Q, Onstein RE, Little SA, Sauquet H. 2018. Estimating divergence times and ancestral breeding system in *Ficus* and Moraceae. *Annals of Botany* mcy159, https://doi.org/10.1093/aob/mcy159

Zwickl DJ, Hillis DM. 2002. Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology* **51**: 588–598.

Zwickl DJ. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. Diss. The University of Texas at Austin. www.bio.utexas.edu/faculty/antisense/garli/Garli.html.

Figure captions

Figure 1. Phylogenetic trees from individual (upper left panel) and combined (main tree) maximum likelihood analyses of Involucraoideae using ITS, *G3pdh*, and *GBSSI*. Thickened branches represent posterior probabilities greater than 0.95, and maximum likelihood bootstrap values are indicated above the branches (main tree only). Genera within Involucraoideae are represented by different colours consistent between the gene trees and combined phylogenetic tree. Within Ficeae, clades corresponding to named sections have been collapsed where possible (full tree not shown). For the three gene trees, *G3pdh*, ITS and *GBSSI*, clades have been collapsed based on genus or clades with a genus to compare relationships among these groups within each tree (all trees are available in TreeBase accession S24008).

Figure 2. Cladogram based on relationships reconstructed from the maximum likelihood analysis of the six-gene *Ficus* dataset (detailed tree in Fig. 3A-F) providing an overview of the current phylogenetic understanding of relationships within *Ficus*. Approximate number of species in each clade indicated to the left of each clade name, and the subgeneric classification based on Berg & Corner (2005) indicated on the right hand side of the coloured boxes. ML bootstrap support indicated as follows: thickened branch = 95-100%,

thin branch = 70-94%, and dashed branches = <69%; posterior probability greater than 0.95 indicated with an asterisk.

Figure 3A-F. Maximum likelihood tree of the combined analysis of six gene regions for 307 species of *Ficus*. ML bootstrap support indicated as follows: thickened branches = 95-100%, thin branches = 70-94%, and dashed branches = <69%; posterior probability greater than 0.95 indicated with an asterisk. Species included in phylogenetic analysis of *Ficus* for the first time marked in bold. Proposed names for monophyletic groups of figs are indicated to the right of each clade throughout the figure.

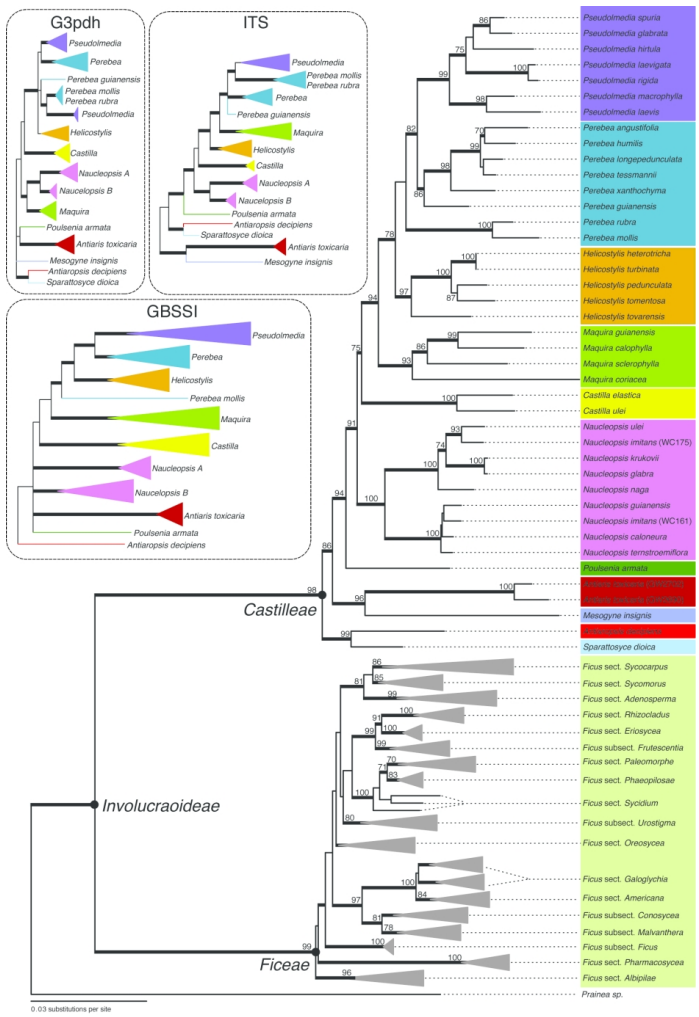


Figure 1. Phylogenetic trees from individual (upper left panel) and combined (main tree) maximum likelihood analyses of Involucraoideae using ITS, *G3pdh*, and *GBSSI*. *Thickened branches represent posterior probabilities greater than 0.95, and maximum likelihood bootstrap values are indicated above the branches (main tree only). Genera within Involucraoideae are represented by different colours consistent between the gene trees and combined phylogenetic tree. Within Ficeae, clades corresponding to named sections have been collapsed where possible (full tree not shown). For the three gene trees, G3pdh, ITS and GBSSI, clades have been collapsed based on genus or clades with a genus to compare relationships among these groups within each tree (all trees are available in TreeBase accession S24008).*

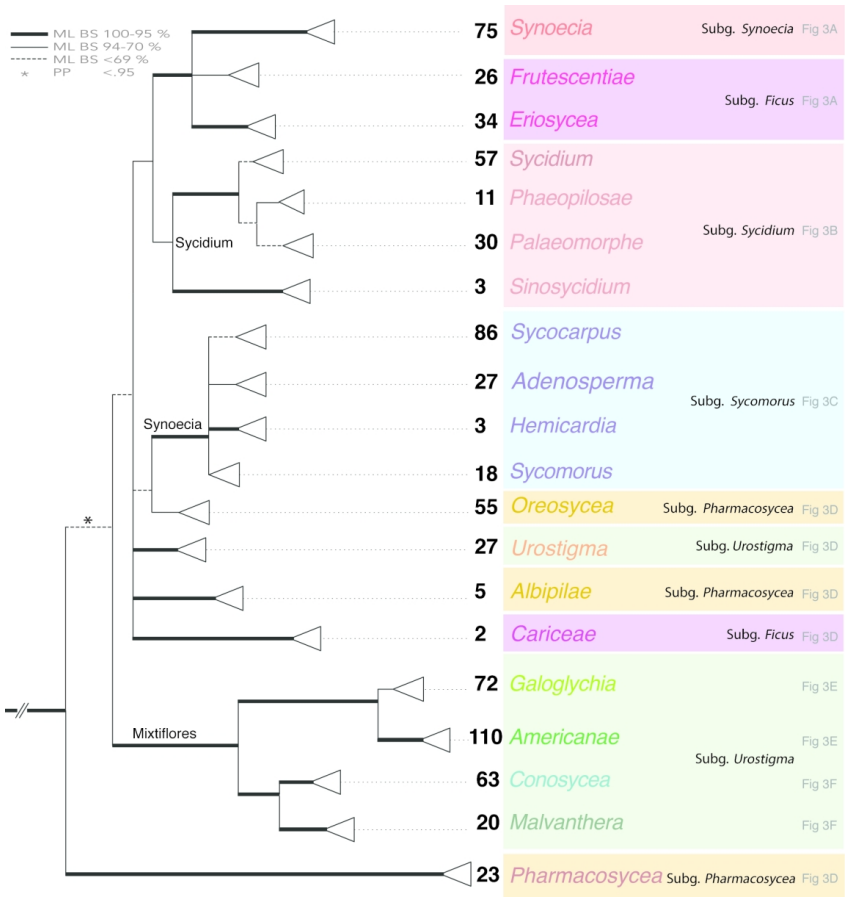


Figure 2. Cladogram based on relationships reconstructed from the maximum likelihood analysis of the six-gene *Ficus* dataset (detailed tree in Fig. 3A-F) providing an overview of the current phylogenetic understanding of relationships within *Ficus*. Approximate number of species in each clade indicated to the left of each clade name, and the subgeneric classification based on Berg & Corner (2005) indicated on the right hand side of the coloured boxes. ML bootstrap support indicated as follows: thickened branch = 95-100%, thin branch = 70-94%, and dashed branches = <69%; posterior probability greater than 0.95 indicated with an asterisk.

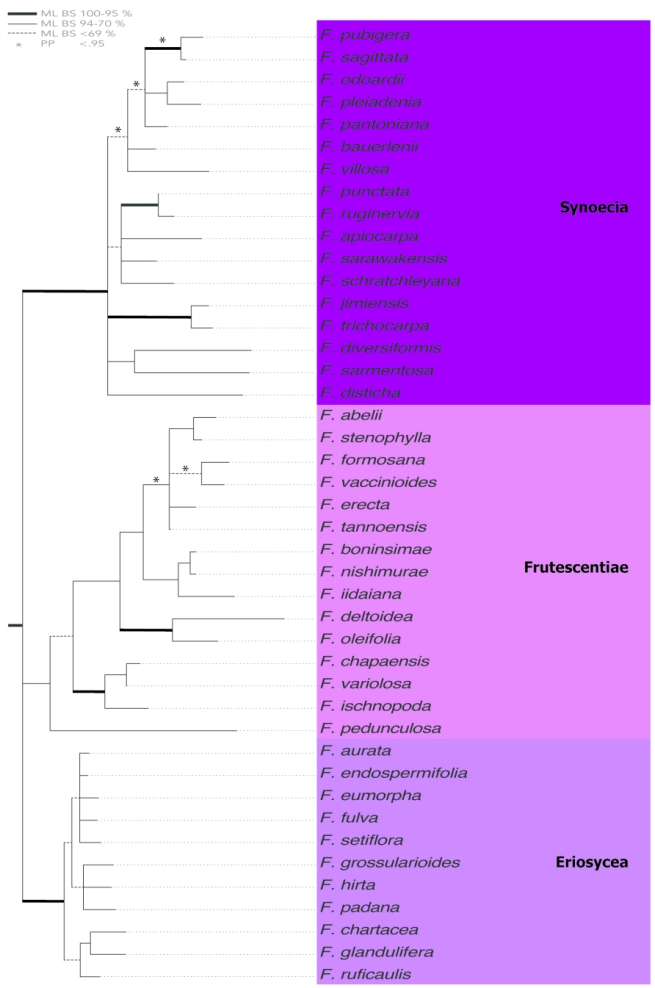


Figure 3A-F. Maximum likelihood tree of the combined analysis of six gene regions for 307 species of *Ficus*. ML bootstrap support indicated as follows: thickened branches = 95-100%, thin branches = 70-94%, and dashed branches = <69%; posterior probability greater than 0.95 indicated with an asterisk. Species included in phylogenetic analysis of *Ficus* for the first time marked in bold. Proposed names for monophyletic groups of figs are indicated to the right of each clade throughout the figure.

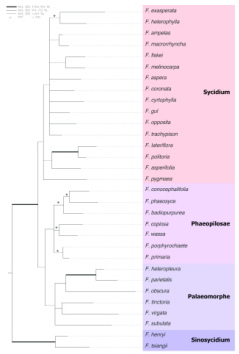


Figure 3A-F. Maximum likelihood tree of the combined analysis of six gene regions for 307 species of *Ficus*. ML bootstrap support indicated as follows: thickened branches = 95-100%, thin branches = 70-94%, and dashed branches = <69%; posterior probability greater than 0.95 indicated with an asterisk. Species included in phylogenetic analysis of *Ficus* for the first time marked in bold. Proposed names for monophyletic groups of figs are indicated to the right of each clade throughout the figure.

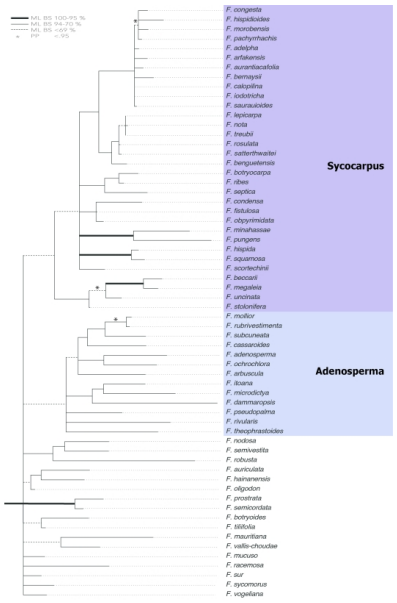


Figure 3A-F. Maximum likelihood tree of the combined analysis of six gene regions for 307 species of *Ficus*. ML bootstrap support indicated as follows: thickened branches = 95-100%, thin branches = 70-94%, and dashed branches = <69%; posterior probability greater than 0.95 indicated with an asterisk. Species included in phylogenetic analysis of *Ficus* for the first time marked in bold. Proposed names for monophyletic groups of figs are indicated to the right of each clade throughout the figure.

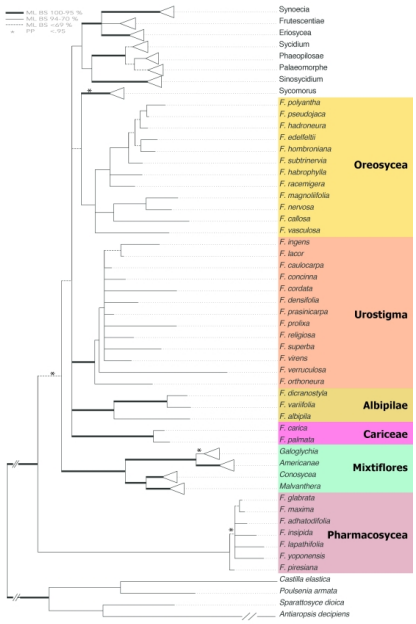


Figure 3A-F. Maximum likelihood tree of the combined analysis of six gene regions for 307 species of *Ficus*. ML bootstrap support indicated as follows: thickened branches = 95-100%, thin branches = 70-94%, and dashed branches = <69%; posterior probability greater than 0.95 indicated with an asterisk. Species included in phylogenetic analysis of *Ficus* for the first time marked in bold. Proposed names for monophyletic groups of figs are indicated to the right of each clade throughout the figure.

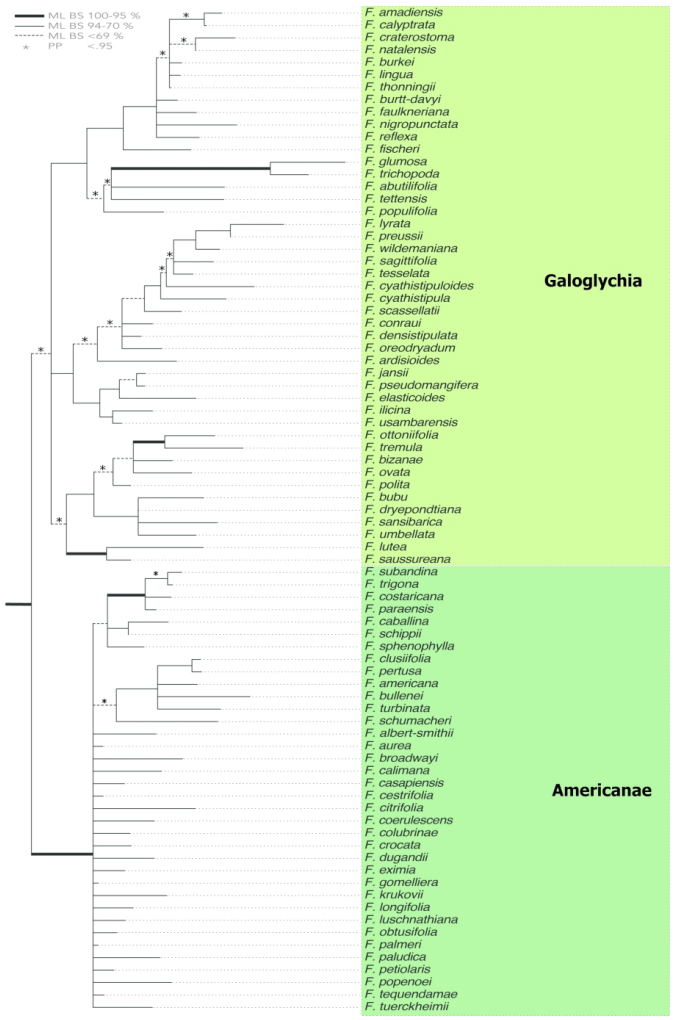


Figure 3A-F. Maximum likelihood tree of the combined analysis of six gene regions for 307 species of *Ficus*. ML bootstrap support indicated as follows: thickened branches = 95-100%, thin branches = 70-94%, and dashed branches = <69%; posterior probability greater than 0.95 indicated with an asterisk. Species included in phylogenetic analysis of *Ficus* for the first time marked in bold. Proposed names for monophyletic groups of figs are indicated to the right of each clade throughout the figure.

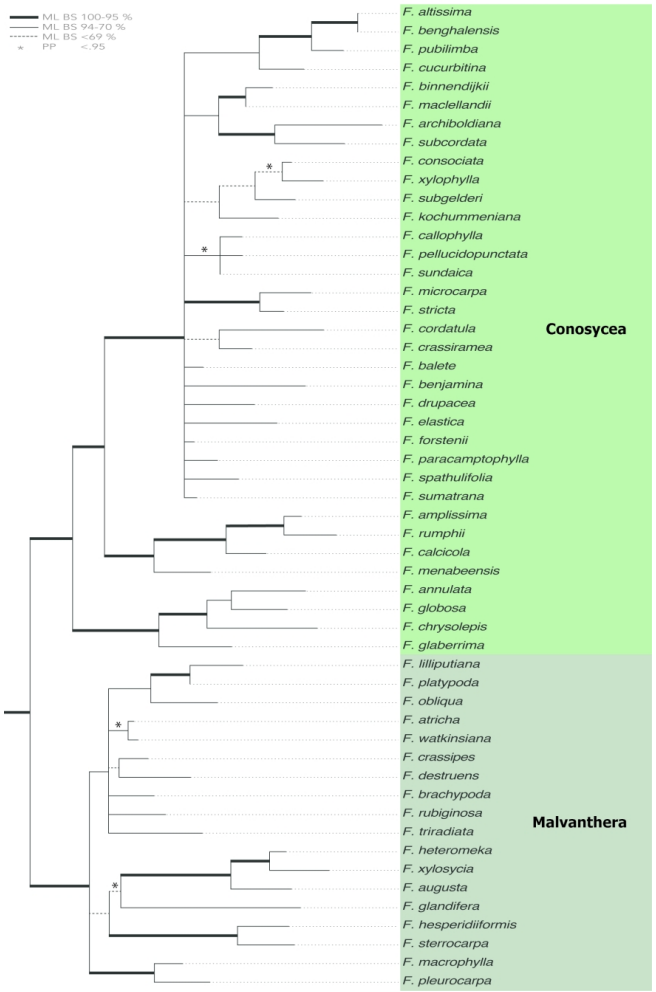


Figure 3A-F. Maximum likelihood tree of the combined analysis of six gene regions for 307 species of *Ficus*. ML bootstrap support indicated as follows: thickened branches = 95-100%, thin branches = 70-94%, and dashed branches = <69%; posterior probability greater than 0.95 indicated with an asterisk. Species included in phylogenetic analysis of *Ficus* for the first time marked in bold. Proposed names for monophyletic groups of figs are indicated to the right of each clade throughout the figure.

Supplementary information for:

Evolution and classification of figs (*Ficus*) and their close relatives (Castilleae) united by involucral bracts

WENDY L. CLEMENT[1], SAM BRUUN-LUND[2], ALANNA COHEN[1,3]; FINN KJELLBERG[4]; GEORGE D. WEIBLEN[5], NINA RØNSTED[2]*

[1]*Department of Biology, The College of New Jersey, Ewing, NJ USA.*

[2]*The Natural History Museum of Denmark, University of Copenhagen, Denmark.*

[3]*Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ USA.*

[4]*CEFE, CNRS, University of Montpellier, University Paul Valéry Montpellier, EPHE, IRD, Montpellier, France.*

[5]*Department of Plant Biology, University of Minnesota, MN USA.*

Supplementary Table S1. Voucher information for Involucraoideae DNA sequences sorted alphabetically by genus and within *Ficus* alphabetically by clade (refer to Fig. 2). Genbank numbers are provided for all data analysed in this study. Additionally, for new data generated in this study the voucher, herbarium, and country of origin are provided. Taxa included for the first time in the phylogeny are indicated in bold text, and new sequences are indicated by an asterisk next to the Genbank number. Herbarium abbreviations are as follows: AAU – Aarhus University; A – Harvard University; BG – University of Bergen; BR – Botanic Garden Meise; C – University of Copenhagen; DNA - Northern Territory Herbarium; F – Field Museum of Natural History; HITBC – Xishuangbanna Tropical Botanical Garden, Academia Sinica; HON – Sichuan Grassland Research Institute; K –Royal Botanic Gardens, Kew; LAE – Papua New Guinea Forest Research Institute; MIN – University of Minnesota; MO – Missouri Botanical Garden; NBG – South African National Biodiversity Institute; PUH – University of the Philippines; REU – Université de la Réunion. Gene region abbreviations are as follows: ITS – internal transcribed spacer region; ETS – external transcribed spacer region; G3phd – glyceraldehyde 3-phosphate dehydrogenase gene; ncpGS – glutamine synthetase gene; GBSSI – granule bound starch synthase gene; At103 – Magnesium-protoporphyrin monomethyl ester cyclase gene. Cult. refers to a plant sampled from a living collection.

Taxon	Voucher	Origin	ITS	ETS	G3pdh	ncpGS	GBSSI	At103
Antiaropsis								
A. decipiens K. Schum.	G.D. Weiblen 1706 (MIN)	Papua New Guinea	AY730142	-	EF092326	-	Pending*	KJ417793*
Antiaris								

A. toxicaria (Pers). Lesch.	G.D. Weiblen 2690 (MIN)	Papua New Guinea	Pending*	-	Pending*	-	Pending*	-
A. toxicaria (Pers). Lesch.	G.D. Weiblen 2702 (MIN)	Malaysia	Pending*	-	Pending*	-	Pending*	-
Castilla								
C. elastica Sess	G.D. Weiblen 1415 (MIN)	Costa Rica	Pending*	-	Pending*	-	Pending*	-
C. elastica Sess	G.D. Weiblen 1433 (MIN)	Panama	AY730143	AY730232	EF092327	-	EU084352	KJ417672*
C. ulei Warb.	G.D. Weiblen 1456 (MIN)	Peru	Pending*	-	Pending*	-	Pending*	-
Ficus								
***Adenosperma* clade**								
F. adenosperma Miq.	G.D. Weiblen 1764 (MIN)	Papua New Guinea	AF165374	KP406987*	EF092374	EU084324	DQ367636	KJ417804*
F. arbuscula Lauterb./K.Schum	G.D. Weiblen 500 (MIN)	Papua New Guinea	EU091617	-	-	-	EU084375	-
F. casearioides King.			EU091618	-	EU087654	-	-	-
F. dammaropsis Diels	G.D. Weiblen 1744 (MIN)	Papua New Guinea	AF165383	KP406988*	DQ367621	-	DQ367641	KJ417755*

F. itoana Diels	G.D. Weiblen 622 (A)	Papua New Guinea	AF165391	KP406989*	EU087655	-	EU084376	KJ417706*
F. microdictya Diels	G.D. Weiblen 954 (MIN)	Papua New Guinea	AF165394	-	EU087656	-	EU084377	-
F. mollior F.Muell. ex. Benth.	G.D. Weiblen 1544 (MIN)	Papua New Guinea	-	KP406990*	DQ367623	-	DQ367643	KJ417764*
F. ochrochlora Ridley	G.D. Weiblen 735 (A)		AF165396	EU084448	-	-	EU084378	-
F. pseudopalma Blanco	R.D. Harrison 610 (PUH)	Philippines	EU091629	EU084457*	EU087664	EU084329	-	KJ417674*
F. rivularis Merr.			EU091619	-	EU087657	-	-	-
F. rubrivestimenta Weiblen & Whitfeld			DQ457093	-	DQ457092	-	-	-
F. subcuneata Miq.	G.D. Weiblen 700 (A)		EU091620	-	-	-	-	-
F. subcuneata Miq.	G.D. Weiblen 2166 (MIN)		-	-	DQ367631	-	DQ367651	-
***Albipilae* clade**								
F. albipila (Miq.) King	B. Chanterasuwan s.n. (no voucher information)	Thailand	AF165375	EU084406*	EF092366	EU084298*	EU084355*	KJ417728*
F. dicranostyla Mildbr.	N. Rønsted 152 (K)	Cult. (BG) 1988-241. Orig. acc Utrecht Bot. Gard.	EU091566*	EU084407*	EF092368	EU084300*	EU084357*	KJ417771*
F. variifolia Warb.	N. Rønsted 131 (K)	Cult. (BG) 1988-240	EF092318	EF092319	EU087618*	EU084301*	EU084360*	KJ417710*

***Americanae* clade**								
F. albert-smithii Standl.	N. Rønsted 105 (K)	Cult. (BG) 1989-519. No. prov.	AY730069	AY730157	EU087635*	DQ455612*	-	KJ417681*
F. americana Aubl.	N. Rønsted 154 (K)	Cult. (BG) 1994-0678. No. prov.	AY730070	AY730158	EF092339	DQ455613	-	KJ417737*
F. aurea Nutt.			EU091598	EU084431	EU087636	-	-	-
F. broadwayi Urb	N. Rønsted 121 (K)	Cult. (BG) 1989-522. Orig. acc. Utrecht Bot. Gard.	AY730072	AY730160	EF092341	-	-	KJ417766*
F. bullenei I.M.Johnst.			EU081758	-	EU089833	-	-	-
F. caballina Standl.	N. Rønsted 101 (K)	Cult. (BG) 1992-0922. Ecuador	AY730073	AY730161	-	-	-	KJ417758*
F. calimana Dugand	N. Rønsted 108 (K)	Cult. (BG) 1989-540 . No. prov.	AY730074	AY730162	-	EU084311*	-	-
F. casapiensis (Miq) Miq.			AY730075	AY730163	-	-	-	-
F. cestrifolia Schott	N. Rønsted 139 (K)	Cult. (BG) 2000.0278. Brazil.	AY730076	AY730164	EF092342	DQ455614*	-	-
F. citrifolia Mill.	N. Rønsted 112 (K)	Cult. (BG) 1989.537. No. prov.	AY730077	AY730165	AY967955	DQ455615	-	KJ417784*

***F. clusiifolia* Schott.**	**N. Rønsted 155 (K)**	**Cult. (BG) 2000-332. Brazil**	**EU091600***	**EU084433***	-	-	-	**KJ417761***
***F. coerulescens* (Rusby) Rossb.**	**N. Rønsted 122 (K)**	**Cult. (BG) 1989-526.** No. prov.	**EU091601***	**EU084434***	**EU087638***	**EU084312***	-	**KJ417765***
F. colubrinae Standl.			EU081764	-	EU089848	-	-	-
F. costaricana (Liebm.) Miq.	K. Oyama s.n. (no voucher information)	Mexico	EU091602	EU084435	AY967952	-	-	KJ417749*
F. crocata (Miq.) Miq.			AY730080	AY730168	EF092343	DQ455618	-	-
F. dugandii Standl.			EU081763	-	AY967957	-	-	-
F. eximia Schott			AY730079	AY730167	EF092344	-	-	-
F. gommelleira Kunth & Bouché	N. Rønsted 153 (K)	Cult. (BG) 2000-0832. Brazil	AY730081	AY730169	-	-	-	KJ417738*
***F. krukovii* Standl.**	**N. Rønsted 141 (K)**	**Cult (BG) 1988-254. Orig. acc. Utrecht Bot. Gard.**	**EU091603***	**EU084436***	-	**EU084313***	-	-
F. longifolia Schott.	N. Rønsted 140 (K)	Cult. (BG) 1999-1021. No prov.	EU091604	KP406978*	-	-	-	KJ417684*
F. luschnatiana (Miq.) Miq	N. Rønsted 151 (K)	Cult. (BG) 1999-1019. Brazil	AY730082	AY730170	EF092345	-	-	KJ417689*

F. nymphaeifolia Mill.	N. Rønsted 144 (K)	Cult. (BG) 1988.0297. Orig. acc. Utrecht Bot. Gard.	AY063566	AY063527	EU089843	-	-	KJ417787*
F. obtusifolia Kunth.			AY730084	AY730172	AY967949	-	-	-
F. palmeri S.Watson			AY730085	AY730173	-	-	-	-
F. paludica Standl.			*GQ504302*	GQ504283	-	-	-	-
F. paraensis Miq.	N. Rønsted 107 (K)	Cult. (BG) 1989-527. No prov.	AY730086	AY730174	AY967954	-	-	KJ417693*
F. pertusa L.			AF165400	AY730176	AY967950	-	-	-
F. petiolaris Kunth.			AY730088	AY730177	-	-	-	-
F. popenoi Standl.			EU081761	-	EU089842	-	-	-
F. schippii Standl.			AY730089	AY730178	-	-	-	-
F. schumacheri (Liebm.) Griseb.	N. Rønsted 123 (K)	Cult. (BG) 1988-274. Orig. acc. Utrecht Bot. Gard.	AY063567	AY063528	EF092346	DQ455616*	-	KJ417676*
***F. sphenophylla* Standley**	**N. Rønsted 98 (K)**	**Cult. (BG)** 1989.0533. Orig. acc. Utrecht Bot. Gard.	**EU091605***	**KP406979***	**-**	**-**	**-**	**-**
F. subandina Dugand	N. Rønsted 104 (K)	Cult. (BG) 1993-1674. No	DQ455668	DQ455687	EU087639*	DQ455617*	-	-

		prov.						
***F. tequendamae* Dugand**	**FB/S3754 (BR)**	**Cult. (BR) s.n.**	**EU091606***	**EU084437***	**-**	**-**	**-**	**-**
F. trigona L.	N. Rønsted 103 (K)	Cult. (BG) 1990-630. Ecuador	DQ455669	DQ455688	-	DQ455619	EU084368	KJ417682*
F. tuerckheimii Standl.	K. Oyama s.n. (no voucher information)	Mexico	EU091608	EU084438	EU087640	-	-	KJ417692*
F. turbinata Pittier			EU081769	-	EU089832	-	-	-
***Asperae* clade**								
F. ampelas Burm.f.	W. Takeuchi 14565 (K)	Papua New Guinea	EU091659	-	HQ890565	-	-	-
***F. aspera* G. Forst**	**N. Rønsted 62 (C)**	**Cult. (C) 1932-5073. No prov.**	**EU091660***	**EU084483***	**-**	**EU084345***	**-**	**-**
F. asperifolia Miq.			EU091661	EU084484	EF092394	-	-	-
F. coronata Spin.	N. Rønsted 71 (C)	Cult. (C) 1953-2381. No prov.	AY730131	AY730218	EF092396	-	EU084391*	KJ417792*
F. cyrtophylla Miq.	N. Rønsted 124 (K)	Cult. (BG) 1989.531. No prov.	EU091664	EU084488	-	EU084346	-	KJ417725*
F. exasperata Vahl	N. Rønsted 217 (K) /	Cameroon	EU091665	-	-	-	-	-
F. exasperata Vahl	N. Rønsted 97 (K)	Cult. (BG) 1992-118. India	-	EU084489	EU087683	EU084347	EU084392	-

***F. fiskei* Elmer**	**R.D. Harrison 613 (PUH)**	**Philippines**	**EU091666***	**EU084490***	**-**	**EU084348***	**-**	**KJ417722***
F. gul Laut. et K. Schum.			AY730132	AY730219	EF092397	EU084349	-	-
***F. heterophylla* L.f.**	**N. Rønsted 247 (HITBC)**	**Cult. (HITBC) s.n.**	**EU091667***	**EU084491***	**EU087684***	**-**	**-**	**KJ417731***
F. lateriflora Vahl			AY063585	AY063546	EF092398	-	-	-
***F. macrorrhyncha* Laut./Schum.**	**T8989 (MIN)**	**Papua New Guinea**	**EU091668***	**KP407005***	**-**	**-**	**-**	**-**
F. melinocarpa Blume	G.D. Weiblen 1705 (MIN)	Papua New Guinea	EU091669	KP407006*	EU087685	-	-	-
F. opposita Miq.			EU091670	-	EU087686	-	-	-
F. politoria Lam.	O. Maurin 74 (K)	Madagascar	EU091671	KP407007*	EU087687	-	-	-
F. pygmaea Hiern.			AY730134	AY730221	EF092399	EU084350	-	-
F. trachypison K. Schum.			EU091674	EU084493	EU087688	-	-	-
***Cariceae* clade**								
F. carica L.	G.D. Weiblen 1072 (MIN)	Cult. USA	EU091637	EU084464	EU087670*	-	EU084382	-
F. carica L.	N. Rønsted 96 (C)	Cult. (C) S1966-0166. No prov.	-	-	-	-	-	KJ417697*
F. iidaiana Rehder & E.H.Wilson			*HQ890765*	-	HQ890581	-	-	-

F. palmata Forssk.	FB/S22786 (BR)	Cult. (BR) s.n.	AY730125	AY730214	EF092383	EU084333*	EU084385*	-
***Conosycea* clade**								
F. altissima Blume	N. Rønsted 282 (HITBC)	Cult. (HITBC) s.n.	AY730064	AY730152	EU087621	-	EU084363	KJ417788*
***F. amplissima* J.E.Sm.**	**Matthew 20582 (K)**	**India**	**EU091577***	**-**	**-**	**-**	**-**	**-**
F. annulata Blume	N. Rønsted 293 (HITBC)	Cult. (HITBC) s.n.	EU091578	EU084417	EU087622	-	-	KJ417781*
F. archboldiana Summerh.	G.D. Weiblen 1695 (MIN)	Papua New Guinea	EU091579*	EU084418	-	-	-	KJ417776*
***F. balete* Mer.**	**R.D. Harrison 631 (PUH)**	**Philippines**	**EU091580***	**-**	**-**	**EU084304***	**-**	**-**
F. benghalensis L.			AY730065	AY730153	-	-	-	-
F. benjamina L.	N. Rønsted 81 (C)	Cult. (C) P1870-5193. No prov.	AY063559	AY063520	EF092333	EU084305	EU084364	KJ417721*
F. binnendijkii (Miq.) Miq.	M.W. Chase 19871 (K)	Cult. (K) 1992-883. No prov.	AY063561	AY063522	EF092334	EU084306*	-	-
***F. calcicola* Corner**	**N. Rønsted 172 (AAU/K)**	**Thailand**	**EU091581***	**EU084419***	**EU087623***	**-**	**-**	**-**
F. callophylla Blume	R.D. Harrison 601 (PUH)	Philippines	EU091582	KP406972*	-	-	-	KJ417778*
F. chrysolepis Miq.	G.D. Weiblen 2353 (MIN)		EU091583	EU084420	EU087624	-	-	KJ417720*
F. consociata Blume			AY063558	AY063519	-	-	-	-

F. cordatula Merr.	R.D. Harrison 606 (PUH)	Philippines	EU091584	EU084421	EU087625	EU084307	-	-
***F. crassiramea* (Miq.) Miq.**	**E. Jousselin 7**	**Brunei**	**EU091585***	**EU084422***	**-**	**-**	**-**	**-**
***F. cucurbitina* King**	**B. Chanterasuwan s.n.** (no voucher information)	**Thailand**	**EU091586***	**EU084423***	**-**	**-**	**-**	**-**
F. drupacea Thunb.	N. Rønsted 114 (K)	Cult. (BG) 1992-755. Orig. acc Paradenya Bot. Gard.	AY730066	AY730154	EF092335	DQ455632*	-	-
F. elastica Roxb.	M.W. Chase 19875 (K)	Cult. (K) 1969-13671. No Prov.	AY063555	AY063516	EF092338	DQ455633*	-	-
F. forstenii Miq.	R.D. Harrison 632 (PUH)	Philippines	EU091587	KP406973*	EU087626	-	-	-
F. glaberrima Blume	N. Rønsted 295 (HITBC)	Cult. (HITBC) s.n.	EU091588	KP406974*	EU087627	-	-	-
***F. globosa* Blume**	**B. Chanterasuwan s.n.** (no voucher information)	**Thailand**	**EU091589***	**KP406975***	**EU087628***	**-**	**-**	**-**
***F. kochummeniana* C. C. Berg**	**B. Chanterasuwan s.n.** (no voucher information)	**Thailand**	**EU091590***	**EU084424***	**-**	**-**	**-**	**-**
F. maclellandii King	N. Rønsted 281(HITBC)	Cult. (HITBC) s.n.	EU091591	EU084425	EU087629	-	EU084365	KJ417789*
F. menabeensis Perrier			AY730067	AY730155	-	-	-	-
F. microcarpa L. f.	N. Rønsted 77 (C)	Cult. (C). P1979-5041. No. prov.	AY063560	AY063521	EU087630*	EU084308*	-	-

F. microcarpa L. f.	N. Rønsted 286 (HITBC)	Cult. (HITBC) s.n.	-	-	-	-	-	KJ417719*
F. paracamptophylla Corner	E. Jousselin 6	Brunei	EU091592	EU084426	-	-	-	-
F. pellucidopunctata Griff.			AF165399	EU084427	-	-	-	-
***F. pubilimba* Merr.**	**B. Chanterasuwan s.n.** (no voucher information)	**Thailand**	**EU091593***	**KP406976***	**-**	**-**	**-**	**-**
F. rumphii Blume	N. Rønsted 74 (C)	Cult. (C). P1953-5105. No prov.	AY730063	AY730151	-	-	-	KJ417785*
F. spathulifolia Corner	G.D. Weiblen 929 (MIN)	Papua New Guinea	EU091594	EU084428	EU087631	-	-	KJ417690*
F. stricta (Miq.) Miq.	N. Rønsted 288 (HITBC)	Cult. (HITBC) s.n.	EU091595	EU084429	EU087632	-	-	KJ417752*
***F. subcordata* Blume**	**B. Chanterasuwan s.n.** (no voucher information)	**Thailand**	**EU091596***	**EU084430***	**EU087633***	**-**	**-**	**KJ417704***
F. subgelderi Corner			AY064556	AY063517	EF092336	-	-	-
F. sumatrana (Miq.) Miq.	R.D. Harrison 629 (PUH)	Philippines	EU091597	KP406977*	EU087634	-	-	-
F. sundaica Blume	G.D. Weiblen 906 (A/LAE)	Borneo	AY730068	AY730156	EF092337	-	-	KJ417717*
F. xylophylla (Wall ex Miq) Miq			AY063557	AY063518	-	-	-	-
***Eriosycea* clade**								
F. aurata Miq.	E. Jousselin 1	Brunei	EU091642	KP406999*	-	-	-	-
F. chartacea King	N. Rønsted 170 (AAU/K)	Thailand	AY730126	EU084470*	EF092384	-	-	-

***F. endospermifolia* Corner**	**G.D. Weiblen 2282 (MIN)**	**Borneo**	**EU091643***	-	-	-	-	-
***F. eumorpha* Corner**	**G.D. Weiblen 2277 (MIN)**		**EU091644***	**EU084471***	**EU087674***	**EU084334***	-	-
F. fulva Reinw. ex Blume	F. Kjellberg 1998-79	Brunei	EU091645	KP407000*	EU087675	EU084335	-	-
***F. glandulifera* (Wa.exMiq) King**	**B. Chanterasuwan s.n.** (no voucher information)	**Thailand**	**EU091646***	**EU084472***	-	-	-	-
F. grossularioides Burm.	E. Jousselin s.n.	Brunei	AY063591	-	EF092385	EU084336	EU084386	KJ417747*
F. hirta Vahl	N. Rønsted 168 (AAU/K)	Thailand	AY730127	EU084473	EF092386	-	-	KJ417760*
F. padana Burm.			AF165398	-	EF092387	-	-	-
F. ruficaulis Merr.	R.D. Harrison 626 (PUH)	Philippines	EU091647	KP407001*	-	EU084337	-	KJ417696*
***F. setiflora* Stapf.**	**G.D. Weiblen 2290 (MIN)**	**Malaysia**	**EU091648***	**KP407002***	-	-	-	-
***Frutescentiae* clade**								
Ficus abelii Miq.			*JF976314*	-	-	-	-	-
F. boninsimae Koidz.			*AB485917*	-	HQ890580	-	-	-
F. chapaensis Gagnepain			EU091638	EU084465	EU087671	-	-	-
F. deltoidea Jack.	N. Rønsted 73 (C)	Cult. (C) E1859-0015. No prov.	AY063579	AY063540	EF092378	-	-	KJ417797*

F. erecta Thunberg			AY730121	AY730211	EF092379	EU084330	-	-
F. formosana Maxim.			*HQ890690*	-	HQ890585	-	-	-
F. ischnopoda Miq.	N. Rønsted 175 (AAU/K)	Thailand	AY730122	AY730212	EF092380	-	EU084383	KJ417713*
F. nishimurae Koidz.			*AB485915*	-	HQ890579	-	-	-
F. oleifolia King	G.D. Weiblen 2287 (MIN)	Borneo	AY730124	EF092322	EF092382	EU084332	EU084384	-
F. pedunculosa Miq.			*HQ890687*	-	HQ890589	-	-	-
F. stenophylla Hemsl.			EU091640	EU084467	-	-	-	-
F. tannoensis Hayata			*HQ890708*	-	HQ890593	-	-	-
F. vaccinioides King			*HQ890693*	-	HQ890591	-	-	-
F. variolosa Lindl. ex Benth.			*JF976332*	-	-	-	-	-
***Galoglychia* clade**								
F. abutilifolia (Miq.) Miq.	F. Forest 326 (NBG)	Cult. (NBG) 508/55	AY730091	AY730180	EF092348	EU084314*	-	KJ417669*
F. amadiensis De Wild.	S. Bidgood et al. 3518 (K)	Tanzania	DQ455646	-	-	EU084315*	-	-
F. ardisioides Warb.	N. Rønsted 222 (K)	Cameroon	DQ455655	DQ455678	EU087641*	-	-	-
F. bizanae Hutch/Burtt-Davy	F. Forest 334 (NBG)	Cult. (NBG) 186/82	DQ455636	DQ455670	-	-	-	KJ417724*
F. bubu Warb.	F. Forest 339 (NBG)	Cult. (NBG) 430-84	DQ455637	DQ455671	EU087642	DQ455628	-	KJ417743*
F. burkei (Miq.) Miq.	F. Forest 345 (NBG)	Cult. (NBG) 509-77	AY730095	AY730184	-	DQ455621	EU084369	-
F. burtt-davyi Hutch.	F. Forest 328 (NBG)	Cult. (NBG) 218/83	DQ455647	DQ455675	EU087643	-	-	KJ417735*

F. calyptrata Vahl	N. Rønsted 133 (K)	Cult. (BG) 1988-236. Orig. acc. Utrecht Bot. Gard.	DQ455648	DQ455676	EU087644*	-	-	KJ417700*
F. conraui Warb.			DQ455656	-	-	-	-	-
F. craterostoma Mildbr./Burret	F. Forest 340 (NBG)	Cult. (NBG) 46/79	AY730097	AY730186	EF092349	DQ455622	-	-
F. cyathistipula Warb.			DQ455657	DQ455679	-	-	-	-
F. cyathistipuloides De Wild.	N. Rønsted 136 (K)	Cult. (BG) 1988-239. Orig. acc. Utrecht Bot. Gard.	AY063563	AY063524	EU087645	-	-	KJ417759*
F. densistipulata De Wild.			DQ455659	DQ455680	-	-	-	-
F. dryepondtiana De Wild.			DQ455638	-	-	-	-	-
F. elasticoides De Wild.	N. Rønsted 128 (K)	Cult. (BG) 1994-426. Côte d'Ivoire	AY730103	AY730192	EF092354	DQ455624*	-	KJ417678*
F. faulkneriana C.C. Berg	Luke & Mbinda 5824 (K)	Kenya	DQ455645	KP406981*	-	-	-	-
F. fischeri Mildbr./Burret	F. Forest 327 (NBG)	Cult. (NBG) s.n.	DQ455649	AY730187	EF092350	DQ455623	-	KJ417802*
F. glumosa Delile			AY063562	AY063523	-	EU084316	-	-
F. ilicina (Sond.) Miq.			*HM746960*	HM746957	-	-	-	-
F. jansii Boutique			*GQ504310*	JQ504292	-	-	-	-
F. lingua DeWild/Durrand	N. Rønsted 208 (K)	Cameroon	AY730099	AY730188	EF092351	-	-	KJ417716*

F. lutea Vahl	N. Rønsted 87 (C)	Cult. (C) P1928-5257. No prov.	AY063564	AY063525	EF092347	-	-	KJ417687*
F. lyrata Warb.	N. Rønsted 85 (C)	Cult. (C) P1909-5032. No prov.	AY730104	AY730193	-	-	-	KJ417806*
F. natalensis Hochst.	F. Forest 333 (NBG)	Cult. (NBG) 386/83	AY730100	AY730189	EF092352	-	-	KJ417733*
F. nigropunctata Mildbr./Burret			DQ455663	-	-	-	-	
F. oreodryadum Mildbr.	N. Rønsted 228 (K)	Cameroon	DQ455652	KP406982*	-	-	-	KJ417774*
F. ottoniifolia (Miq.) Miq.	N. Rønsted 117 (K)	Cult (BG) 1988-263. Orig. acc. Utrecht Bot. Gard.	AY730109	AY730198	EF092358	-	-	KJ417800*
F. ovata Vahl	N. Rønsted 132 (K)	Cult. (BG) 1994-425. Côte d'Ivoire.	DQ455640	DQ455672	-	-	-	KJ417705*
F. polita Vahl	N. Rønsted 150 (K)	Cult. (BG) 1988-0259. Orig. acc. Utrecht Bot. Gard.	DQ455642	DQ455673	-	-	-	KJ417673*
F. populifolia Vahl	Thulin & Warfa 5542 (K)	Somalia	AY730093	AY730182	-	-	-	KJ417751*
F. preussii Warb.	N. Rønsted 138 (K)	Cult. (BG) 1992-1123. Orig. acc. Turku Bot. Gard.	AY730105	AY730194	EF092355	DQ455625	-	KJ417750*

F. pseudomangifera Hutch.			GQ504309	GQ504291	-	-	-	-
F. reflexa Thunb.	O. Maurin 76 (K)	Madagascar	DQ455650	GQ504278	EU087646	-	-	KJ417754*
F. sagittifolia Mildbr. & Burret			AY730106	AY730195	EF092356	DQ455626	-	
F. sansibarica Warb.	N. Rønsted 127 (K)	Cult. (BG) 1988-271. Orig. acc. Utrecht Bot. Gard.	AY730110	AY730199	EF092359	-	-	KJ417746*
F. saussureana DC.			AY730090	AY730179	-	-	-	-
F. scassellatii Pamp.	N. Rønsted 110 (K)	Cult. (BG) 1988-293. Orig. acc. Utrecht Bot. Gard.	AY730107	AY730196	EF092357	DQ455627*	EU084370*	-
F. tesselata Warb.	N. Rønsted 143 (K)	Cult. (BG) 1988-281. Orig. acc. Utrecht Bot. Gard.	DQ455662	DQ455682	EU087647	-	-	KJ417794*
F. tettensis Hutch.			DQ455665	DQ455683	-	DQ455620	-	-
F. thonningii Bl.	F. Forest 341 (NBG)	Cult. (NBG) 534/77	AY730102	AY730191	EF092353	-	-	KJ417775*
F. tremula Warb.	F. Forest 331 (NBG)	Cult. (NBG) 544/87	AY730111	AY730200	-	-	-	KJ417712*
F. trichopoda Baker	N. Rønsted 118 (K)	Cult. (BG) 1988-267. No prov.	DQ455666	DQ455684	EU087648	-	-	KJ417772*
F. umbellata Vahl	FB/S22813 (BR)	Cult. (BR) s.n.	DQ455644	DQ455674	-	DQ455629	-	KJ417736*

F. usambarensis Warb.			DQ455653	DQ455677	-	-	-	-
F. wildemaniana DeWild/Durand			AY730108	AY730197	-	-	-	-
***Hemicardia* clade**								
F. prostrata (Wall ex. Miq.) Miq.	N. Rønsted 268 (HITBC)	Cult. (HITBC) s.n.	EU091612	KP406997*	-	-	-	-
F. semicordata Buch-Ham ex Sm	B. Chanterasuwan s.n. (no voucher information)	Thailand	EU091613	-	EU087652	EU084322	-	KJ417777*
F. semicordata Buch-Ham ex Sm	N. Rønsted 267 (HITBC)	Cult. (HITBC) s.n.	-	KP406998*	-	-	-	-
***Malvanthera* clade**								
F. atricha D.J. Dixon			AY730116	AY730206	EF538786	-	-	-
F. augusta Corner			EF545651	EF538767	EF538787	-	-	-
F. brachypoda (Miq.) Miq.	Dixon 1101 (DNA)	Australia	EF545652	EF538768	EF538788	EU084309	-	KJ417779*
F. crassipes Bailey			AY730112	AY730201	EF538789	-	-	-
F. destruens C.T. White			AF165384	EF538769	EF538790	-	-	-
F. glandifera Summerh.			AY730113	AY730202	EF092361	-	-	-
F. hesperidiiformis King			AF165387	AY730203	EF092362	-	-	-
F. heteromeka Corner			EF545656	EF538772	EF538791	-	-	-
F. liliputiana D.J. Dixon			EF545657	EF538773	-	-	-	-
F. macrophylla DesFicus ex. Pers.			AY730115	AY730205	EF538792	-	-	-
F. obliqua G. Forst.	J.M. Cook 2003-6	Australia	EF545659	EF538774	EF538793	-	-	KJ417796*

F. platypoda D. J. Dixon			AY730114	AY730204	EF538794	-	-	-
F. pleurocarpa Ficus Muell.			AY063568	AY063529	EF538795	DQ455634	-	-
F. rubiginosa Desf. ex. Vent.	N. Rønsted 89 (C)	Cult. (C). E1859-0014. No prov.	AY063569	AY063530	EF092363	DQ455635	EU084366	KJ417762*
F. sterrocarpa Diels			EF545665	EF538779	EF538796	-	-	-
F. triradiata Corner	Hayland 8336 (K)	Australia	AY730117	AY730207	EF092364	-	-	KJ417742*
F. watkinsiana F.M. Bailey	N. Rønsted 83 (C)	Cult. (C). S1959-1912. No prov.	AY730118	AY730208	EF092365	EU084310	EU084367	KJ417730*
F. xylosycia Diels			AF165419	EF538784	-	-	-	-
***Oreosycea* clade**								
F. callosa Willd.	N. Rønsted 109 (K)	Cult. (BG) 1999-561. Indonesia.	AY063565	AY063536	EF092367	EU084299*	EU084356*	-
F. edelfeltii King			EF165385	AY730209	-	-	-	-
F. habrophylla Seem.	G.D. Weiblen 1224 (MIN)	New Caledonia	EU091567	EU084408	EU087612	-	EU084358*	KJ417714*
***F. hadroneura* Diels**	**G.D. Weiblen 1857 (MIN)**	**Papua New Guinea**	**EU091568***	**-**	**EU087613***	**-**	**-**	**-**
F. hombroniana Corner	G.D. Weiblen 1859 (MIN)	Papua New Guinea	pending	KP406970*	EF092369	-	-	KJ417685*

***F. magnoliifolia* Blume**	**R.D. Harrison 619 (PUH)**	**Philippines**	**EU091569***	**EU084409***	**EU087614***	**-**	**EU084359***	**KJ417767***
F. nervosa Roth	B. Chanterasuwan s.n. (no voucher information)	Thailand	EU091570	EU084410	EU087615	-	-	KJ417698*
F. polyantha Corner	G.D. Weiblen 2174 (MIN)	Papua New Guinea	EU091571	KP406971*	EU087616	-	-	KJ417707*
F. pseudojaca Corner	G.D. Weiblen 2341 (MIN)		EF092317	EF092320	EF092370	-	-	-
F. racemigera Bureau			AY063587	AY063554	-	-	-	-
F. subtrinervia Laut./K.Schum.	G.D. Weiblen 1543 (MIN)		AY730119	EU084411	EU087617	-	-	-
***F. vasculosa* Miq.**	**B. Chanterasuwan s.n.** (no voucher information)	**Thailand**	**EU091572***	**EU084412***	**-**	**EU084302***	**-**	**KJ417782***
F. vasculosa Miq.	W.L. Clement 27 (MIN)	Borneo	-	-	-	-	EU084361*	-
***Paleomorphe* clade**								
***F. heteropleura* Bl.**	**R.D. Harrison 599 (PUH)**	**Philippines**	**EU091663***	**EU084487***	**EU087682***	**-**	**-**	**KJ417718***
F. obscura Blume	R.D. Harrison 602 (PUH)	Philippines	EU091676	-	EU087689	-	-	KJ417683*
F. parietalis Bl.	R. de Kok 1026 (K)	Borneo	AY063583	AY063544	EF092401	-	-	KJ417695*

F. subulata Blume	B. Chanterasuwan s.n. (no voucher information)	Thailand	EU091677	EU084495	EU087690	-	-	KJ417770*
F. tinctoria Forst.f.	N. Rønsted 99 (K)	Cult. (BG) 1989.551. No prov.	AF165413	AY730223	EF092403	-	-	KJ417756*
F. virgata Reinv. ex Blume	N. Rønsted 65 (C)	Cult. (C) P1986-5148. No prov.	AF165417	AY730224	EF092404	EU084351	EU084393	KJ417670*
***Phaeopilosae* clade**								
***F. badiopurpurea* Diels**	**G.D. Weiblen 2010 (MIN)**	**Papua New Guinea**	**EU091662***	**EU084485***	**EU087681***	**-**	**-**	**-**
F. conocephalifolia Ridley	G.D. Weiblen 1754 (MIN)	Papua New Guinea	AF165381	EU084486	-	-	-	KJ417739*
F. copiosa Steud.	G.D. Weiblen 57 (A)	Papua New Guinea	AF165382	EF092324	EF092395	-	EU084390	-
F. phaosyce Laut. et K. Schum.			AF165401	-	-	-	-	-
***F. poropyrochaete* Corner**	**CR1428 (MIN)**	**Papua New Guinea**	**EU091672***	**-**	**-**	**-**	**-**	**-**
***F. primaria* Corner**	**G.D. Weiblen 2375 (MIN)**	**Papua New Guinea**	**EU091673***	**EU084492***	**-**	**-**	**-**	**-**
F. wassa Roxb.			AF165418	EF092325	DQ367635	-	DQ367655	-
***Pharmacosycea* clade**								
F. adhatodifolia Schott	N. Rønsted 148 (K)		EU091563	EU084404	EU087608	-	-	-

F. glabrata Kunth.			AY063593	AY063550	AY967960	-	-	-
F. insipida Willd.	N. Rønsted 119 (K)	Cult. (BG) 1989-523. No prov.	AY063592	AY063549	AY967961	EU084296	EU084354	KJ417740*
F. lapathifolia (Liebm.) Miq.	K. Oyama s.n. (no voucher information)	Mexico	EU091564	EU091564	EU091564	-	-	KJ417715*
F. maxima Mill.	W.L. Clement 215 (MIN)	Ecuador	AY063595	AY063551	AY967958	-	-	KJ417763*
***F. piresiana* Avila & C.C.Berg**	**W.L. Clement 216 (MIN)**	**Ecuador**	**EU091565***	**KP406969***	**EU087610***	**-**	**-**	**KJ417807***
F. tonduzii Standl.	W.L. Clement 188 (MIN)	Ecuador	AY730140	AY730230	EU087611	EU084297	-	KJ417780*
F. yoponensis Desvaux	K. Oyama (no voucher information)	Mexico	AY063594	AY063552	AY967959	-	-	KJ417679*
***Sinosycidium* clade**								
F. henryi Diels			EU091639	EU084466	EU087672	EU084331	-	-
***F. tsiangii* Merr. ex Corner**	**N. Rønsted 298 (HITBC)**	**Cult. (HITBC) s.n.**	**EU091675***	**EU084494***	**-**	**-**	**-**	**KJ417726***
***Synoecia* clade**								
***F. apiocarpa* (Miq.) Miq.**	**B. Chanterasuwan s.n.** (no voucher information)	**Thailand**	**EU091655***	**EU084480***	**-**	**EU084341***	**-**	**-**
F. bauerlenii King			AY063377	EU084474	-	-	-	-

***F. disticha* Bl.**	**G.D. Weiblen 625 (MIN)**	**Papua New Guinea**	**EU091656***	**EU084481***	**-**	**EU084342***	**-**	**-**
F. diversiformis Miq.			AY730128	AY730215	EF092392	-	-	-
F. jimiensis C. C. Berg	G.D. Weiblen 792 (A/LAE)	Papua New Guinea	AY730129	AY730216	EF092388	-	-	KJ417694*
F. odoardii King			AF165397	-	EF092389	-	-	-
F. pantoniana King	G.D. Weiblen 2079 (MIN)	Papua New Guinea	EU091649	KP407004*	-	-	-	KJ417791*
***F. pleiadenia* Diels**	**G.D. Weiblen 2315 (MIN)**	**Papua New Guinea**	**EU091650***	**EU084475***	**EU087676***	**EU084338***	**EU084387***	**KJ417745***
***F. pubigera* (Wall. ex Miq.) Miq.**	**N. Rønsted 270 (HITBC)**	**Cult. (HITBC) s.n.**	**EU091651***	**EU084476***	**EU087677***	**-**	**-**	**KJ417702***
F. punctata Thunb.			AF165403	AY063545	-	EU084343	-	-
F. ruginervia Corner			AF165407	EF092323	EF092393	-	-	-
F. sagittata J. König ex Vahl	N. Rønsted 266 (HITBC)	Cult. (HITBC) s.n.	EU091652	EU084477	EU087678	EU084339	-	KJ417699*
***F. sarawakensis* Corner**	**E. Jousselin s.n.**	**Brunei**	**EU091657***	**KP407003***	**-**	**-**	**-**	**KJ417729***
F. sarmentosa Buch-Ham ex Sm.	N. Rønsted 263 (HITBC)	Cult. (HITBC) s.n.	EU091653	EU084478	EU087679	-	-	KJ417675*
***F. scratchleyana* King**	**G.D. Weiblen 1834 (MIN)**	**Papua New Guinea**	**EU091658***	**-**	**EU087680***	**EU084344***	**-**	**-**
***F. trichocarpa* Blume**	**B. Chanterasuwan s.n.**	**Thailand**	**EU091654***	**-**	**-**	**-**	**-**	**KJ417691***

	(no voucher information)							
F. villosa Bl.	M.W. Chase 19851 (K)	Cult. (K) 1984-2930. Indonesia.	AY730130	AY730217	EF092391	EU084340	EU084389	KJ417753*
***Sycocarpus* clade**								
F. adelpha Laut. et K. Schum.	G.D. Weiblen 1689 (MIN)	Papua New Guinea	DQ367656	-	DQ367615	-	-	KJ417744*
F. arfakensis King	G.D. Weiblen 1726 (MIN)		DQ367657	-	DQ367617	-	DQ367637	-
F. aurantiacafolia Weiblen & Whitfeld	G.D. Weiblen 1545 (MIN)	Papua New Guinea	DQ367662	-	DQ367632	-	DQ367652	-
F. beccarii King			EU091621	-	EU087658	-	-	-
F. benguetensis Merr.			*AB485854*	-	HQ890561	-	-	-
F. bernaysii King			AF165378	-	DQ367618	-	DQ367638	-
F. botryocarpa Miq.	G.D. Weiblen D3 (A)	Papua New Guinea	AF165379	-	DQ367619	-	DQ367639	KJ417790*
F. calopilina Diels	G.D. Weiblen 2135 (MIN)	Papua New Guinea	EU091622	-	-	-	-	-
F. condensa King	E. Jousselin s.n.	Brunei	AY063577	KP406991*	-	EU084325	-	-
F. congesta Roxb.	G.D. Weiblen B1 (A)	Papua New Guinea	AY730136	-	DQ367620	-	DQ367640	KJ417703*
F. fistulosa Reinw. ex Bl.	N. Rønsted 111 (K)	Cult. (BG) s.n.	-	KP406992*	-	-	-	-

F. fistulosa Reinw. ex Bl.	R.D. Harrison (PUH)		AY730137	-	EF092375	-	EU084379	-
F. hispida L. f.	N. Rønsted 88 (C)	Cult. (C) E1859-0002	EU091623	KP406993*	EU087659	EU084326	-	KJ417711*
F. hispidioides S. Moore	G.D. Weiblen B2 (MIN)	Papua New Guinea	AF165388	-	DQ367622	-	DQ367642	KJ417668*
F. iodotricha Diels			EU091624	-	EU087660	-	-	-
F. lepicarpa Bl.	N. Rønsted 162 (AAU/K)	Thailand	AY730138	KP406994*	EF092376	-	-	KJ417741*
F. megaleia Corner			EU091625	-	EU087661	-	-	-
F. minnahassae (Teijs/Vrie) Miq			-	-	EU087662	-	EU084380	-
F. morobensis C. C. Berg	G.D. Weiblen 2228 (MIN)	Papua New Guinea	DQ367659	-	DQ367624	-	DQ367644	-
F. nota (Blanco) Merr.			EU091626	-	EU087663	EU084327	-	-
F. obpyrimidata King	B. Chanterasuwan s.n. (no voucher information)	Thailand	EU091627	KP406995*	-	-	-	-
F. pachyrrachis Lauterb./Schum.	G.D. Weiblen 2377 (MIN)	Papua New Guinea	EU091628	-	DQ367626	EU084328	DQ367646	-
F. pungens Reinw. ex Blume	G.D. Weiblen 1756 (MIN)	Papua New Guinea	AF165404	-	DQ367627	-	DQ367647	KJ417732*
F. ribes Reinw. ex Blume			EU091630	-	EU087665	-	-	-
F. satterthwaitei Elmer	G.D. Weiblen 2100 (MIN)	Philippines	EU091631	-	EU087666	-	-	-

F. saurauroides Diels	G.D. Weiblen 2006 (MIN)	Papua New Guinea	EU091632	-	EU087667	-	EU084381	KJ417801*
F. rosulata Berg			EU091633	-	-	-	-	-
F. scortechinii King	N. Rønsted 167 (AAU/K)	Thailand	AY730139	KP406996*	EF092377	-	-	-
F. septica Burm.f.	G.D. Weiblen 836 (HON)	Solomon Islands	AF165409	-	DQ367630	-	DQ367650	KJ417786*
F. squamosa Roxb.	N. Rønsted 262 (HITBC)	Cult. (HITBC) s.n.	EU091634	-	-	-	-	KJ417680*
F. stolonifera King			EU091635	-	-	-	-	-
F. theophrastoides Seem.			AF165412	-	-	-	-	-
F. treubii King			EU091636	-	EU087668	-	-	-
F. uncinata King			AY063576	-	EU087669	-	-	-
***Sycomorus* unresolved section**								
F. auriculata Lour.	D. Dewsnap (MIN)	Cult. Florida	AF165376	FJ812281	EU087653	-	EU084374	-
F. botryoides Baker			AF165380	-	-	-	-	-
F. hainanensis Merrill & Chun			EU091614	-	-	-	-	-
F. mauritiana Lam.	Fournel & Michenaud JF89 (REU) 10241)	Réunion Island	AY063570	AY063531	EF092371	-	-	KJ417723*
F. mucuso Ficalho			AY730120	AY730210	EF092372	EU084317	-	-
F. nodosa Teysm. et Binn.	G.D. Weiblen 2175 (MIN)	Papua New Guinea	AF165395	-	DQ367625	-	DQ367645	KJ417803*

F. nodosa Teysm. et Binn.	R.D. Harrison 603 (PUH)	Philippines	-	KP406983*	-	-	-	-
F. oligodon Miq.	N. Rønsted 306 (HITBC)	Cult. (HITBC). s.n.	EU091615	KP406984*	-	-	-	-
F. racemosa L.	N. Rønsted 116 (K)	Cult. (BG) 1989-235. No prov.	AF165405	KP406985*	-	EU084318	EU084371	KJ417769*
F. robusta Corner	N. Rønsted 952 (A)	Papua New Guinea	AF165406	EU084442	DQ367628	-	DQ367648	KJ417795*
F. semivestita Corner	G.D. Weiblen 2380 (MIN)	Papua New Guinea	EU091616	EU084443	DQ367629	-	DQ367649	KJ417727*
F. sur Forssk.	N. Rønsted 76 (C)	Cult. (C) 1992-0213. No prov.	AF165411	KP406986*	EU087649	EU084319	EU084372	KJ417748*
F. sycomorus L.	N. Rønsted 72 (C)	Cult. (C). P1965-5118. No prov.	AY063575	AY0635362	-	EU084320	-	KJ417798*
F. tiliifolia Baker			EU091609	EU084439	-	-	-	-
F. vallis-choudae Delile	N. Rønsted 234 (K)	Cameroon	AY063574	AY063535	EF092373	EU084321	EU084373	KJ417701*
F. vogeliana (Miq.) Miq.	N. Rønsted 201 (K)	Cameroon	EU091610	EU084440	EU087650		-	KJ417688*
F. caulocarpa (Miq.) Miq.	G.D. Weiblen 2384 (MIN)	Papua New Guinea	EU091573	EU084413	EU087619	-	-	KJ417708*
F. concinna (Miq.) Miq.	N. Rønsted 180 (BG)	Thailand	AY730059	AY730145	EF092328	-	-	KJ417805*
F. cordata Thunb.	KBG 510 (NBG)	Cult. (NBG). s.n.	AY730060	AY730146	EF092329	-	EU084362*	-
F. cordata Thunb.	F. Forest 329 (NBG)	Cult. (NBG) s.n.	-	-	-	DQ455630*	-	-

***F. densifolia* Miq.**	**J. Fournel 117 (REU)**	**Mauritius**	**EU091574***	**EU084414***	**-**	**-**	**-**	**KJ417757***
F. ingens (Miq.) Miq.			AY730061	AY730147	EF092330	EU084303	-	-
F. lacor Buch.-Ham.			AY730062	AY730148	-	-	-	-
***F. orthoneura* H. Lév & Vaniot**	**N. Rønsted 291 (HITBC)**	**Cult. (HITBC) s.n.**	**EU091575***	**EU084415***	**EU087620***	**-**	**-**	**KJ417768***
F. prasinicarpa Elmer	G.D. Weiblen 827 (HON)	Solomon Islands	-	-	-	-	-	KJ417799*
F. prasinicarpa Elmer			AF165402	-	-	-	-	-
F. prolixa G. Forst.	G.D. Weiblen 1218 (MIN)	Noumea BG, New Caledonia	AY063581	AY063542	-	-	-	KJ417674*
F. religiosa L.	N. Rønsted 86 (C)	Cult. (C). P1951-5144. No prov.	AY063582	AY063543	EF092331	-	-	KJ417677*
F. superba (Miq.) Miq.			AF165410	AY730149	EF092332	DQ455631	-	
***F. verruculosa* Warb.**	**N. Rønsted 115 (K)**	**Cult. (BG) 1990-1239. South Africa.**	**EU091576***	**EU084416***	**-**	**-**	**-**	**KJ417686***
F. virens Aiton	N. Rønsted 137 (BG)		AF165416	AY730150	DQ367634	-	DQ367654	KJ417671*
Helicostylis								
***H. heterotricha* Ducke**	**Rudas & Aguila Joaquin 1614 (MO)**	**Colombia**	**Pending***	**-**	**-**	**-**	**-**	**-**
H. pedunculata Benoist	G.D. Weiblen 1680	French Guiana	Pending*	-	Pending*	-	Pending*	-

	(MIN)							
H. tomentosa (Poepp. & Endl.) J.F.Macbr.	G.D. Weiblen 1475 (MIN)	Peru	Pending*	-	Pending*	-	Pending*	-
***H. tovarensis* (Klotzsch & H.Karst.) C.C.Berg**	**Bascopé et al. 82 (MO)**	**Bolivia**	**Pending***	**-**	**Pending***	**-**	**-**	**-**
***H. turbinata* C.C.Berg**	**Ribiero 1850 (F)**	**Brazil**	**Pending***	**-**	**-**	**-**	**-**	**-**
Maquira								
***M. calophylla* (Poepp. & Endl.) C.C.Berg**	**W.L. Clement 66 (MIN)**	**Ecuador**	**Pending***	**-**	**Pending***	**-**	**Pending***	**-**
***M. coriacea* (H.Karst.) C.C.Berg**	**Solomon 6450 (MO)**	**Bolivia**	**Pending***	**-**	**-**	**-**	**-**	**-**
***M. guianensis* Aubl.**	**G.D. Weiblen 1439 (MIN)**	**Panama**	**Pending***	**-**	**Pending***	**-**	**Pending***	**-**
***M. sclerophylla* (Ducke) C.C.Berg**	**Costa 23 (F)**	**Brazil**	**Pending***	**-**	**-**	**-**	**-**	**-**
Mesogyne								
M. insignis Engl.	M.A. Mwangoka 161 (MO)	Tanzania	Pending*	-	Pending*	-	-	-
Naucleopsis								
N. caloneura (Huber) Ducke	G.D. Weiblen 1517 (MIN)	Brazil	Pending*	-	Pending*	-	Pending*	-

***N. glabra* Spruce ex Pittier**	**W.L. Clement 60 (MIN)**	**Ecuador**	**Pending***	**-**	**Pending***	**-**	**Pending***	**-**
N. guianensis (Mildbr.) C.C.Berg	G.D. Weiblen 1683 (MIN)	French Guiana	Pending*	-	Pending*	-	Pending*	-
***N. imitans* (Ducke) C.C.Berg**	**W.L. Clement 175 (MIN)**	**Ecuador**	**Pending***	**-**	**Pending***	**-**	**Pending***	**-**
***N. imitans* (Ducke) C.C.Berg**	**W.L. Clement 161 (MIN)**	**Ecuador**	**Pending***	**-**	**-**	**-**	**Pending***	**-**
N. krukovii (Standl.) C.C.Berg	W.L. Clement 44 (MIN)	Ecuador	Pending*	-	Pending*	-	-	
N. naga Pittier	G.D. Weiblen 1404 (MIN)	Costa Rica	Pending*	-	Pending*	-	Pending*	-
N. ternstroemiiflora (Mildbr.) C.C.Berg	G.D. Weiblen 1518 (MIN)	Brazil	Pending*	-	Pending*	-	Pending*	-
N. ulei (Warb.) Ducke	G.D. Weiblen1509 (MIN)	Brazil	Pending*	-	Pending*	-	Pending*	-
Perebea								
P. angustifolia (Poepp. & Endl.) C.C.Berg	G.D. Weiblen 1403 (MIN)	Costa Rica	Pending*	-	Pending*	-	Pending*	-
***P. guianensis* Aubl.**	**G.D. Weiblen 1676 (MIN)**	**French Guiana**	**Pending***	**-**	**Pending***	**-**	**Pending***	**-**
P. humilis C.C. Berg	G.D. Weiblen 1468	Peru	Pending*	-	Pending*	-	Pending*	-

	(MIN)							
P. longepedunculata C.C. Berg	G.D. Weiblen 1477 (MIN)	Peru	Pending*	-	Pending*	-	Pending*	-
***P. mollis* (Poepp. & Endl.) Huber**	**W.L. Clement 218 (MIN)**	**Ecuador**	**Pending***	**-**	**Pending***	**-**	**Pending***	**-**
P. rubra (Trécul)C.C. Berg	G.D. Weiblen 1685 (MIN)	French Guiana	Pending*	-	Pending*	-	-	-
***P. tessmannii* Mildbr.**	**W.L. Clement 48 (MIN)**	**Ecuador**	**Pending***	**-**	**Pending***	**-**	**Pending***	**-**
P. xanthochyma H.Kartst.	W.L. Clement 80 (MIN)	Ecuador	Pending*	-	Pending*	-	Pending*	-
Poulsenia								
P. armata (Miq.) Standl.	G.D. Weiblen 1238 (MIN)	Panama	AY730144	AY730233	-	-	EU084353	-
P. armata (Miq.) Standl.	W.L. Clement 69 (MIN)	Ecuador	Pending*	-	Pending*	-	Pending*	-
Pseudolmedia								
***P. glabrata* (Liebm.)C.C. Berg**	**Brokaw 274 (MO)**	**Belize**	**Pending***	**-**	**Pending***	**-**	**-**	**-**
***P. hirtula* Kuhlm.**	**H.C. de Lima 2298**	**Brazil**	**Pending***	**-**	**Pending***	**-**	**-**	**-**
P. laevigata Trécul	W.L. Clement 71 (MIN)	Ecuador	Pending*	-	Pending*	-	Pending*	-
P. laevis (Ruiz & Pav.)J.F.Macbr.	W.L. Clement 57 (MIN)	Ecuador	Pending*	-	Pending*	-	Pending*	-
P. macrophylla Trécul	G.D. Weiblen 1472	Peru	Pending*	-	Pending*	-	Pending*	-

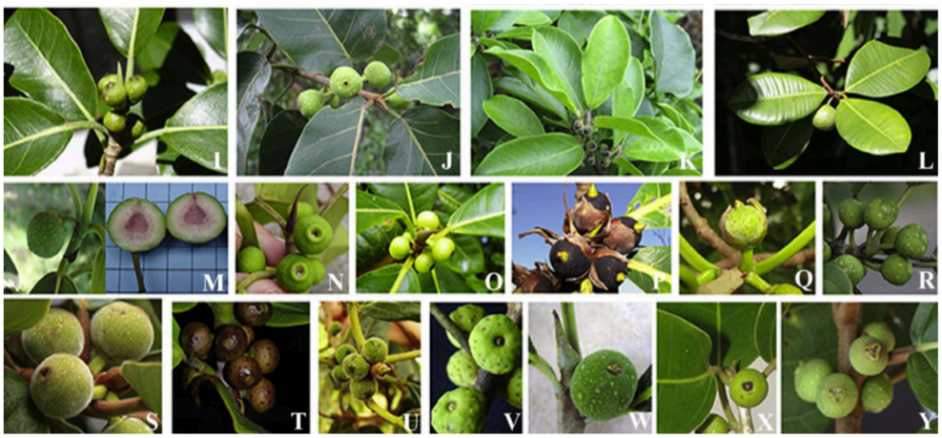
	(MIN)							
***P. rigida* (Klotzsch & H.Karst.)Cuatrec.**	**W.L. Clement 56 (MIN)**	**Ecuador**	**Pending***	**-**	**Pending***	**-**	**Pending***	**-**
P. spuria (Sw.) Griseb.	G.D. Weiblen 1427 (MIN)	Costa Rica	Pending*	-	Pending*	-	Pending*	-
Sparattosyce								
S. dioica Bur.	G.D. Weiblen 1223 (MIN)	New Caledonia	AY730141	AY730231	EU087607	-	-	KJ417734*

Supplementary Table S2. Amplification and sequencing primers used with ITS, *G3pdh,* and GBSSI. Grouped by gene region, the name, description, sequence, and reference is provided for each primer.

Gene	Name	Description	Sequence (5' - 3')	Reference
ITS	17SE	External amplification primer	ACGAATTCATGGTCCGGTGAAGTGTT	Sun et al. 2004
	26SE	External amplification primer	TAGAATTCCCCGGTTCGCTCGCCGTT	Sun et al. 2004
	ITS1R	Reverse internal sequencing primer	GCTACGTTCTTCATCGATGC	Modified from White et al. 1990
	ITS2F	Forward internal sequencing primer	GCATCGATGAAGAACGTAGC	Modified from White et al. 1990
	ITS4	Reverse amplification/sequencing primer	TCCTCCGCTTATTGATATGC	White et al. 1990
	ITS5	Forward amplification/sequencing primer	GGAAGTAAAAGTCGTAACAAGG	White et al. 1990
G3pdh	GPDX7F	Forward amplification primer	GATAGATTTGGAATTGTTGAGG	Strand et al. 1997
	GPDX9R	Reverse amplification/sequencing primer	AAGCAATTCCAGCCTTGG	Strand et al. 1997
	7FCast	Forward internal sequencing primer	TGAGGGTCTCATGACTACCG	This study
	338FCast	Forward internal sequencing primer	ACACAGGCCGTCGGAAAGGTTCTG	This study
	437R	Reverse internal sequencing primer	TTCTGAAGCCTGACAGTGAGG	This study
GBSSI	3F-Moraceae	Forward amplification primer	YAMAARMGMGGRGTTGATCG	Modified from Evans et al. 2000

	10R-Moraceae	Reverse amplification primer	GCAACTGAATGAGACCACA	This study
	375F-Moraceae	Forward internal sequencing primer	GCAGCTYTAGAAGYDCCAAGG	This study
	539R-Moraceae	Reverse internal sequencing primer	GGTARAAGGSYACTGTGCCAG	This study
	728F-Moraceae	Forward internal sequencing primer	CATACACAACATTGCYTACCAGG	This study
	942R-Castilleae	Reverse internal sequencing primer	CTTCACCGGTTTCTCATAGCTGC	This study
	1226F-Castilleae	Forward internal sequencing primer	GATATGATCAGGTGATGYCTGC	This study
	1348R-Moraceae	Reverse internal sequencing primer	ACHGGGATATTTCTATCYACTGGC	This study
At103	At103-F	Forward amplification primer	CTTCAAGCCMAAGTTCATCTTCTA	Li et al. 2008
	At103-R	Reverse amplification primer	TTGGCAATCATTGAGGTACATNGTMACATA	Li et al. 2008
ETS	ETS-Hel1	Forward amplification primer	GCTCTTTGCTTGCGCAACAACT	Baldwin & Markos 1998
	ETS-Fic1	New *Ficus* specific forward primer	GACCCTTGGTTCCTGTGTTGC	This study
	18S-ETS	Reverse amplification primer	ACTTACACATGCATGGCTTAATCT	Baldwin & Markos 1998
ncpGS	3F	Moraceae-specific forward amplification primer	GTTGTGATTWACCATGCT	Cruaud et al. 2012
	4R	Moraceae-specific reverse amplification primer	AGATTCAAAATCGCCTTC	Cruaud et al. 2012

Chapter II



Syconia diversity in Neotropical figs (from figure 6)

Atlantic forests to the all Americas: Biogeographical history and divergence times of Neotropical *Ficus* (Moraceae)

Molecular Phylogenetics and Evolution 122 (2018) 46–58

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Atlantic forests to the all Americas: Biogeographical history and divergence times of Neotropical *Ficus* (Moraceae)

Anderson Ferreira Pinto Machado[a,*], Nina Rønsted[b], Sam Bruun-Lund[b], Rodrigo Augusto Santinelo Pereira[c], Luciano Paganucci de Queiroz[a]

[a] *Programa de Pós-graduação em Botânica, Universidade Estadual de Feira de Santana, Av. Transnordestina s/n, Novo Horizonte, 44.036-900, Feira de Santana, BA, Brazil*

[b] *Natural History Museum of Denmark, University of Copenhagen, Sølvgade 83S, DK-1307 Copenhagen K, Denmark*

[c] *Departamento de Biologia/FFCLRP-Universidade de São Paulo, Av. Bandeirantes, 3900, 14040-901 Ribeirão Preto, SP, Brazil*

ARTICLE INFO

Keywords:
Americanae
Ancestral range estimation
Biogeography
Character evolution
Neotropics
Pharmacosycea

ABSTRACT

Ficus (Moraceae) is well diversified in the Neotropics with two lineages inhabiting the wet forests of this region. The hemiepiphytes of section *Americanae* are the most diversified with c. 120 species, whereas section *Pharmacosycea* includes about 20 species mostly with a terrestrial habit. To reconstruct the biogeographical history and diversification of *Ficus* in the Americas, we produced a dated Bayesian phylogenetic hypothesis of Neotropical *Ficus* including two thirds of the species sequenced for five nuclear regions (At103, ETS, G3pdh, ITS/5.8S and Tpi). Ancestral range was estimated using all models available in Biogeobears and Binary State Speciation and Extinction analysis was used to evaluate the role of the initial habit and propagule size in diversification. The phylogenetic analyses resolved both Neotropical sections as monophyletic but the internal relationships between species in section *Americanae* remain unclear. *Ficus* started their diversification in the Neotropics between the Oligocene and Miocene. The genus experienced two bursts of diversification: in the middle Miocene and the Pliocene. Colonization events from the Amazon to adjacent areas coincide with the end of the Pebas system (10 Mya) and the connection of landmasses. Divergence of endemic species in the Atlantic forest is inferred to have happened after its isolation and the opening and consolidation of the Cerrado. Our results suggest a complex diversification in the Atlantic forest differing between postulated refuges and more instable areas in the South distribution of the forest. Finally the selection for initial hemiepiphytic habit and small to medium propagule size influenced the diversification and current distribution of the species at Neotropical forests marked by the historical instability and long-distance dispersal.

1. Introduction

Neotropical rainforests are among the most diverse global biomes (Pennington et al., 2004) and exhibit higher biodiversity compared to African and Asian rainforests (Koenen et al., 2015). Explanations for the high diversity in rainforests have been subject to debate as to whether they are old or recent radiations (e.g. Antonelli and Sanmartín, 2011; Hughes et al., 2013; Pennington et al., 2015; Bruun-Lund et al., 2017). The "museum model" of diversification (Stebbins, 1974) explains their hyper diversity as old and evolving by constant rates of diversification and low rates of extinction with relative ecological stability over a long time. Another hypothesis considers rainforest as cradles of diversification with recent and elevated speciation rates (e.g. Richardson et al., 2001) and this so-called "cradle model" gained further support with the "refuge theory" suggesting that drier climates during glacial periods caused speciation by vicariance in rainforest species (Haffer, 1969; Prance, 1982; Carnaval and Moritz, 2008). The two models may not be mutually exclusive (Pennington et al., 2015) and consequently a recent hypothesis suggests that the diversity of rainforests could result from recent radiations from a large stock of higher-level taxa (Koenen et al., 2015).

Several studies of pantropical rainforest tree families have supported the museum (Couvreur et al., 2011; Wang et al., 2012), whereas generic level studies focused on the Neotropics have supported the cradle model. (Richardson et al., 2001; Kay et al., 2005; Erkens et al., 2007; Särkinen et al., 2007) suggesting that Neotropical rainforests may be cradles of recent diversity within a pantropical museum model.

Compared to other typical rainforest genera, *Ficus* L. possess unique diversification patterns. With approximately 750 species distributed in tropical and subtropical regions worldwide is one of the most important

* Corresponding author.
E-mail address: machadoafp@gmail.com (A.F.P. Machado).

https://doi.org/10.1016/j.ympev.2018.01.015
Received 7 April 2017; Received in revised form 20 January 2018; Accepted 20 January 2018
Available online 31 January 2018

Table 1
Main differences between the two Neotropical fig clades: *Ficus* section *Americanae* vs. *Ficus* section Pharmacosycea.

	Pharmacosycea	Americanae
Number of species	~20	~120
Distribution	Rainforests, rare in other biomes	All biomes
Initial habit/habit	Terrestrial (one case of hemi-epiphytic habit)/Tall trees	Hemiepiphytic, hemiepilhitic, rupicolous, terrestrial Shrubs, small trees or large "banyan trees"
Glandular spot(s)	A pair in the axils of basal lateral veins	One at the top of the petiole
Syconia position	Axillary and generally solitary	Axillary or along the branches. Generally in pairs or grouped
Syconia size and color	Large (> 2.5 cm diam.) rare small. Often green	Small (0.5–1 cm diam.); medium (1–2.4 cm diam.) or large (> 2.5 cm diam.). Yellow, red, purple
Basal bracts	3	2
Male flowers	2-stamens	1-stamen

plant genera in lowland tropical rainforests with high alpha-diversity (Harrison, 2005). A diversification rate analysis of *Ficus* by Bruun-Lund et al. (in press) found that they generally follow the museum model of evolution with a gradual accumulation of species over time and with very low extinction rates and no significant evolutionary shifts. However, a shift in diversification rate related to the lineage including Neotropical section *Americaneae* was detected, although with a non-significant probabability.

Neotropical *Ficus* appear clearly separated into two monophyletic sections (Rønsted et al., 2005, 2007, Rønsted et al., 2008a,b; Cruaud et al., 2012): a speciose section *Americanae* (Miq.) Corner (~120 spp.) consisting of hemi-epiphytic stranglers and section *Pharmacosycea* (Miq.) Benth. & Hook.f. (~20 spp.) including mostly large free-standing trees (Table 1; Fig. 6) (Berg and Simonis, 1981; Berg, 1989).

The morphological variation, complexity, and the massive diversification of *Ficus* in the Neotropics make this genus an important element for understanding diversification in species rich biomes such as the two major blocks of Neotropical rainforests, Amazonia (AM) and Atlantic Forest (AF).

Amazonian *Ficus* show greater morphological diversity compared to Atlantic Forest elements. The Amazonian figs comprise about 50 species of large and small trees with a great variety of leaves and syconia (fig inflorescences, propagule) sizes, but endemism in this area is low and most Amazonian species are also found in the Antilles, Mesoamerica and North America. Some taxa exhibit a disjunct distribution with the Atlantic rainforest (e.g. *F. castellviana* Dugand, *F. mariae* C.C.Berg et al., *F. pulchella* Schott, *F. trigona* L.f.) and others are widely distributed in the Americas (e.g. *F. gomelleira* Kunth, *F. obtusiuscula* (Miq.) Miq., *F. obtusifolia* Kunth., *F. pertusa* L.f. and *F. citrifolia* Mill.).

Although less diverse, the Atlantic Forest (with about 35 species; Carauta, 1989; BFG, 2015) the Atlantic forest harbors mostly endemic species of *Ficus* (e.g. *F. bahiensis* C.C.Berg & Carauta, *F. cestrifolia* Schott, *F. hirsuta* Schott, *F. enormis* Mart. ex Miq., *F. luschnathiana* (Miq.) Miq., and *F. guaranitica* Chodat), some also ocurring in adjacent areas.

In spite of the diversity and importance of *Ficus* in the Neotropics, phylogenetic studies have focused only on the small section *Pharmacosycea* (Honorio Coronado et al., 2014; Pederneiras et al., 2015; Costa et al., 2017). No published study to date has focused on the phylogeny and diversification of the megadiverse section *Americanae*. Consequently, little is known about the origin and diversification of Neotropical *Ficus* in general and this could improve our understanding of diversification patterns and processes in the tropical forests. About half of the fig species are hemiepiphytic woody plants (Berg and Corner, 2005; Harrison, 2005), a growth habit that could have evolved independently three times in *Ficus* (Jousselin et al., 2003), and which may have influenced the ability to diversify and colonize new areas in neotropical rainforests.

Likewise, the dispersal ability of propagules may have influenced the diversification of *Ficus* in the Neotropics. Two major morphofunctional propagule types are found in Neotropical *Ficus*. The mostly bat-dispersed type is characterised by larger, green or yellow syconia, which provide little color contrast against the background, but instead produce many volatile compounds and large peduncles evident among the leaves (Lomáscolo et al., 2008, 2010). In contrast, the mostly bird-dispersed syconia are small with bright colors (red, orange or purple, never green) at maturity, growing between the foliage and procucing few volatile compounds (Lomáscolo et al., 2010). Syconia with intermediate traits may be dispersed both by bats and birds (Lomáscolo et al., 2010). Species with small and intermediate syconia are expected to have higher dispersal range/efficiency due the bird dispersal effectiveness (Traveset, 1998; Jacomassa and Pizo, 2010; Lomáscolo et al., 2008, 2010).

A biogeographical analysis of *Ficus* in a phylogenetic context provides the opportunity to investigate the diversification patterns of the South American rainforests and its correlation with the main biogeographical events that might have influenced their high species diversity. From a comprehensively sampled dated phylogeny of Neotropical figs we adresse the following questions (1) When and where did the most recent common ancestor (MRCA) of both sections of Neotropical figs originate? (2) How is Neotropical *Ficus* phylogeny geographically structured in Amazonia and Atlantic rainforests? (3) How did geological events influence the divergence history of *Ficus*? Could the initial growth habit and the propagule traits have influenced diversification rates of *Ficus* in Neotropical rainforests?

Table 2
Estimated ages (Mya; median and 95% HPD) for crown nodes of the main lineages for selected nodes and their corresponding support values (BS, maximum-likelihood bootstrap; PP, Bayesian posterior probability).

Node	Cruaud et al. (2012) Median ages Ma (95% HPD)	This study Calibration with crown ages	This study Calibration with crown and stem ages	Support (BS/PP)
Crown *Ficus*	74.9(101.9–60.0)	63.4(80.8–47.8)	63.5(83.5–45.4)	100/1
Crown *Americanae*	20.5(29.3–13.1)	25.9(32.45–20.3)	26.7(33.2–20.5)	88/0.99
Stem *Americanae*	32.3(46.1–22.1)	31.36(40.5–23.6)	32.6(42.65–24.2)	–
Crown *Pharmacosycea*	16.2(25.7–8.2)	19.87(26.6–13.6)	20.1(27.0–13.9)	100/1
Stem *Pharmacosycea*	74.9(101.9–60.0)	63.37(80.8–47.8)	55.4(74.9–38.0)	–

2. Material and methods

2.1. Sampling

To assess the biogeographical history and diversification of *Ficus* in the Neotropics we assembled the largest and most representative data matrix to date including 66 species of section *Americanae* and 11 species of section *Pharmacosycea* (more than 65% of the Neotropical *Ficus*). Samples were collected in the field or from herbarium specimens and a list of material and vouchers is included in Appendix A. We produced 188 new sequences for *Ficus* expanding the sampling of Neotropical species available in GenBank for phylogenetic studies considerably from 31 species included in Cruaud et al. (2012) to 77 species (65%) included in the present study. Our sampling comprises species from all areas of the Americas including 28 of approximately 50 species found in Amazonia and 30 of 35 species of the Atlantic Rainforest (see Carauta, 1989; BFG, 2015; Berg and Villavicencio, 2004). *Antiaropsis decipiens, Castilla elastica* and *Sparratosyce dioica* representing Castilleae, the sister tribe of *Ficus*, were used as outgroups following previous studies (Rønsted et al., 2005; Zerega et al., 2005).

2.2. DNA extraction, amplification, and sequencing

Total genomic DNA was extracted using the CTAB protocol of Doyle and Doyle (1987) from 15 to 30 mg of silica dried leaf-fragments or herbarium material. Since plastid regions provide little phylogenetic information within *Ficus* (Rønsted et al., 2008a), phylogenetic studies of *Ficus* have focused on the more informative single- or low copy nuclear regions. For the present study, we sequenced five nuclear markers: the Internal Transcribed Spacer including the coding (ITS) region, the External Transcribed Spacer (ETS) region, the Glycerol-3-phosphate dehydrogenase gene (G3pdh), the Magnesium-protoporphyrin IX monomethyl ester cyclase (At103) and the Triosephosphate isomerase gene (Tpi).

We chose the ITS, ETS and the G3pdh regions because they have provided good resolution of *Ficus* in previous studies (Rønsted et al., 2008a). Additionally, At103 and Tpi were selected because they have provided both good amplification and resolution in other studies (Strand et al., 1997; Li et al., 2008).

Amplification of ITS, ETS and G3pdh followed Rønsted et al. (2008a). Amplification of At103 followed Li et al. (2008) except that 1 μL each of DMSO and BSA were added to all reactions. Amplification of Tpi was performed in 25 μL reactions adding 0.5 μL DNA to a reaction mixture of 1x Buffer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.5 mM of each primer, 1U of Taq polymerase (VWR international, Haasrode, Belgium), 1 μL of BSA and DMSO. The PCR protocol for Tpi consisted of 2 min at 95 °C, followed by 35 cycles of 1 min at 95 °C, 1.30 min at 46.5 °C, 2 min at 72 °C and a final extension for 9 min at 72 °C. PCR products were purified using a Qiagen PCR purification kit (Qiagen Inc., Valencia, California, USA) except for PCR products of Tpi which were purified using ExoSAP-IT® (Affymetrix UK Ltdn, High Wycombe, UK), following the manufacturers protocols.

PCR products were sequenced using Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Austin, Texas, USA) and purified sequencing products were run on an AB3130x1 Genetic Analyzer (Applied Biosystems/HITACHI, Tokyo, Japan) at the Laboratório de Sistemática Molecular de Plantas of the Universidade Estadual de Feira de Santana (LAMOL/UEFS) or at the National Sequencing Centre, Natural History Museum of Denmark. Forward and reverse sequences were edited and assembled in Geneious V. 7 to V. 8 (http://www.biomatters. com) or with Sequencer 4.8™ software (Gene Codes, Ann Arbor, MI, USA). Alignments were conducted using MUSCLE (Edgar, 2004) with default settings and were inspected in Mesquite (Maddison and Maddison, 2015).

2.3. Phylogenetic analyses

Despite some missing data all major clades are represented by all DNA regions sampled. Prior to analysis, the best fitting model of sequence evolution was determined using jModeltest v. 2.1.7. (Darriba et al., 2012) following the AIC criterion (Posada and Buckley, 2004). The HKY + G model of sequence evolution was selected for At103, ETS, G3pdh and Tpi; and GTR + G was selected for ITS.

Maximum-likelihood (ML) analysis was performed in RAxML v. 8 (Stamatakis, 2014) in the CIPRES Science Gateway v.3.3 (Miller et al., 2010). We executed 1000 rapid bootstrap inferences and, thereafter, a thorough ML search following Stamatakis (2014).

Bayesian analyses were conduced in MrBayes v.3.2.5 (Ronquist and Huelsenbeck, 2003) in the CIPRES Science Gateway v.3.3 (Miller et al., 2010). Two separate MCMC runs each initiated with a randon tree and eight simultaneous chains set at default temperatures (Ronquist and Huelsenbeck, 2003). Markov chains were run for 2×10^7 generations sampling every 1000th generation. Convergence of runs was tested by inspecting whether the standard deviation of split frequencies of the runs was < 0.01 and by using the effective sample sizes (ESS) > 200 in Tracer v.1.6 (Rambaut and Drummond, 2013). We then used MrBayes command "sumt" to summarize 75% of the trees sampled from post burn-in generations into a 50% majority rule consensus tree that included posterior probabilities (PP) as branch support estimates. The trees were annotated for presentation using FigTree v.1.4.2 (Rambaut, 2014).

2.4. Divergence time estimation

To obtain the chronograms we used the uncorrelated log-normal relaxed clock implemented in the BEAST package v.1.8.0 (Drummond et al., 2012) in the CIPRES Science Gateway v.3.3 (Miller et al., 2010) with a null tree prior (Yule speciation), and a random starting tree. A XML file was generated in BEAUti v.1.8.0 (Drummond et al., 2012). We conducted 2 runs of 60×10^6 generations, sampling trees every 6×10^3 generations. The output files were examined in Tracer v.1.6 (Rambaut and Drummond, 2013) to assess convergence of the runs and that the ESS values were > 200 for all parameters. The runs were combined using LogCombiner v.1.8.2 (Drummond et al., 2012). Following the removal of 10% burn-in, the sampled posterior trees were summarized using TreeAnnotator v.1.8.2 (Drummond et al., 2012) to generate a maximum clade credibility tree and calculate the mean ages, 95% highest posterior density intervals (95% HPD) and PP. The chronogram was visualized and annotated using FigTree v.1.4.2 (Rambaut, 2014).

For the molecular dating, calibration points were selected based on sampling of lineages and the ages derived from the most comprehensive analyses of *Ficus* (Cruaud et al., 2012). We adopted the dates from Cruaud et al. (2012) for the Most Recent Common Ancestor (MRCA) of the genus *Ficus*; the MRCA of section *Americanae*; section *Conosycea*; section *Galoglychia*; section *Malvanthera*; section *Pharmacosycea* and section *Sycomorus* using a normal prior for each node.

2.5. Biogeographical analyses

For biogeographical analyses we defined nine areas based on the current species distribution (Table 3) and the biogeographical regions of the neotropics defined by Morrone (2014). Species occurrence data were compiled from the literature including taxonomic revisions, floras and checklists (Berg and Villavicencio, 2004, Carauta, 1989; BFG, 2015) as well as from inspections of scientific collections of plants of Neotropical area mainly in Brazilian herbaria: ALCB, ASE, B, BHCB, BOTU, C, CEPEC, CESJ, CGMS, CVRD, EAC, ESA, FLOR, FUEL, FURB, GUA, HB, HRB, HUEFS, HUESB, HUFU, IAC, IAN, IAL, INPA, K, MBM, MBML, NX, NY, PACA, PAMG, R, RB, SP, SPF, SPFR, UEC, VIC and VIES. Herbarium acronyms follow Thiers (continuously updated).

Table 3
Geographical areas used in the Biogeographical analysis.

Code	Geographical areas
A	Antilles
B	Mesoamerica
C	Amazonia
D	South American transition zone
E	Brazilian seasonally dry tropical forest: Caatinga
F	Brazilian Atlantic forest
G	Cerrado, Chacoan and Pampean provinces (*Sensu* Morrone, 2014)
H	Afrotropics
I	Australia and Asia

Reconstructions were inferred using the MCC tree obtained in BEAST. The maximum number of areas was restricted to the maximum number of regions observed among extant taxa (five) and dispersion probabilities among areas were equally weighted (unconstrained model).

The analyses were conducted in the package BioGeoBEARS 0.2.1 (Matzke, 2013a) implemented in R 3.1.2 (R Core Team, 2016), which allows comparison of different models of ancestral range estimation (ARE). We used six different models, namely DEC, DEC + J, DIVA, DIVA + J, BayArea, BayArea + J (J models include a j parameter controlling founder event speciation). Massana et al. (2015) suggest the DEC model underestimates local extinction because the model allows observed species to transition into being present in no areas (i.e., null range). Consequently, we also accounted for the impact of the null range for each tested model to improve inference of local extinction (Massana et al., 2015, Matzke 2013b). These modified models are identified with an asterisk (*) in Table 4. Fit of the models was compared using AIC values in BioGeoBEARS.

2.5.1. Diversification through time

The temporal accumulation of lineages was assessed with a lineage-through-time plot (LTT) using the R package phytools 0.4–45 (Revell, 2012). We use the MCC tree and 1000 ultrametric trees randomly sampled from the posterior distribution of trees obtained in the BEAST analyses to obtain the 95% confidence interval (CI).

2.5.2. Diversification and evolution of life-history traits and propagule (syconia) size in Neotropical figs

Neotropical figs are composed of two distinct lineages, sections *Americanae* and *Pharmacosycea* (Rønsted et al., 2005; Jackson et al., 2008; Cruaud et al., 2012), which both have different life forms in the initial phase of development (see Table 1; Fig. 6).

To explore the importance of the initial life-history traits to the diversification of Neotropical *Ficus*, we reconstructed the evolution of habit (terrestrial vs hemiepiphyte). We also evaluate propagule traits and their association with diversification patterns. A matrix was constructed scoring syconia as small (0.5–1 cm diam.) or medium (1–2.4 cm diam.) yellow, orange or purple at maturity vs. large syconia (> 2.5 cm diam.) usually green at maturity. Data for life histories and syconia size were obtained from taxonomic literature (Berg and Villavicencio, 2004; Carauta, 1989) supplemented by measurements made by the authors.

We tested the hypothesis that small and medium syconia being yellow, red or purplish at maturity (state 1) have allowed for greater diversification compared to larger syconia green at maturity (state 2). To test whether life history or propagule characteristics are associated with differential rates of diversification, ancestral-state reconstruction on the MCC tree and the 1000 subsampled posterior trees was assessed in phytools 0.4.98 using stochastic character mapping (Huelsenbeck et al., 2003). We also implemented the Binary State Speciation and Extinction model (BiSSE) (Maddison et al., 2007) using the R package Diversitree 0.9–7 (FitzJohn, 2012). BiSSE estimates speciation and extinction rates among lineages with different states of a binary trait. We compared eight models: full; equal.L (lambda0 = lambda1); equal.m (mu0 = mu1); equal.q (q01 = q10); equal.lm; equal.lq; equal.mq; equal.lmq. Fit of alternative models were evaluated using LnL and AIC scores. The estimated speciation and extinction rates are plotted using a R Package plot3D (Soetaert, 2013). All analyses were run on the 1000 random trees as well as on the maximum credibility tree.

3. Results

Both Neotropical *Ficus* sections *Americanae* and *Pharmacosycea* were supported as monophyletic and with sect. *Pharmacosycea* sister to all other sampled sections (Fig. 1). Most speciation events yielding the current diversity of these sections occurred between 15.0 and 2.0 Mya and the greater diversification was observed from the middle Miocene (Fig. 1; Table 2).

Ancestral range estimation recovered the BayArea* + J (Landis et al., 2013) as the best-fit model for our data set (lnL = −369.2, AIC = 744.3) followed by the BayArea* (Table 4; Fig. 2). The use of parameter j (founder-event speciation) and the impact of the null range (Massana et al., 2015) both significantly improved the fit of all models tested (Table 4).

Both molecular dating analysis and ancestral range estimation indicate that Neotropical *Ficus* sections arrived in eastern Brazilian Atlantic Forest from Old World ancestors between the late Oligocene and earlier Miocene (Fig. 2; Table 2). *Pharmacosycea* likely diverged in the Paleogene 58.1 Mya (95% HPD 34.6–60.8) from an Asian ancestor (clade I in Fig. 2) and the MRCA of section *Pharmacosycea* was dated to the early Miocene (19.9 Mya; 95% HPD: 26.6–13.6 Mya). The divergence of section *Americanae* is dated to around 31.4 Mya (95% HPD 23.6–40.5) (Fig. 1). Atlantic Forest is the most probable ancestral area of this clade in the Oligocene (Fig. 2II) and the MRCA of *Americanae* was dated to the Oligocene (25.9 Mya; 95% HPD: 32.5–20.3 Mya) suggesting that both lineages of Neotropical *Ficus* arrived on the continent between the later Oligocene and the early Miocene (Fig. 2). We

Table 4
Likelihood (LnL) and Akaike information criterion (AIC) scores from each of the models tested in BioGeoBEARS. The best model is highlighted in bold.

	LnL	Numparams	d	e	J	AIC
DEC	−459.3	2	0.010	0.0046	0	922.6
DEC + J	−458.1	3	0.0098	0.0032	0.0060	922.2
DIVALIKE	−478.8	2	0.012	0.0088	0	961.5
DIVALIKE + J	−477.8	3	0.010	1.0e−12	0.0071	961.5
BAYAREALIKE	−405	2	0.0059	0.054	0	814
BAYAREALIKE + J	−399.7	3	0.0056	0.050	0.0022	805.4
DEC*	−436.1	2	0.019	0.10	0	876.2
DEC* + J	−436.1	3	0.020	0.11	1.0e−05	878.2
DIVALIKE*	−445.1	2	0.024	0.15	0	894.2
DIVALIKE* + J	−436.1	3	0.020	0.11	1.0e−05	878.2
BAYAREALIKE*	−372.8	2	0.0063	0.080	0	749.7
BAYAREALIKE* + J	**−369.2**	**3**	**0.0060**	**0.076**	**0.0017**	**744.3**

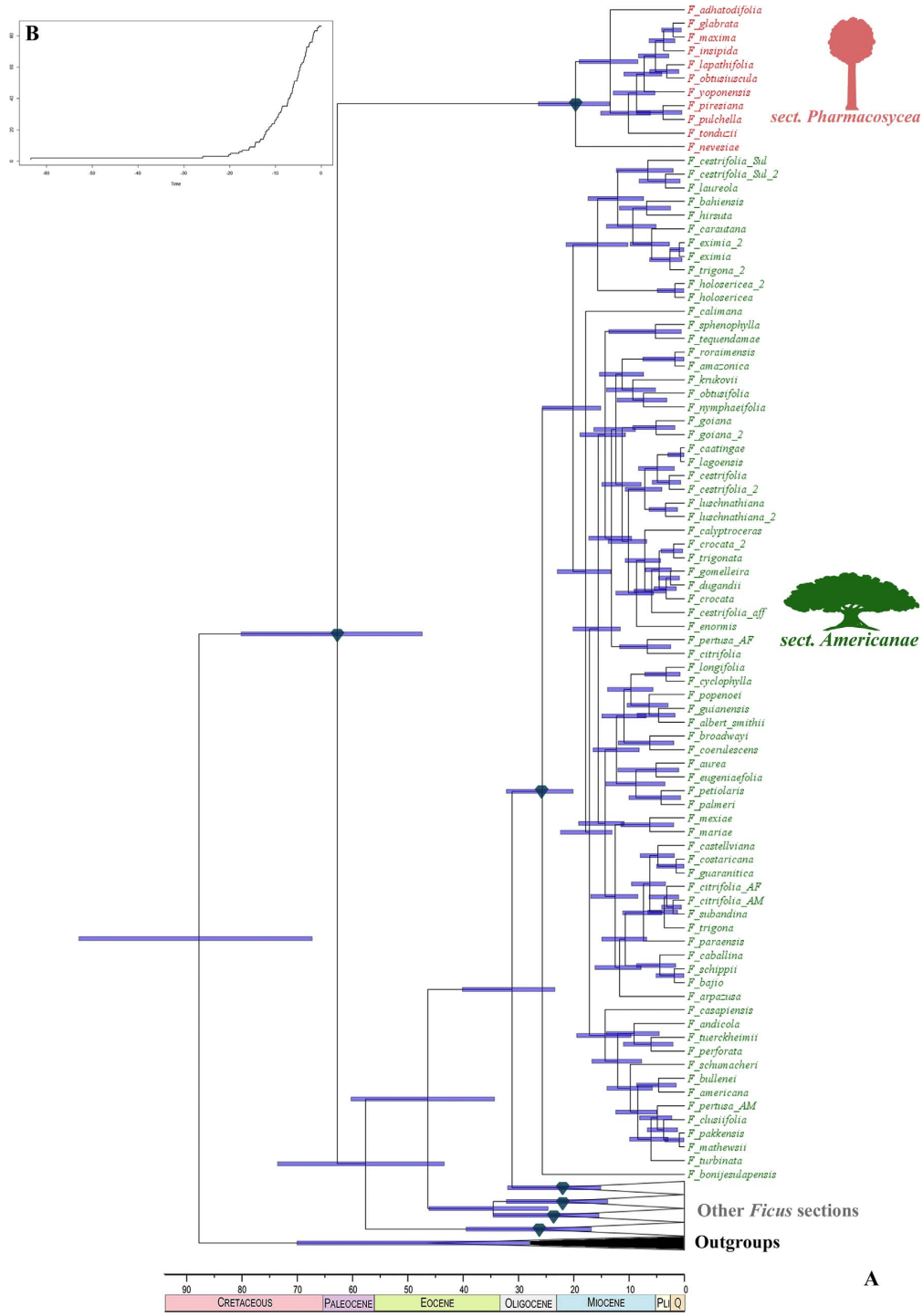


Fig. 1. A. MCC tree of Neotropical *Ficus* derived from divergence time estimation in Beast. Black diamonds refer to the calibration points. Section *Americanae* is highlighted in blue and section *Pharmacosycea* in red. Shaded horizontal bars show the 95% highest posterior densities of divergence times. B. Log-lineages-through time-plot (LTT) following the same time scale as the chronogram. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

156

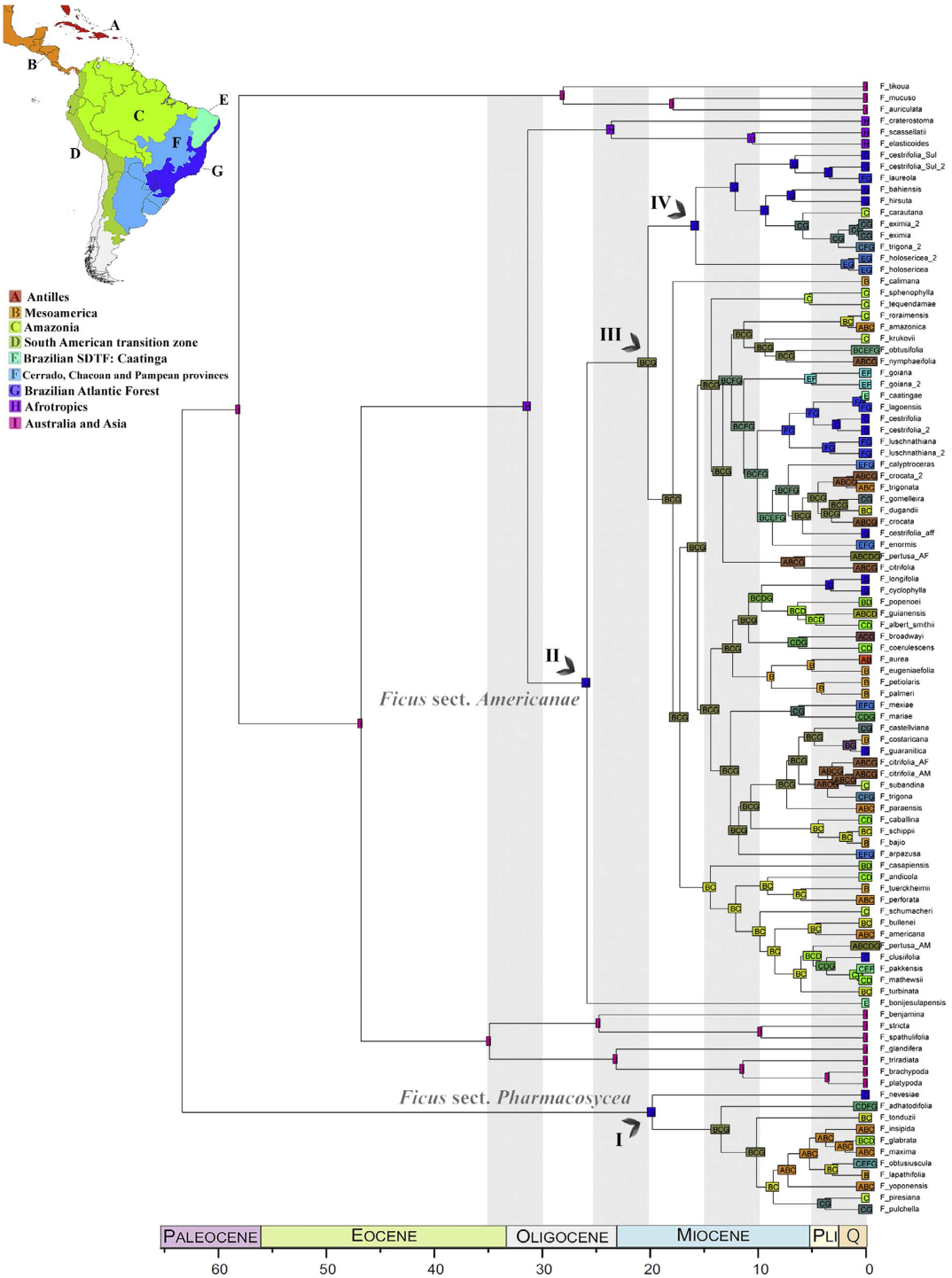


Fig. 2. Ancestral range estimations for the Neotropical *Ficus* using the BAYAREA*+j model in BioGeoBEARS -LnL = 369.2. States at nodes (squares) represent the most probable ancestral area before the speciation event. Squares with more than one letter refer to ancestral areas composed of more than one biogeographical area. Stratigraphical time in millions of years ago (Mya) is indicated on the time-scale.

detected a first diversification event resulting in *F. bonijesulapensis* a species endemic to Seasonally Tropical Dry Forests (STDF) of Brazil. A second lineage consisting of all other *Americanae* species diverged in the early Miocene with inferred origin in both rainforests (Atlantic Forest, Amazonia, Central America).

At the end of the early Miocene there was a split resulting in the origin of a group with Atlantic rainforest species (Fig. 2IV). In the middle Miocene we detected a divergence between lineages of *Ficus* in the Atlantic Forest and the Amazon or Central America (Fig. 2III). In the late Miocene (11.6–5.3 Mya) our analysis (Fig. 2) identified colonization events from Amazonia to the Antilles, Central American and Andes.

Neotropical figs expanded to other American rain forests by the early to middle Miocene, evidenced by the reconstruction of a composed BCG area for the MRCA (Fig. 2) of the most speciose clades in both sections. *Ficus* also had an increase in diversification rate between 16 and 5 Mya (Fig. 1 B). From the middle Miocene, diversification rate increases in both sections and remained high until the Pleistocene (2.6–0.0 Mya), which was marked by a slightly smaller *Ficus* diversification in rainforests (Fig. 2B). However, there was diversification to adjacent biomes e.g. *F. lagoensis* (Cerrado and Atlantic Forest) and *F. caatingae* (endemic to Caatinga) (Fig. 2).

The current endemic species of the Atlantic Forest and Amazonia started diversification in the late Miocene (Fig. 2). We also detected dispersal events between Amazonia and Atlantic Forest and diversification in a group of endemic species to Central and North America at this time. At the end of the Miocene and Pliocene, dispersal and diversification of *Ficus* was inferred from rain forests to seasonally dry forests, and savannas on the continent. In the Pliocene (5.3–2.6 Mya) we detected the first occurrence of species in the Antillean and the Andes region (Fig. 2). The analysis detects an expansion of Amazonian and Atlantic species to the Cerrado and Caatinga and intensification of floristic exchanges between the Amazon and the Atlantic Forest in this period.

We detected two transitions between the different initial habits (terrestrial vs. hemiepiphyte) in Neotropical *Ficus* (Fig. 3A and B). Only one species in section *Pharmacosycea* (*F. crassivenosa*, not sampled here), initially exhibits the hemiepiphytic habit. DAIC scores from BiSSE analyses recovered equal.l (lambda1 = lambda0) as the best-fit model (Fig. 5B and C; Table 5). We did not find evidence for different rates of speciation associated with the hemiepiphytic habit. However, posterior density of the parameters q (transitions) and mu (extinction) were differentiated (Fig. 5B and C) suggesting lower extinction associated with the hemiepiphytic habit (see Fig. 6).

Eleven transitions among syconia size were inferred in Neotropical *Ficus*. The best-fit model was equal.lq (lambda1 = lambda0; q01 = q10) (Fig. 5E–G; Table 6) meaning that the speciation and transitions between traits are equal. However, we found lower extinction rates associated with small or medium and coloured syconia (mu0) compared to large green syconia (mu1).

4. Discussion

4.1. Phylogeny, divergence times and diversification of Neotropical figs

In agreement with previous molecular studies including Neotropical samples (Rønsted et al., 2005, Rønsted et al., 2007, Rønsted et al., 2008a; Cruaud et al., 2012) we found a strong statistical support for monophyly of sections *Americanae* and *Pharmacosycea*. However the internal relationships among the species, mainly in section *Americanae*, are not strongly supported. This lack of phylogenetic resolution with the short branch lengths is probably a consequence of recent diversification observed in Neotropical *Ficus*. This pattern is also found in other species-rich lineages in Neotropical rainforests [e.g. *Inga* (Richardson et al., 2001); Gesneriaceae (Perret et al., 2013); *Astrocaryum* (Arecaceae, Roncal et al., 2013); *Attalea* (Arecaceae, Freitas et al., 2016); *Philodendron* (Araceae, Loss-Oliveira et al., 2016)] and

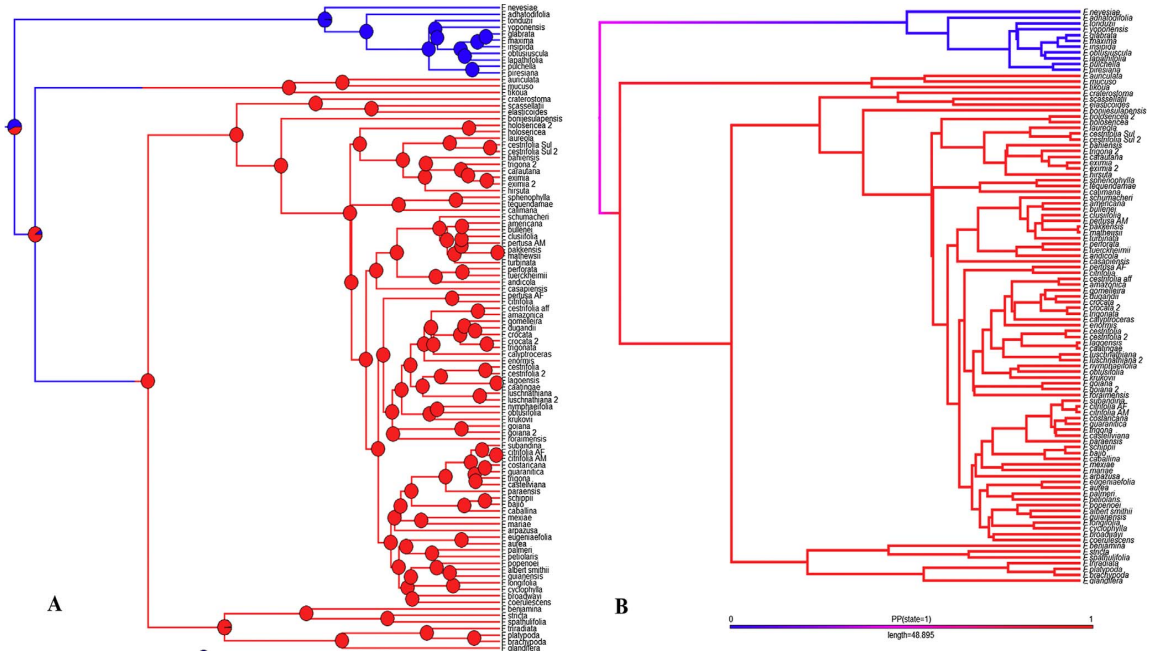


Fig. 3. A. Result from 10,000 stochastic character-mapping reconstructions of the life forms (terrestrial vs. hemiepiphyte) on the MCCT of the Neotropical *Ficus* using Phytools. Red indicates hemiepiphytic habit. B. Result from 10,000 stochastic character-mapping reconstructions of the life forms (terrestrial vs. hemiepiphyte) from 1000 subsampled posterior trees using Phytools. The colour of branches in the tree indicates the posterior probability along the branches. Red indicates high posterior probability of hemiepiphytic habit. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

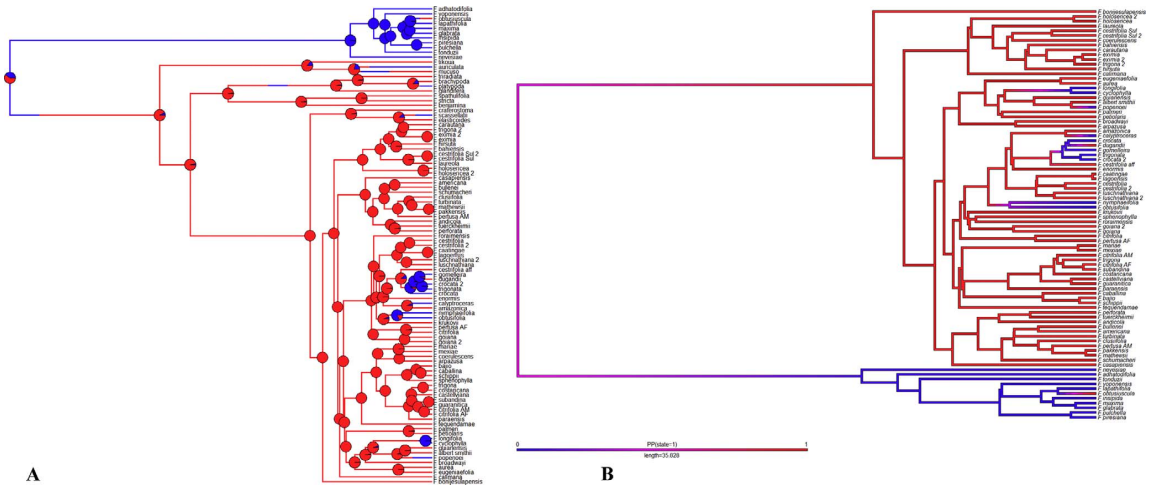


Fig. 4. A. Result from 10,000 stochastic character-mapping reconstructions of the syconia size (small vs. medium and large) on the MCCT of the Neotropical *Ficus* using Phytools. Red indicates small size. B. Result from 10,000 stochastic character-mapping reconstructions of the syconia size (small vs. medium and large) from 1000 subsampled posterior trees using Phytools. The colour of branches in the tree gives the posterior probability of each size along the branches. Red indicates high posterior probability of small syconia. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

have been interpreted as a consequence of incomplete lineage sorting due to their recent diversification since the late Miocene (Hughes et al., 2013), as we found here in *Ficus*.

Most of the divergence in Neotropical *Ficus* took place in rainforests and happened in the last 16 Mya. In fact, high levels of rainfall, temperature, and habitat heterogeneity have been correlated with high species richness in general (Kreft and Jetz, 2007) and were the prevalent conditions by the middle Miocene climatic optimum, allowing the expansion of rainforests in America (Morley, 2000; Zachos et al., 2001). All idiosyncratic geographical and climatic events in the Neotropical region probably variously influenced the diversification of *Ficus* in different domains of the region.

4.2. Diversification of Ficus in Amazonia

Amazonian species of *Ficus* appeared scattered across the phylogeny (Fig. 3). The lack of geographic structure was observed in other Amazonian groups (Hughes et al., 2013) such as *Clusia* (Gustafsson and Bittrich, 2002) , *Guatteria* (Erkens et al., 2007), *Inga* (Richardson et al., 2001) and *Swartzia* (Torke and Schaal 2008). This pattern suggests that vicariance events (e.g. the Pebas System) were not the major driver of speciation in this region (Hughes et al., 2013) and that these communities were assembled mostly by dispersal. The importance of immigrants in the composition of the Amazonian flora has already been suggested (Pennington et al., 2004; Pennington and Dick, 2004).

We found species endemic to the Amazon diversifying only in the last 10 Mya (late Miocene), which coincides with the end of the Pebas system and the establishment of the current course of the Amazon River (Hoorn et al., 2010). Biotic interchange between South America, the Antilles and Central America occurred mainly at 23–20 and 8–6 Mya (Bacon et al., 2015). Events of colonization of adjacent areas from Amazonian taxa were inferred at this time. Diversification in Mesoamerica is detected for a clade of endemic species diverging from 10 Mya. *Ficus* reached the Antillean region around 8–5 Mya, coinciding with the closure of the Isthmus of Panama connecting these areas.

Species shared between the Amazon and the South American Transition Zone (*sensu* Morrone, 2014) are rare and probably colonized the latter area around 10 Mya coinciding with the end of the Pebas System (Hoorn et al., 2010) as a potential barrier to colonization of these areas. Other colonization events between these areas were detected in the Pliocene, which were probably due to the final phase of the Andes uplift creating new niches. Despite the biotic interchange between Amazonia, the Antilles and the South American Transition zone we did not find an increase in diversification rates in these areas as reported for Arecaceae (Roncal et al., 2013; Freitas et al., 2016).

4.3. Diversification of Ficus in the Atlantic Forest

Neotropical *Ficus* of both sections diversified first in Atlantic Forest and expanded to other areas during the Oligocene and Miocene in accordance with favorable climatic conditions and the continuity of Neotropical wet forests during this time (Morley, 2000; Zachos et al., 2001). Thenceforth *Ficus* in the Atlantic Forest experienced two bursts of diversification in the middle Miocene (16.0–11.6 Mya) and in the Pliocene (5–3 Mya), respectively. We compared the patterns observed in *Ficus* with the different available hypotheses of diversification in this area reflecting the complexity and different histories of Atlantic Forest areas.

Historically situated in an area with less drastic tectonic and hydrological changes the Atlantic Forest was considered more stable than Amazonia (Pennington et al., 2004; Carnaval and Moritz, 2008; Hoorn et al., 2010; Hughes et al., 2013; Pennington et al., 2015). The first diversification burst of neotropical *Ficus* in the middle Miocene (16.0–11.6 Mya) is chronologically coherent with the Miocene climatic optimum (Morley, 2000; Zachos et al., 2001) and continuous rainforests in the Americas. Examples of wet forest Neotropical taxa originating in the Miocene are well documented: *Inga* (Richardson et al., 2001) Gesneriaceae (Perret et al., 2013), *Astrocaryum* (Arecaceae, Roncal et al., 2013); *Attalea* (Arecaceae, Freitas et al., 2016); *Philodendron* (Araceae, Loss-Oliveira et al., 2016) and the Miocene favorable climatic conditions also appears to have driven the diversification of *Myrcia* (Lucas et al., 2011) in AF and Amazonia.

The Pliocene was characterized by a gradual decrease in temperature and humidity resulting in contraction of wet forests and expansion of dry forests (Zachos et al., 2001). Some AF areas served as *refugia* during this time (Carnaval and Moritz, 2008; Thomé et al., 2010). As an effect of these fluctuations, *Ficus* experienced a second burst of diversification in the Pliocene (5–3 Mya) resulting in the majority of species endemic to the Atlantic Forest and the southern part of the domain. Besides identifying an increase in diversification, our analyses

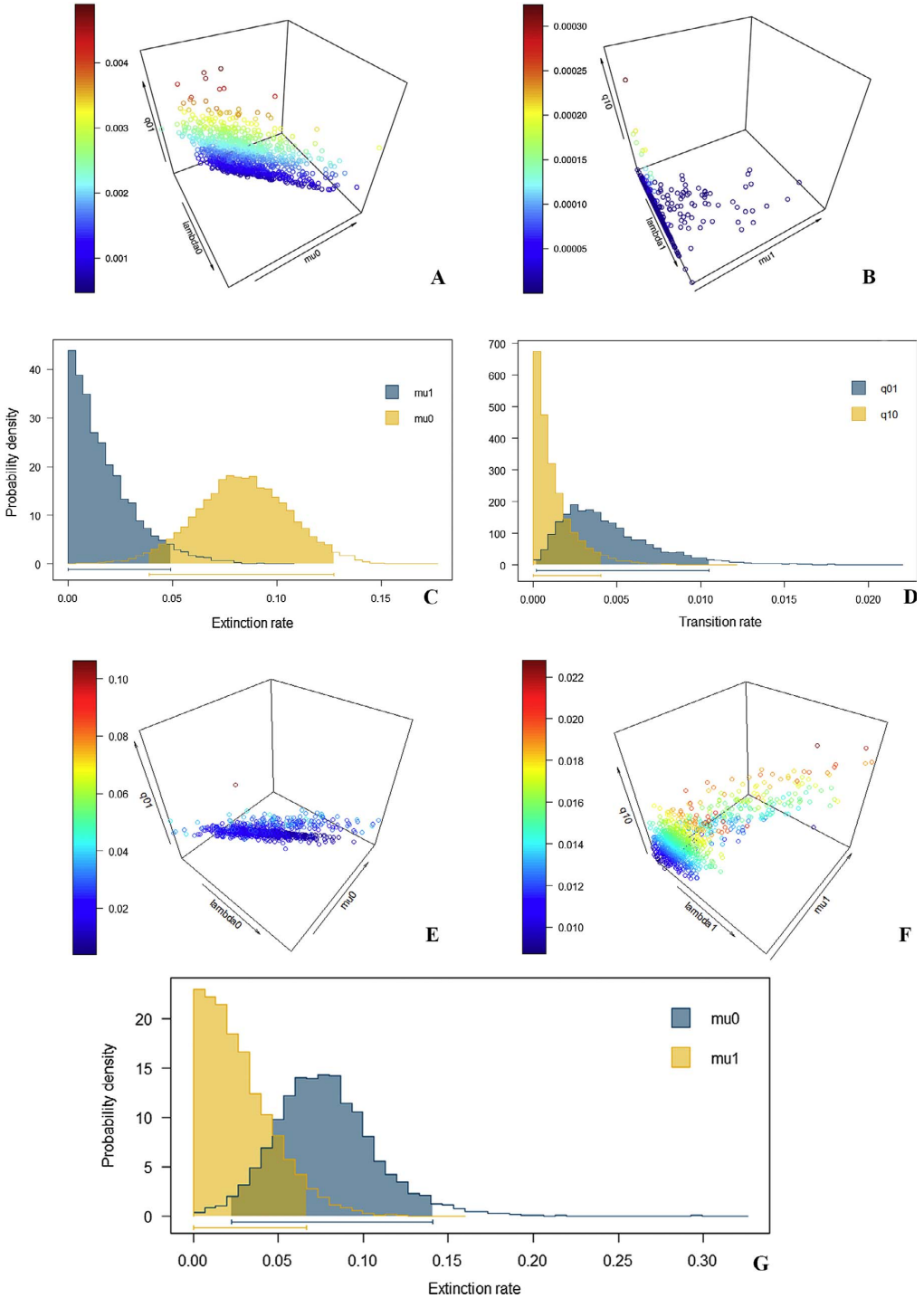


Fig. 5. Result of Binary State Speciation and Extinction (BiSSE analysis). A–D: Habit (0-Terrestrial, 1-Hemiepiphyte). A. Distributions of DAIC values to q01, *lambda*0 and *mu*0; B. Distributions of DAIC to q01, *lambda*0 and *mu*0; C. Posterior probability distributions of extinction (*mu*0 and *mu*1); D. Posterior probability distributions transition rates (q01, q10). E-G: Syconia traits (0-Large, green; 1-Medium to small, not green). E. Distributions of DAIC values to q01, *lambda*0 and *mu*0; F. Distributions of DAIC to q01, *lambda*0 and *mu*0; G. Posterior probability distributions of extinction (*mu*0 and *mu*1). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

160

Table 5
Likelihood ratio test and AIC of each tested model in BiSSE analyses of correlated diversification. Initial habit (terrestrial vs hemiepiphytic) of the Neotropical *Ficus*. The best model is highlighted in bold.

	Df	lnLik	AIC	ChiSq	Pr(> |Chi|)
full	6	−358.78	729.56		
equal.l	**5**	**−358.79**	**727.57**	**0.0113**	**0.91**
equal.m	5	−360.55	731.10	3.5383	0.059
equal.q	5	−360.13	730.26	2.698	0.10
equal.lm	4	−365.00	738.00	12.44	0.001
equal.lq	4	−360.20	728.40	2.835	0.24
equal.mq	4	−363.07	734.13	8.573	0.013
equal.lmq	3	−366.34	738.68	15.11	0.001

also detected evidence of colonization of the newly opened formations followed by *in situ* diversification. Dispersal events from Atlantic Forest to adjacent areas and Amazonia are common (DaSilva and Pinto-da-Rocha, 2011) and were reported in *Attalea* (Freitas et al., 2016) and other genera occurring in dry forests (Pennington et al., 2004).

The distinct histories between the southern and northern part of AF generated different patterns in species distribution in the AF. The existence of a more stable area in the northern part of the AF, the Bahian refugium (Carnaval and Moritz, 2008), is supported by recent studies (e.g. Martins, 2011; Staggemeier et al., 2015). The occurrence of endemic species of *Ficus* is coherent with *in situ* speciation into refuges (Staggemeier et al., 2015). The shared species between Amazonia and the northern part of AF could have resulted from past forest connections between Amazonia and AF in climatically favorable times at the Neogene and Quaternary (Santos et al., 2007).

In contrast, the southern part of the AF was more unstable due to neotectonic events in the Pliocene. These were followed by a decrease in rainfall and a fragmentation of the southern parts (Ricommini and Assumpção, 1999), a scenario, which would have prevailed until the end of the Pliocene (Grazziotin et al., 2006). A recent hypothesis suggests that the southern part of the AF would have expanded (not contracted) during the last glacial maximum (21 kya) (Leite et al., 2016). In spite of the unfavorable climatic conditions in the southern part of the AF there were probably more terrestrial habitats in areas at lower elevation and an expansion of the forest would have promoted immigration between areas of the Atlantic Forest. This hypothesis is partially corroborated by a palynological study (Freitas et al., 2013), which supports the occurrence of a tropical forest in a region of the Brazilian continental shelf during this epoch.

4.4. Evolution of traits: the importance of being hemiepiphyte

Hemiepiphytism in *Ficus* is thought to be an adaptation to avoid deep shade in the forest understory (Ramírez, 1977; Harrison, 2005). This trait enables greater initial access to light in the canopy and the hemiepiphytic woody species are adapted to adverse water conditions, which possibly could have been an advantage in colonizing unstable areas and dry regions. We hypothesize that the hemiepiphytic habit provides the species of section *Americanae* with greater ability for establishment under adverse lighting conditions and water stress (Hao et al., 2012).

According to our results hemiepiphytism reduces the extinction rates (Fig. 4A–D) if compared with primarily terrestrial lineages. The hemiepihiphytic habit can reduce risks related to terrestrial growth such as high competition, flooding, and terrestrial herbivores, which are common in tropical rainforests (Hao et al., 2012). This could have been particularly important for the diversification of *Ficus* in highly unstable and newer habitats such as the Amazonia.

In addition, all Neotropical *Ficus* species occurring in drier regions today are hemiepiphytes or hemiepilithics/lithophytes ("rock splitters"), as has also been found in a lineage of Australian and Australasian *Ficus* (Rønsted et al., 2008b). Hao et al. (2012) suggest that an important characteristic related to drought resistance is the loss and regeneration of the canopy. In the Neotropics several of the section *Americanae* are deciduous (e.g., *F. bahiensis*, *F. bonijesulapensis*, *F. caatingae*, *F. enormis*, *F. eximia*, *F. gomelleira*, *F. hirsuta*, *F. holosericea* and *F. mexiae*; AFPM, pers. obs.; Pereira et al., 2007; Bianchini et al., 2015). However, it is not known if all Neotropical hemiepiphytic species have this characteristic.

4.5. Evolution of traits: the importance of having small (and more) propagules for dispersal ability

Ficus shows traits typical of pioneer species such as small seeds, high fecundity, flexible rooting habits and high growth rates (Harrison, 2005). However, the obligatory pollination mutualism between Agaonidae wasps and *Ficus* is a limiting factor for reproduction (Janzen, 1979; Ramírez, 1970; Wiebes, 1979).

Our results indicate that occurrence of small and medium propagules reduce extinction rates (Fig. 5E–G). This may be related with the fact that species with small syconia having red shades or yellow ones are likely dispersed by birds and syconia with intermediate characteristics are probably dispersed both by bats and birds (Lomáscolo et al., 2008, 2010). Species with small syconia have more infructescenses along the branches and higher reproductive success compared to species with large, bat-dispersed syconia (Lomáscolo et al., 2008, 2010). While small syconia can be swallowed whole and dispersed over long distances, larger syconia are dispersed by larger animals over a shorter distance. It also seems to be an advantage to have the seeds in many infructescences instead of having seeds concentrated in few infructescences.

In addition, larger syconia generally have far fewer inflorescences per individual compared to species with smaller syconia. Large syconia are also more susceptible to attacks by non-pollinating wasps. Additionally, large syconia tend to be dispersed by bats, which are less effective dispersers compared to birds and bats also eat immature syconia as well (Jacomassa and Pizo, 2010). Finally, many seeds loose viability when passing through the digestive tract of bats, which happens at a much lower rate when passing through birds (Traveset, 1998).

5. Conclusions and perspectives

The first biogeographic approach focused in Neotropical *Ficus* is presented here. This study provides a phylogenetic background, which allows for addressing questions about the history of diversification and the importance of traits of *Ficus* in this process.

Despite the inclusion of both more species and more DNA regions, the internal relationships between the species of section *Americanae* are still not completely resolved. Future studies should preferably include several terminals per taxa and use a high-throughput NGS approach to fully resolve the phylogeny and clarify the relationship of most of its species. Despite uncertainty in some of the species relationships, the present study showed that the two lineages of *Ficus* with different traits arrived in the Americas at the Atlantic Forest, began to diversify in the Miocene, and expanded to other regions of the continent.

The Amazonian *Ficus* species diversity is a result of immigrations to this region occurring during unstable periods. Our results are consistent with well-documented events that influenced the history of these forests (the Andean uplifts, the end of Pebas system, and the closure of the Panama Isthmus). Dispersal events from the Amazon to the Andes, Antilles and Central America after the end of the Pebas system supports the hypothesis of colonization of new habitats in the Andes after the end of this barrier. The diversification of *Ficus* in Atlantic Forest was significantly different, being marked by endemisms related to its isolation after the separation of two blocks of Neotropical wet forests. Our results point to a complex diversification in the Atlantic Forest during two periods in the middle Miocene and the Pliocene as a result of

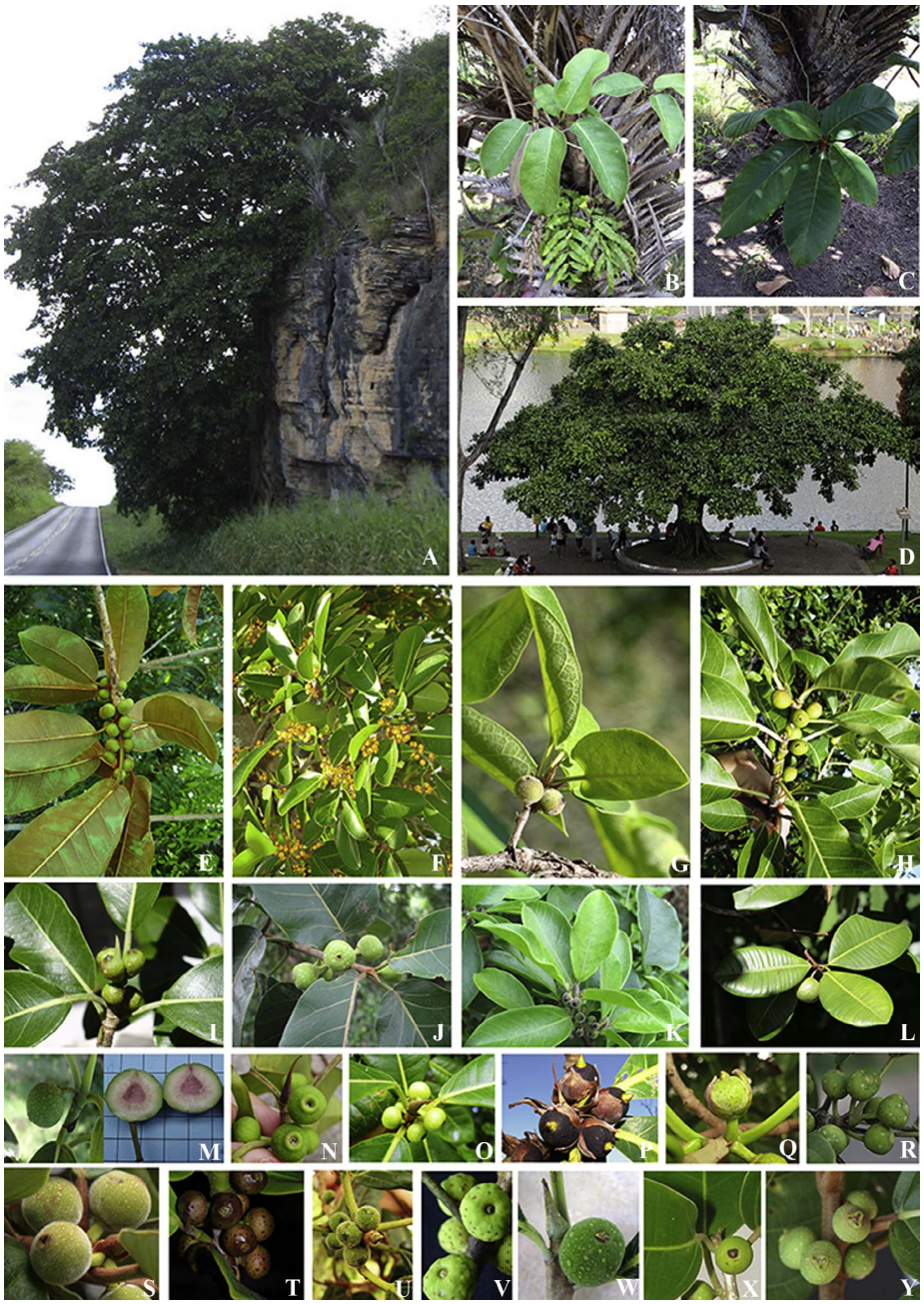


Fig. 6. Habit, branches and syconia diversity in Neotropical figs. A–D: Habit diversity. A: *F. bonijesulapensis* growing on stones, B. *F. arpazusa* hemiepiphytic, C. *F. gomelleira* hemiepiphytic; D. *F. cyclophylla* in an urban area. E–L: Branches. E: *F. castellviana*, F. *F. clusiifolia*, G. *F. holosericea*, H. *F. luschnathiana*, I. *F. cestrifolia*, J. *F. crocata*, K. *F. hirsuta*, L. *F. pulchella*, M–Y: Syconia diversity, M. *F. adhatodifolia* (Section *Pharmacosycea*), N. *F. arpazusa*, O. *F. bahiensis*, P. *F. cyclophylla*, Q. *F. eximia*, R. *F. guaranitica*, S. *F. gomelleira*, T. *F. hirsuta*, U. *F. lagoensis*, V. *F. mariae*, W. *F. mexiae*, X. *F. pertusa*, Y. *F. trigona*. Photos A, D, T by E. Matos; B–C, E–I, N–Q, U–X by A. Machado; J–K. by R. Lacerda; L, R, S, Y by R. Pereira and M, V. by G. Siqueira.

neotectonic events in the southern and southeastern distribution and a more stable area in the northern distribution of this forest.

Finally, our BiSSE analysis of specific traits in Neotropical *Ficus* suggests that some traits may also have influenced the diversification and current distribution of the species. The hemiepiphytic habit of section *Americanae* and the adaptation to adverse water conditions (i.e., apparently common feature of losing and regenerating the canopy during the dry season) allowed further expansion of this lineage in the Neotropical biomes. Additionally, small and medium sized propagules possibly provided a long-distance dispersal advantage in Neotropical forests allowing greater diversification success.

The successful diversification of Neotropical *Ficus* was likely the

Table 6
Likelihood ratio test and AIC of each tested model in BiSSE analyses of correlated diversification. Syconia size (Large vs. Small) of the Neotropical *Ficus*. The best model is highlighted in bold.

	Df	lnLik	AIC	ChiSq	Pr(> |Chi|)
full	6	−395.75	803.51		
equal.l	5	−396.25	802.49	0.9876	0.3203
equal.m	5	−396.18	802.35	0.8478	0.3572
equal.q	5	−395.76	801.53	0.0195	0.8890
equal.lm	4	−397.33	802.66	3.1548	0.2065
equal.lq	**4**	**−396.38**	**800.76**	**1.2493**	**0.5354**
equal.mq	4	−396.42	800.84	1.3335	0.5134
equal.lmq	3	−398.65	803.30	5.7917	0.1222

result of a complex and different history in each Neotropical area. This study provides new insights into the biogeographical history of *Ficus* in the Neotropical region and to the broader understanding of diversification of large genera. Future studies in *Ficus* or other megadiverse genera should include phylogeography and species distribution modelling approaches to better understand diversification in tropical forest communities.

Acknowledgements

We thank Alessandra Schnadelbach, Ariane Barbosa, Cassio Van den Berg and Lizandro Peraza for the help with analysis; Bruno Ferraz for help with the Graphical abstract; and Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) for issuing collecting permits. Fieldwork and DNA sequencing were partly sponsored by the PRONEX (FAPESB PNX0014/2009), PROTAX (CNPq 562354/2010-7). This study is part of the first author's PhD thesis developed in the Programa de Pós-Graduação em Botânica (UEFS) and supported by PhD fellowship grants from FAPESB (No. BOL0563/2014) and CAPES. The work was also supported by an International Network Grant from the Danish Ministry of Science, Innovation and Education, 2012 to NR, AFPM, RASP and LPQ. RASP was funded by CNPq (grant # 306078/2014-7) and FAPESP (grant #2004/10299-4).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2018.01.015.

References

Antonelli, A., Sanmartín, I., 2011. Why are there so many plant species in the Neotropics? Taxon 60 (2), 403–414.
Bacon, C.D., Silvestro, D., Jaramillo, C., Smith, B.T., Chakrabarty, P., Antonelli, A., 2015. Biological evidence supports an early and complex emergence of the Isthmus of Panama. Proc. Natl. Acad. Sci. U.S.A. 112, 6110–6115.
Berg, C.C., 1989. Classification and distribution of *Ficus*. Experientia 45, 605–611.
Berg, C.C., Corner, E.J.H., 2005. Moraceae – Ficus, in: Nooteboom, H.P. (Ed.). Flora Malesiana Series I – Seed Plants Vol. 17/Part 2. National Herbarium of the Netherlands, Leiden, pp. 1–152.
Berg, C.C., Simonis, E., 1981. The *Ficus* flora of Venezuela: five species complexes discussed and two new species described. Ernstia 6, 1–12.
Berg, C.C., Villavicencio, X., 2004. Taxonomic studies on *Ficus* (Moraceae) in the West Indies, extra-Amazonian Brazil, and Bolivia. Ilicifolia 5, 1–132.
Bianchini, E., Emmerick, J., Messetti, A., Pimenta, J., 2015. Phenology of two Ficus species in seasonal semi-deciduous forest in Southern Brazil. Braz. J. Biol. 75, 206–214.
BFG –The Brazilian Flora Group, 2015. Growing knowledge: an overview of seed plant diversity in Brazil. Rodriguésia 66, 1085–1113.
Bruun-Lund, S., Clement, W.L., Kjellberg, F., Rønsted, N., 2017. First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). Mol. Phylogen. Evol. 109, 93–104.
Bruun-Lund, S., Verstraete, B., Kjellberg, F., Rønsted, N. Rush hour in the Museum – Diversification patterns provide new clues for the success of Ficus L. (Moraceae). Acta Oecol., in press. doi:http://dx.doi.org/10.1016/j.actao.2017.11.001
Carauta, J.P.P., 1989. *Ficus* (Moraceae) no Brasil: Conservação e Taxonomia. Albertoa 2, 1–365.
Carnaval, A.C., Moritz, C., 2008. Historical climate modelling predicts patterns of current biodiversity in the Brazilian Atlantic forest. J. Biogeograp. 35, 1187–1201.
Costa, P.C., Lorenz-Lemke, A.P., Furini, P.R., Honorio-Coronado, E.N., Kjellberg, F., Pereira, R.A.S., 2017. The phylogeography of two disjunct Neotropical *Ficus* (Moraceae) species reveals contrasted histories between the Amazon and the Atlantic Forests. Bot. J. Linn. Soc. 185, 272–289.
Couvreur, T.L.P., Pirie, M.D., Chatrou, L.W., Saunders, R.M.K., Su, Y.C.F., Richardson, J.E., Erkens, R.H.J., 2011. Early evolutionary history of the flowering plant family Annonaceae: steady diversification and boreotropical geodispersal. J. Biogeogr. 38, 664–680.
Cruaud, A., Rønsted, N., Chantarasuwan, B., Chou, L.S., Clement, W.L., Couloux, A., Cousins, B., Genson, G., Harrison, R.D., Hanson, P.E., Hossaert-McKey, M., Jabbour-Zahab, R., Jousselin, E., Kerdelhué, C., Kjellberg, F., Lopez-Vaamonde, C., Peebles, J., Peng, Y.-Q.Q., Pereira, R.A., Schramm, T., Ubaidillah, R., van Noort, S., Weiblen, G.D., Yang, D.-R.R., Yodpinyanee, A., Libeskind-Hadas, R., Cook, J.M., Rasplus, J.-Y.Y., Savolainen, V., 2012. An extreme case of plant-insect codiversification: figs and fig-pollinating wasps. Syst. Biol. 61, 1029–1047.
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9 772-772.
DaSilva, M.B., Pinto-da-Rocha, R., 2011. História biogeográfica da Mata Atlântica: opiliões (Arachnida) como modelo para sua inferência, in: Carvalho, C.J.B., Almeida, E.A.B. (Orgs.). Biogeografia da América do Sul - Padrões e Processos Roca, São Paulo, pp. 221–238.
Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation method for small quantities of fresh tissues. Phytochem. Bull. 19, 11–15.
Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. 29, 1969–1973.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl. Acids Res. 32, 1792–1797.
Erkens, R.H.J., Chatrou, L.W., Maas, J.W., Van-der-Niet, T., Savolainen, V., 2007. A rapid diversification of rainforest trees (*Guatteria*: Annonaceae) following dispersal from Central into South America. Mol. Phylogenet. Evol. 44, 399–411.
Freitas, A.G., Carvalho, M.A.de, Mendonça, C.B.F., Gonçalves-Esteves, V., 2013. Pollen grains in quaternary sediments from the Campos Basin, state of Rio de Janeiro, Brazil: Core BU-91-GL-05. Acta Bot. Bras. 27, 761–772.
Freitas, C., Meerow, A., Pintaud, J., Henderson, A., Noblick, L., Costa, F., Barbosa, C., Barrington, D., 2016. Phylogenetic analysis of *Attalea* (Arecaceae): insights into the historical biogeography of a recently diversified Neotropical plant group. Bot. J. Linn. Soc. 2, 1–16.
FitzJohn, R.G., 2012. Diversitree: comparative phylogenetic analyses of diversification, in R. Methods Ecol. Evol. 3, 1084–1092.
Grazziotin, F.G., Monzel, M., Echeverrigaray, S., Bonatto, S.L., 2006. Phylogeography of the Bothrops jararaca complex (Serpentes: Viperidae): past fragmentation and island colonization in the Brazilian Atlantic forest. Mol. Ecol. 15, 3969–3982.
Gustafsson, M., Bittrich, V., 2002. Evolution of morphological diversity and resin secretion in flowers of Clusia L. (Clusiaceae): insights from ITS sequence variation. Nord. J. Bot. 22, 183–203.
Haffer, J., 1969. Speciation in Amazonian forest birds. Science 165, 131–137.
Hao, G.Y., Wang, A.Y., Sack, L., Goldstein, G., Cao, K.F., 2012. Is hemiepiphytism an adaptation to high irradiance? Testing seedling responses to light levels and drought in hemiepiphytic and non-hemiepiphytic *Ficus*. Physiol. Plantarum. 148, 74–86.
Harrison, R.D., 2005. Figs and the diversity of tropical rainforests. Bioscience 55, 1053–1064.
Honorio Coronado, E.N., Dexter, N.G., Poelchau, M., Hollingsworth, P.M., Phillips, O.L., Pennington, T.R., 2014. *Ficus insipida* subsp. *insipida* (Moraceae) reveals the role of ecology in the phylogeography of widespread Neotropical rainforest tree species. J. Biogeogr. 41, 1697–1709.
Hoorn, C., Wesselingh, F.P., ter Steege, H., Bermudez, M.A., Mora, A., Sevink, J., Sanmartin, I., Sanchez-Meseguer, A., Anderson, C.L., Figueiredo, J.P., Jaramillo, C., Riff, D., Negri, F.R., Hooghiemstra, H., Lundberg, J., Stadler, T., Sarkinen, T., Antonelli, A., 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science 330, 927–931.
Huelsenbeck, J.P., Nielsen, R., Bollback, J.P., 2003. Stochastic mapping of morphological characters. Syst. Biol. 52, 131–158.
Hughes, C.E., Pennington, R.T., Antonelli, A., 2013. Neotropical plant evolution: assembling the big picture. Bot. J. Linn. Soc. 171, 1–18.
Jackson, A.P., Machado, C.A., Robbins, N., Herre, E.A., 2008. Multi-locus phylogenetic analysis of neotropical figs does not support co-speciation with the pollinators: the importance of systematic scale in fig/wasp cophylogenetic studies. Symbiosis 45, 57–72.
Jacomassa, F.A.F., Pizo, M.A., 2010. Birds and bats diverge in the qualitative and quantitative components of seed dispersal of a pioneer tree. Acta Oecol. 36, 493–496.
Janzen, D.H., 1979. How to be a fig. Annu. Rev. Ecol. Syst. 10, 13–51.
Jousselin, E., Rasplus, J.Y., Kjellberg, F., 2003. Convergence and Coevolution in a mutualism: evidence from a molecular phylogeny of *Ficus*. Evolution 57, 1255–1272.
Kay, K.M., Reeves, P.A., Olmstead, R.G., Schemske, D.W., 2005. Rapid speciation and the evolution of hummingbird pollination in neotropical Costus subgenus Costus (Costaceae): evidence from nrDNA ITS and ETS sequences. Am. J. Bot. 92, 1899–1910.
Koenen, E.J.M., Clarkson, J.J., Pennington, T.D., Chatrou, L.W., 2015. Recently evolved diversity and convergent radiations of rainforest mahoganies (Meliaceae) shed new light on the origins of rainforest hyperdiversity. New Phytol. 207, 327–339.
Kreft, H., Jetz, W., 2007. Global patterns and determinants of vascular plant diversity. Proc. Natl. Acad. Sci. U.S.A. 104, 5925–5930.
Landis, M.J., Matzke, N.J., Moore, B.R., Huelsenbeck, J.P., 2013. Bayesian analysis of biogeography when the number of areas is large. Syst. Biol. 62, 789–804.
Leite, Y., Costa, L., Loss, A., Rocha, R., Batalha-Filho, H., Bastos, A., Quaresma, V.,

Fagundes, V., Paresque, R., Passamani, M., Pardini, R., 2016. Neotropical forest expansion during the last glacial period challenges refuge hypothesis. Proc. Natl. Acad. Sci. U.S.A. 113, 1008–1013.
Li, M., Wunder, J., Bissoli, G., Scarponi, E., Gazzani, S., Barbaro, E., Saedler, H., Varotto, C., 2008. Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. Cladistics 24, 727–745.
Lomáscolo, S.B., Levey, D.J., Kimball, R.T., Bolker, B.M., Alborn, H.T., 2010. Dispersers shape fruit diversity in *Ficus* (Moraceae). Proc. Natl. Acad. Sci. U.S.A. 107, 14668–14672.
Lomáscolo, S.B., Speranza, P., Kimball, R., 2008. Correlated evolution of fig size and color supports the dispersal syndromes hypothesis. Oecologia 156, 783–796.
Loss-Oliveira, L., Sakuragui, C., Soares, M., Schrago, C., 2016. Evolution of *Philodendron* (Araceae) species in Neotropical biomes. PeerJ 4, 1–18.
Lucas, E.J., Matsumoto, K., Harris, S.A., Nic Lughadha, E.M., Benardini, B., Chase, M.W., 2011. Phylogenetics, morphology, and evolution of the large genus *Myrcia* s.l. (Myrtaceae). Int. J. Pl. Sci. 172, 915–934.
Maddison, W.P., Maddison, D.R., 2015. Mesquite: a Modular System for Evolutionary Analysis. Version 3.02. < http://mesquiteproject.org > .
Maddison, W.P., Midford, P.E., Otto, S.P., 2007. Estimating a binary character's effect on speciation and extinction. Syst. Biol. 56, 701–710.
Martins, F.M., 2011. Historical biogeography of the Brazilian Atlantic forest and the Carnaval-Moritz model of Pleistocene refugia: what do phylogeographical studies tell us? Biol. J. Linn. Soc. 104, 499–509.
Massana, K.A., Beaulieu, J.M., Matzke, N.J., O'Meara, B.C. 2015. Non-null effects of the null range in biogeographic models: Exploring parameter estimation in the DEC model, bioRxiv. September 16, 2015. https://doi.org/10.1101/026914, http://biorxiv.org/content/early/2015/09/16/026914.abstract.
Matzke, N.J., 2013a. BioGeoBEARS: Biogeography with Bayesian (and likelihood) evolutionary analysis in R Scripts. CRAN: The Comprehensive R Archive Network. Vienna, Austria. http://cran.r-project.org/package=BioGeoBEARS.
Matzke, N.J., 2013b. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Front. Biogeogr. 5, 242–248.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), November 2010, New Orleans, pp. 1–8.
Morley, R.J., 2000. Origin and Evolution of Tropical Rain Forests. John Wiley & Sons Ltd., Chichester, England, pp. 1–362.
Morrone, J., 2014. Biogeographical regionalisation of the Neotropical region. Zootaxa 3782, 1–110.
Pederneiras, L.C., Romaniuc-Neto, S., Mansano, V.F., 2015. Molecular phylogenetics of *Ficus* section *Pharmacosycea* and the description of *Ficus* subsection *Carautaea* (Moraceae). Syst. Bot. 40, 504–509.
Pennington, R.T., Dick, C.W., 2004. The role of immigrants in the assembly of the South American rainforest tree flora. Philos. Trans. R. Soc. Lond. B, Biol. Sci. 359, 1611–1622.
Pennington, R., Cronk, Q., Richardson, J., 2004. Introduction and synthesis: plant phylogeny and the origin of major biomes. Philos. Trans. Roy. Soc. B 359, 1455–1464.
Pennington, R.T., Hughes, M., Moonlight, P.W., 2015. The origins of tropical rainforest hyperdiversity. Trends Plant Sci. 20, 693–695.
Pereira, R., Rodrigues, E., Menezes, A., 2007. Phenological patterns of Ficus citrifolia (Moraceae) in a seasonal humid-subtropical region in Southern Brazil. Plant Ecol. 188, 265–275.
Perret, M., Chautems, A., Araujo, O.A., Salamin, N., 2013. Temporal and spatial origin of Gesneriaceae in the New World inferred from plastid DNA sequences. Bot. J. Linn. Soc. 171, 61–79.
Prance, G.T., 1982. Forest refuges: evidence from woody angiosperms. In: Prance, G.T. (Ed.), Biological Diversification in the Tropics. Columbia University Press, New York, pp. 137–158.
Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.
R Core Team, 2016. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. < http://www.R-project.org/ > .
Rambaut, A., 2014. FigTree version 1.4.2: tree figure drawing tool. Version 1.4.2. < http://tree.bio.ed.ac.uk/software/figtree > .
Rambaut, A., Drummond, A.J., 2013. Tracer version 1.6. < http://beast.bio.ed.ac.uk/Tracer > .
Ramírez, W.B., 1970. Host specificity of fig wasps (Agaonidae). Evolution 24, 680–691.
Ramírez, B.W., 1977. Evolution of the strangling habit in *Ficus* L. subgenus *Urostigma* (Moraceae). Brenesia 12, 11–19.
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.
Ricommini, C., Assumpção, M., 1999. Quaternary tectonics in Brazil. Episodes 22, 221–225.
Roncal, J., Kahn, F., Millan, B., Couvreur, T.L.P., Pintaud, J.-C., 2013. Cenozoic colonization and diversification patterns of tropical American palms: evidence from *Astrocaryum* (Arecaceae). Bot. J. Linn. Soc. 171, 120–139.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Rønsted, N., Salvo, G., Savolainen, V., 2007. Biogeographical and phylogenetic origins of African fig species (*Ficus* section *Galoglychia*). Mol. Phylogenet. Evol. 43, 190–201.
Rønsted, N., Weiblen, G.D., Clement, W.L., Zerega, N.J.C., Savolainen, V., 2008a. Reconstructing the phylogeny of figs (*Ficus*, Moraceae) to reveal the history of the fig pollination mutualism. Symbiosis 45, 45–55.
Rønsted, N., Weiblen, G.D., Cook, J.M., Salamin, N., Machado, C.A., Savolainen, P., 2005. 60 million years of co-divergence in the fig-wasp symbiosis. Proc. R. Soc. B. 272, 2593–2599.
Rønsted, N., Weiblen, G.D., Savolainen, V., Cook, J.M., 2008b. Phylogeny, biogeography, and ecology of *Ficus* section *Malvanthera* (Moraceae). Mol. Phylogenet. Evol. 48, 12–22.
Santos, A.M.M., Silva, J.M.C., Tabarelli, M., 2007. Biogeographical relationships among tropical forests in northeastern Brazil. J. Biogeograp. 34, 437–446.
Soetaert, K., 2013. plot3D: Plotting multi-dimensional data. R package version 1.0.
Staggemeier, V.G., Diniz-Filho, J.A.F., Forest, F., Lucas, E., 2015. Phylogenetic analysis in Myrcia section Aulomyrcia and inferences on plant diversity in the Atlantic rainforest. Ann. Bot. 115, 747–761.
Stamatakis, A., 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Stebbins, G.L., 1974. Flowering Plants: Evolution above the Species Level. Harvard University Press, Cambridge, MA, pp. 1–399.
Strand, A.E., Leebensmack, J., Milligan, B.H., 1997. Nuclear DNA-based markers for plant evolutionary biology. Mol. Ecol. 6, 113–118.
Särkinen, T.E., Newman, M.F., Maas, P.J.M., Maas, H., Poulsen, A.D., Harris, D.J., Richardson, J.E., Clark, A., Hollingsworth, M., Toby Pennington, R., 2007. Recent oceanic long-distance dispersal and divergence in the amphi-Atlantic rain forest genus Renealmia L.f. (Zingiberaceae). Mol. Phylogenet. Evol. 44, 968–980.
Thiers, B. [continuously updated]. Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. < http://sweetgum.nybg.org/science/ih/ > .
Thomé, M.T.C., Zamudio, K.R., Giovanelli, J.G.R., Haddad, C.F.B., Baldissera Jr, F.A., Alexandrino, J., 2010. Phylogeography of endemic toads and post-Pliocene persistence of the Brazilian Atlantic Forest. Mol. Phylogenet. Evol. 55, 1018–1031.
Torke, B.M., Schaal, B.A., 2008. Molecular phylogenetics of the species-rich Neotropical genus *Swartzia* (Leguminosae, Papilionoideae) and related genera of the swartzioid clade. Am. J. Bot. 95, 215–228.
Traveset, A., 1998. Effect of seed passage through vertebrate frugivores' guts on germination: a review. Persp. Plant Ecol. Evol. Syst. 1, 151–190.
Wang, W., Ortiz, R., del, C., Jacques, F.M.B., Xiang, X.G., Li, H.L., Lin, L., Li, R.Q., Liu, Y., Soltis, P.S., Soltis, D.E., Chen, Z.D., 2012. Menispermaceae and the diversification of tropical rainforests near the Cretaceous-Paleogene boundary. New Phytol. 195, 470–478.
Wiebes, J.T., 1979. Co-evolution of figs and their insect pollinators. Annu. Rev. Ecol. Syst. 10, 1–12.
Zachos, J., Pagani, M., Sloan, L., Thomas, E., Billups, K., 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. Science 292, 686–693.
Zerega, N., Clement, W., Datwyler, S., Weiblen, G., 2005. Biogeography and divergence times in the mulberry family (Moraceae). Mol. Phylogenet. Evol. 37, 402–416.

Chapter III

Rush hour at the Museum – Diversification patterns provide new clues for the success of figs (*Ficus* L., Moraceae)

Acta Oecologica 90 (2018) 4–11

Contents lists available at ScienceDirect

Acta Oecologica

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

Rush hour at the Museum – Diversification patterns provide new clues for the success of figs (*Ficus* L., Moraceae)

Sam Bruun-Lund[a,*], Brecht Verstraete[a], Finn Kjellberg[b], Nina Rønsted[a]

[a] Natural History Museum of Denmark, University of Copenhagen, Sølvgade 83S, DK-1350 Copenhagen, Denmark
[b] CEFE UMR 5175, CNRS, Université de Montpellier, Université Paul-Valéry Montpellier, EPHE, F-34293 Montpellier Cédex 5, France

ARTICLE INFO

Keywords:
Diversification rates
Ficus
Key-innovations
Rainforest diversity
Hemi-epiphytes

ABSTRACT

Tropical rainforests harbour much of the earth's plant diversity but little is still known about how it evolved and why a small number of plant genera account for the majority. Whether this success is due to rapid turnover or constant evolution for these hyper-diverse plant genera is here tested for the species-rich genus *Ficus* L. (figs). The pan-tropical distribution of figs makes it an ideal study group to investigate rainforest hyper-diversification patterns. Using a recently published, dated and comprehensive phylogenetic hypothesis, we infer that figs are an old lineage that gradually accumulated species and exhibits very low extinction rates, which corresponds to the 'museum model' of evolution. Overall, no major significant shifts in evolutionary dynamics are detected, yet two shifts with lower probability are found. Hemi-epiphytism, monoecy, and active pollination are traits that possibly are associated with the hyper-diversity found in figs, making it possible for the plants to occupy new niches followed by extensive radiation over evolutionary time scales. Figs possess unique diversification patterns compared to other typical rainforest genera.

1. Introduction

Rainforests harbor the majority of biodiversity on Earth and are viewed as old ecosystems that have been prevalent throughout much of the Cenozoic (65-0 Ma). However, rainforest hyper-diversity is not well understood and it has variously been attributed to be a museum of diversity, showing constant speciation and low extinction (Wallace, 1878; Stebbins, 1974), or a cradle of diversity, referring to more recent and rapid speciation (Pennington et al., 2015; Richardson et al., 2001). Previous phylogenetic studies of diversification patterns in typical rainforest plant families, such as Annonaceae, Arecaceae, and Menispermaceae, have suggested that they are old lineages having experienced a gradual and more or less constant lineage accumulation over time, supporting the museum model (Couvreur et al., 2011a, 2011b; Wang et al., 2012). In contrast, other studies have inferred more recent radiations of rainforest lineages at the genus level for *Inga, Costus, Guatteria*, and *Renealmia*, supporting the cradle model of recent diversification (Erkens et al., 2007; Kay et al., 2005; Richardson et al., 2001; Särkinen et al., 2007). Further complicating the picture, Koenen et al. (2015) recently documented that despite the Meliaceae family originated already in the Eocene (56-34 Ma), the majority of the rainforest species diversity is recent and associated with higher speciation rates compared to non-rainforest lineages.

In line with Couvreur et al. (2011b) who suggested a mixed model of steady processes and mixed diversification, Koenen et al. (2015) proposed a concept of highly dynamic diversification processes across ecosystems that are linked to environmental changes (Xing et al., 2014) rather than to climatic stability, as suggested by the museum model. This view is, for example, supported by the Miocene (23-5 Ma) radiation of succulents (Arakaki et al., 2011) and grasses (Edwards et al., 2010), which was interpreted as an effect of global cooling and the subsequent spread of seasonally dry ecosystems. Furthermore, fossil evidence has shown that Eocene rainforests were more diverse than contemporary rainforests but experienced a period of decreasing diversity up to the Early Miocene, implying both ancient rapid diversification and high past extinction rates (Jaramillo et al., 2010; Wilf et al., 2003). More examples are needed to help clarify if there are general patterns explaining rainforest diversification or if individual plant groups exhibit specific patterns.

Figs (*Ficus* L., Moraceae) are an old Cretaceous (75-90 Ma; Cruaud et al., 2012; Rønsted et al., 2005; Zerega et al., 2005) but mega-diverse (~800 species; Berg and Corner, 2005) and a significant component of tropical forests with many species having wide distribution ranges and ubiquitously high alpha-diversity in lowland tropical rainforests (Berg and Corner, 2005; Harrison, 2005). Harrison (2005) argued that no diversity pattern of any other genus resembles that of figs: other diverse

* Corresponding author. Tel.: +45 35 32 22 22.
E-mail address: sam.bruunlund@snm.ku.dk (S. Bruun-Lund).

https://doi.org/10.1016/j.actao.2017.11.001
Received 17 February 2017; Received in revised form 11 October 2017; Accepted 1 November 2017
Available online 13 November 2017

genera have either geographically restricted distributions or are especially diverse in only one region. This indicates that figs possess certain traits that make them exceptionally well adapted to a broad range of niches, which has led to their high diversity all over the tropics (Corner, 1961).

A range of different traits have been suggested as possible drivers for fig diversification (Harrison, 2005; Jousselin et al., 2003a; Weiblen, 2004). They exhibit a variety of growth forms, including typical rainforest habits like hemi-epiphytes (stranglers and banyans), trees (with/without buttress roots), and large woody climbers. Furthermore, many species are pioneers of forest succession and possess many pioneer traits, such as small seeds (Harrison, 2005). Furthermore, fig wasps have evolved a unique active pollination behaviour that is very uncommon in most other plants except a few known cases such as in *Yucca*. Fig wasps use coxal combs to collect pollen into pollen pockets and deposit the pollen onto receptive flowers before oviposition. Active pollination is found in two thirds of the species. For the other third of the genus, passive pollination happens when wasps carry pollen dispersed on their bodies without specialized behaviour (Cook and Rasplus, 2003; Kjellberg et al., 2001). Due to their obligate pollination mutualism with short-lived wasps (Agaonidae), crops are available at population level year-round, providing an important food source for over 1200 mammal species globally (Shanahan et al., 2001).

In addition, quick responses to climatic changes and/or the presence of key innovations could promote high diversification (Drummond et al., 2012; Heard and Hauser, 1995; Hodges et al., 1995; Mayhew, 2007; Sanderson and Donoghue, 1994; Silvestro et al., 2014; Vamosi and Vamosi, 2011). The idea of key innovations as a single causal explanation for the rapid diversification of a lineage has largely been abandoned in favour of more complex models developed across angiosperms (Smith et al., 2011) and using more nuanced explanations involving multiple interacting traits assembled step-wise through evolution referred to as *synnovation* (Donoghue and Sanderson, 2015). Therefore, traits that are shared among species from diverse genera could very well be responsible for the rise of high species richness in general (Lovette et al., 2002). Certain comparative phylogenetic methods allow for testing and comparing the influence of certain traits on diversification patterns. Using those techniques to study the evolutionary history of a pan-tropical diverse genus such as the figs, and to identify diversification rates shifts and/or specific traits linked to higher diversification and survival might provide insights into the evolution of tropical biodiversity across ecosystems and continents in general.

Using a time-calibrated phylogenetic tree of figs, we address the following questions: (1) does the diversification pattern of figs support the museum model of ancient gradual diversification or the cradle model of recent rapid radiations? (2) Do significant diversification shifts within *Ficus* exist and if so, (3) are they correlated with specific traits representing one or more key-innovations?

2. Materials and methods

2.1. Taxon sampling

As input for the different diversification rate analyses, the most recently published dated phylogenetic hypothesis of figs was used (Cruaud et al., 2012). This phylogenetic hypothesis is also the most comprehensive, including 200 species (> 25% of the diversity of the genus) and was constructed based on the 5 low or single copy nuclear markers ITS, ETS, *G3pdh*, *ncpGS*, and *waxy*. Sampling broadly covers the diversity of *Ficus* with section *Oreosycea* and subgenus *Synoecia* being the least inclusive with 12% of the species of these lineages included (Supplemetary data S1). Although support and resolution of certain infrageneric relationships within *Ficus* are still low, the phylogenetic tree of Cruaud et al. (2012) represents by far the best hypothesis available to date rooted with four taxa of Castilleae, the sister lineage of *Ficus* (Clement and Weiblen, 2009).

2.2. Diversification rate analyses

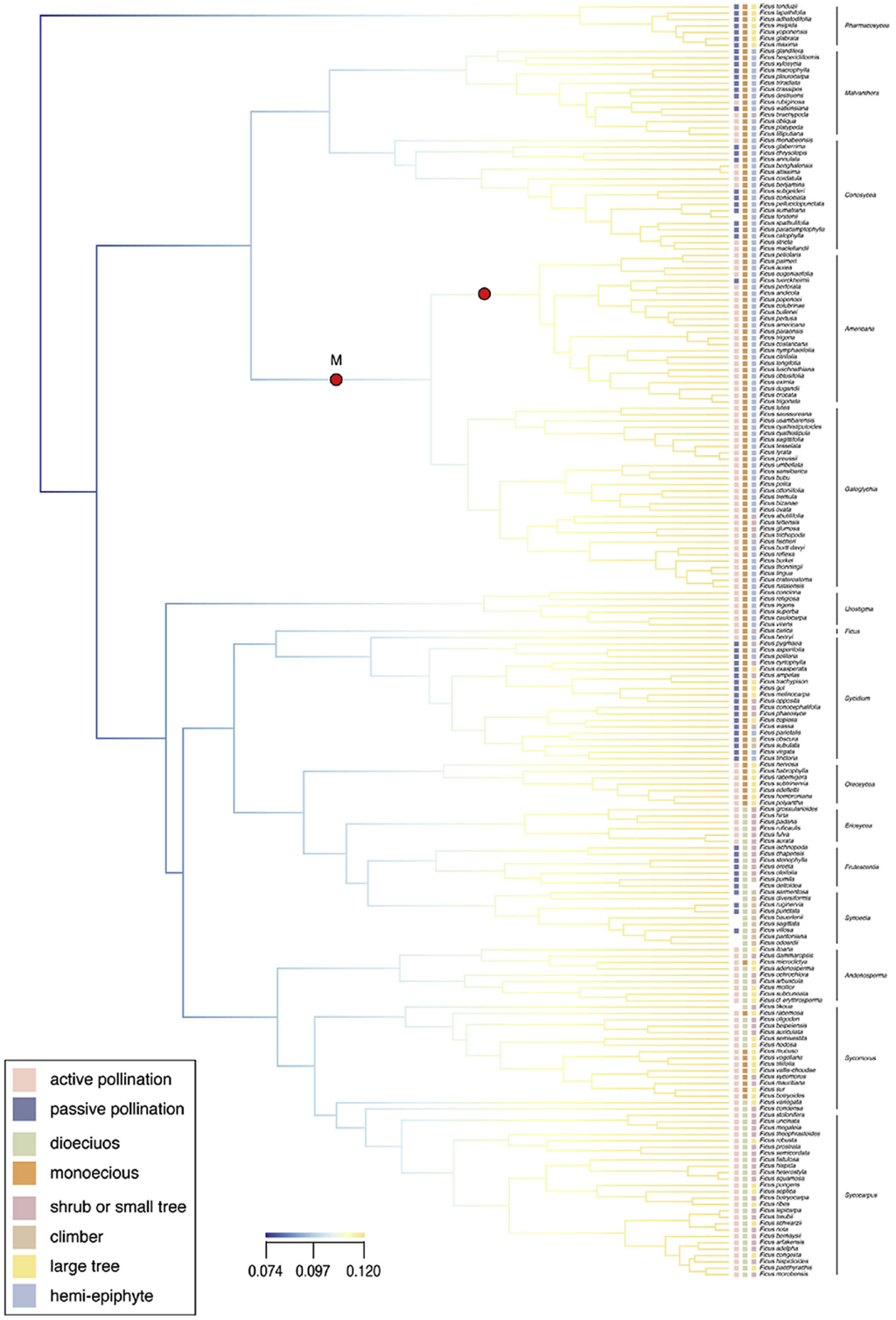
A lineage through time (LTT) plot for the entire genus was generated with the R package paleotree 2.7 (Bapst, 2012). One thousand random post-burnin trees from the BEAST inference analysis by Cruaud et al. (2012) were used as input. A median curve with the 95% confidence interval is displayed.

A Bayesian Analysis of Macro-evolutionary Mixtures (BAMM 2.5) (Rabosky et al., 2014) was used to infer shifts in speciation and extinction rates across the phylogenetic tree. To account for non-random incomplete taxon sampling, total diversity and sampling fractions for all clades (see Supplementary data file S1) within figs were obtained from the literature (Berg and Corner, 2005; Rønsted et al., 2008a). The MCMC analysis was run for 1 million generations with sampling every 1000 generations and, after checking for convergence with the R package coda 0.19–1 (Plummer et al., 2006), the first 10% was discarded as burn-in. We used the R package BAMMtools 2.1.6 (Rabosky et al., 2014) and ape 4.0 (Paradis et al., 2004) to summarize rates over each branch of the phylogenetic tree (visualized in the so-called 'phylorate' plot), to plot the 95% credible shift set (CSS) with sampling frequencies of the different shift configurations, and to obtain the shift configuration with the maximum a posteriori (MAP) probability.

Problems with estimating extinction rates have been reported for earlier versions of BAMM (Moore et al., 2016) and although they have been accounted for in the latest version, we decided to include a MEDUSA analysis as a second method to infer diversification rate shifts (Alfaro et al., 2009). MEDUSA requires species richness to be assigned to unresolved clades and we therefore pruned the original dated phylogenetic tree of Cruaud et al. (2012) to section level and assigned diversity to each terminal. The R package geiger 2.0.6 (Alfaro et al., 2009) was used to perform the MEDUSA analysis.

2.3. State-dependent diversification rate analyses

State-dependent diversification rate analyses were performed with the Binary-State Speciation and Extinction (BiSSE) model (Maddison et al., 2007) and the MultiState Speciation and Extinction (MuSSE) model, as implemented in the R package diversitree 0.9–9 (Fitzjohn, 2012). The influence of two traits on the diversification rates was analysed using the BiSSE model: 1) pollination mode with the states active versus passive pollination, and 2) monoecious versus (gyno)-dioecious plants. The influence of habit was analysed using the MuSSE model and four different states were defined: shrub or small tree (< 30 m, understory), climber, large tree (≥ 30 m, canopy), and hemi-epiphyte. Information on the trait states for each taxon was extracted from the literature (Fig. 1; Berg and Villavivencio, 2004; Berg, 2012; Berg and Corner, 2005; Berg and Wiebes, 1992; Burrows and Burrows, 2003; Corner, 1967; Zhengyi et al., 2003; supplemented by our own taxonomic knowledge). For each analysis, we corrected for state-specific incomplete taxon sampling. Likelihood ratio tests were performed to test for significant differences between the unconstrained and constrained rate models and these found the unconstrained model to be the best-fit model for all characters. The MCMC analysis was performed with a standard Cauchy distribution as hyper-prior to avoid zero rates (Burin et al., 2016) and was run for 10,000 generations with sampling every 1000 generations. Chain convergence and effective sampling size (ESS) parameters were inspected with the R package coda 0.19–1 (Plummer et al., 2006). The mean speciation, extinction, and net diversification rates and their respective 95% credibility intervals were calculated for each state of each character. A Cohen's d effect size test was performed to test for differences between these means. The results were visualized in posterior probability distribution plots.



(caption on next page)

Fig. 1. Phylorate plot of *Ficus* with branches colored according to speciation rate (lineages/Ma), resulting from the Bayesian Analysis of Macro-evolutionary Mixtures (BAMM) analysis. For 70% of the samples in the posterior distribution no shift is observed. However, at a lower probability, a shift in diversification dynamics is observed at the lineage leading to section *Americanae* (20% of the samples in the posterior distribution) and at the lineage containing both section *Americanae* and section *Galoglychia* (3.6% of the samples in the posterior distribution), as indicated with red circles. The M is located on the branch where a rate shift occurs according to the MEDUSA analysis. The scored traits along with the species are indicated on the tips of the tree. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Results

3.1. Diversification shifts

The BAMM analysis revealed that 70% of the samples in the posterior distribution exhibit no shifts in speciation rates, thus a constant model of evolution is favoured with no sudden increase in diversification rates for any clades (Fig. 1). However, shifts in diversification rates are detected with lower probability at the branch leading to section *Americanae* (20% of the samples in the posterior distribution) and at the branch leading to the clade containing both section *Americanae* and section *Galoglychia* (3.6% of the samples in the posterior distribution) (Fig. 1; red circles & Supplementary data file S3). The detected shifts with a lower probability (20% and 3.6%) are at a lower marginal shift probability and therefore not as well supported and significant as a model of constant evolutionary rate. MEDUSA identified significant acceleration of net speciation rates at one point in the tree where a branch leading to the clade containing both section *Americanae* and section *Galoglychia* (Fig. 1; red circle with an M). Overall net diversification rates inferred by MEDUSA and BAMM are very similar, confirming the robustness of our result.

The LLT plot shows a gradual increase in lineage accumulation through time, meaning that no sudden radiation of lineages is detected at any point (Fig. 2). This demonstrates consistency between the patterns inferred by BAMM, where no strongly supported shifts in rate dynamics is observed either.

3.2. State-dependent diversification rates

Following the methodology of the Cohen d effect size test, the differences in diversification rates between traits found in the BiSSE analysis are defined as 'very small', 'small', 'large', 'very large', and 'huge' and this terminology is used in the description of the results. See supplementary data file S2 for all inferred rates by BiSSE and MuSSE analysis.

Pollination mode. The BiSSE analysis inferred a huge difference in net diversification rates between active and passive pollination: species with active pollination ($r = 0.071$) diversify faster than species with passive pollination ($r = 0.052$) – see supplementary data file S2. There is a large difference in extinction and speciation rate between the two traits, with extinction rate being higher for passively pollinated fig species and speciation faster for actively pollinated figs, as also suggested in Kjellberg et al. (2001). The transition rate from passive to active pollination was found to be larger than the reverse.

Monoecy and dioecy. When comparing monoecious and dioecious figs, the BiSSE method inferred a huge difference in net diversification rate between monoecious and dioecious fig species. We found that monoecious figs evolve faster ($r = 0.076$) than dioecious species ($r = 0.055$). Extinction rate is similar with a very small difference between the two trait states (thus, less different than for the mode of pollination). The speciation rate is hugely different between dioecious and monoecious species, with monoecious species evolving faster than the dioecious ones (see supplementary data file S2). The transition rate from dioecy to monoecy is also much higher than the reverse.

Life forms (habits) of Figs. In order to account for the various life forms of figs, a state-dependent diversification analysis using the MuSSE model was performed and four trait states were defined: shrubs/small trees, large trees, climbers, and hemi-epiphytes. Net diversification rate of hemi-epiphytes ($r = 0.076$) is much higher than net diversification for the three other life forms (shrubs/small trees, large trees, climbers) which also show very similar rates (r between 0.049 and 0.057). The speciation and extinction rates for all life forms are somewhat similar, however, the hemi-epiphytic figs show a slightly higher speciation rate and a slightly lower extinction rate, resulting in the generally higher net diversification rate.

Transition rates between life forms. The rate of transitions between life forms varies considerably (see Fig. 3). The highest transition rates are inferred for figs transitioning from large canopy trees (≥ 30 m) to smaller understory trees (< 30 m) or shrubs (mean = 0.033), as well as for the opposite direction (mean = 0.029). The transition rate from either of these life forms as well as climbers to hemi-epiphytic figs is low (from 0.006 up to 0.035). However, the transition rates from hemi-epiphytes to any of the other three life forms are much lower – all ranging between 0.001 and 0.003.

4. Discussion

Our results, which are based on the largest and most recent phylogenetic tree of figs, provide possible explanations for some of the diversification patterns that have been hypothesized up till now. Our aim was to (1) test if the diversification patterns of *Ficus* support a museum model of ancient gradual diversification or a cradle model of recent, rapid radiation. Also, (2) to see if any significant shifts in diversifications patterns occur within *Ficus* and (3) if they are linked with specific traits.

4.1. Diversification patterns

To investigate diversification patterns in figs, we constructed a lineage-through-time (LTT) plot and performed an analysis using both BAMM and MEDUSA to test how lineages evolve and diversify through time. Diversification of figs seems to correspond with a museum model of constant evolution and little extinction, leading to the gradual accumulation of lineages, possibly in response to long-lasting and stable tropical ecosystems (Figs. 1 and 3). However, it is worth noting that other examples of rainforest lineages supporting the museum model of gradual diversification are detected for pan-tropical rainforest tree families (Couvreur et al., 2011a,b; Wang et al., 2012), whereas generic level studies have primarily focused on the Neotropics and largely support the cradle model of recent diversification (Richardson et al., 2001; Kay et al., 2005; Erkens et al., 2007; Särkinen et al., 2007). These patterns could suggest that Neotropical rainforests may be cradles of recent diversity within a pantropical museum model.

Consequently, the diversification patterns we have found lend support to the statement by Harrison (2005) that no other genus compares to the diversity pattern of *Ficus* as other diverse genera have either geographically restricted distributions or are especially diverse in only one region. Furthermore, the overall net diversification rate for the genus ($r = 0.1$) is higher, largely explained by a lower extinction rate ($\mu = 0.009$), than the overall net diversification rate inferred across all angiosperms ($r = 0.07–0.09$) (Magallon and Sanderson, 2001).

Overall, the BAMM analysis inferred no significant rate shifts in most of the samples from the 95% credibility set, but at a lower marginal probability two shifts towards increased speciation rates are found (Fig. 1). The shift with the second highest probability is found at the lineage leading to section *Americanae* and the third rate shift is detected at the branch leading towards the clade containing both section *Americanae* and section *Galoglychia* (Fig. 1; red circles). The latter rate shift was also detected by the MEDUSA analysis on the branch leading towards a clade with both sections. Even though the analysis do detect a

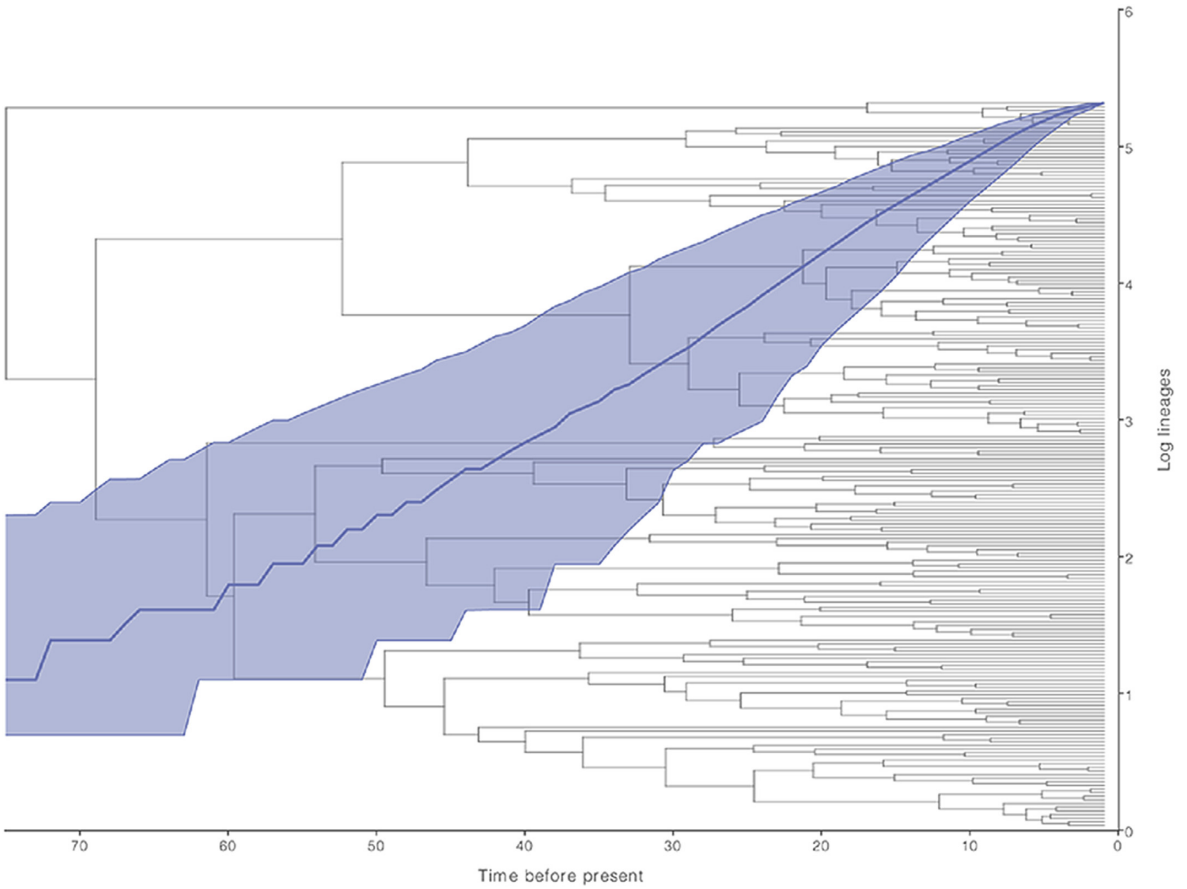


Fig. 2. Lineage-though-time plot depicted on the dated phylogenetic tree from Cruaud et al. (2012). A gradual increase of lineages through time is observed and no sudden bursts of speciation are present within the genus.

shift in rate diversification this is only at a very low probability and the most supported analysis detects no shifts in the genus.

4.2. State-dependent diversification patterns

The aim of the diversification analyses was to explore the potential link between traits or key innovations and varying diversification rates. Our separate results show higher diversification rates linked to figs that are actively pollinated, monoecious, and hemi-epiphytes compared to passively pollinated, dioecious figs exhibiting other life forms. Furthermore, these traits all occur in clades associated with higher diversification rates from BAMM and MEDUSA (sections *Americanae* and *Galoglychia*; see Fig. 1). It should also be noted that the traits were analysed separately and are perhaps not necessarily linked together.

Active pollination in fig species is positively linked with higher diversification rates, i.e. species that are actively pollinated evolve faster. It has previously been hypothesized that active pollination could be a driver of speciation, which is in line with our results (e.g. Harrison, 2005; Jousselin et al., 2003b; Kjellberg et al., 2001). It has also been suggested that active pollination might have evolved as a way for wasps to ensure fertilization of the flowers in which they oviposit (Jousselin et al., 2003a). Harrison (2005) suggests that maintaining high levels of heterozygosity due to ensured outcrossing, together with a very high proportion of inflorescences pollinated could also be a driver of faster speciation in actively pollinated figs. Also, active pollination could be maintained over evolutionary time by actively pollinating species being less prone to extinction than passively pollinating ones (Kjellberg et al., 2001), which is confirmed by results exhibiting a lower extinction rate.

The efficiency of active pollination has enabled figs to occupy rare (micro)-niches and it is therefore an important factor for the high alpha-diversity in figs, since the ability to outcross at extremely low population densities may also have lowered extinction rates (Harrison, 2005). Our results support these hypotheses by inferring an extinction rate almost twice as low for actively pollinating species (0.007 vs. 0.016) compared to passively pollinated species.

Monoecy or (gyno)-dioecy. Being monoecious is correlated with increased rates of diversification in figs. In contrast to angiosperms in general (Käfer et al., 2014), monoecious fig species have a higher diversification rate and this has also been shown elsewhere for other plant species (Laenen et al., 2016), which has been explained by the long-term advantages of outcrossing. In figs, however, this does not seem to be the case. It could instead be explained by self-compatible bisexual species being more likely to establish new colonies following long-distance dispersal than unisexual or self-incompatible ones, potentially promoting allopatric speciation and finding new niches (Heilbuth, 2000; Laenen et al., 2016).

Life forms. Figs that are hemi-epiphytic have an increased diversification rate compared to other life forms, which all show a somewhat similar net diversification rate (see supplementary data file S2). Around half of the species are hemi-epiphytic, further suggesting that this life

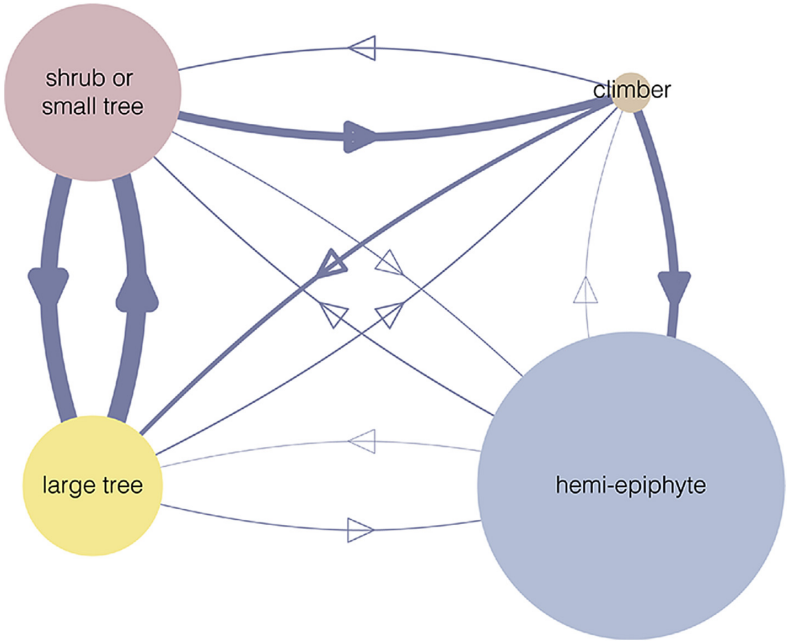


Fig. 3. Transitions rates between life forms (habits). The circle size is proportional to the number of species with that trait state and the thickness of the arrows is proportional to the transition rates. Hardly any species evolve from a hemi-epiphytic life form to any of the three other states.

form is advantageous for figs (Cruaud et al., 2012; Harrison, 2005). They seem to possess some key innovations or a syndrome of advantageous traits explaining their diversification success. The ability to grow on rocky outcrops using aerial roots has probably been very important in the diversification and evolution of figs. Even though hemi-epiphytic figs are present in low population densities, most hemi-epiphytic species have large ranges related to a more diffuse dispersal mechanism as opposed to terrestrial species, which tend to have more locally restricted dispersal abilities (Harrison and Rasplus, 2006). Dispersal range also differs between dioecious figs, which may be more restricted compared to monoecious figs. Whereas pollinators of monoecious figs disperse by ascending above the canopy and drifting in the wind until they reach the fume of receptive fig odour emitted by a tree, the pollinators of dieocious figs disperse more locally and do not use the wind for dispersal (Harrison and Rasplus, 2006).

The hemi-epiphytic life form allows for occupation of niches that most plants do not readily colonize. Hemi-epiphytes are adapted to grow in the canopy layer with low water availability and high light conditions. Adaptive root flexibility may also explain transitioning from forest habitats to similar dry environments expressed by lithophytes in Australia (Harrison, 2005; Harrison et al., 2003; Harrison and Shanahan, 2005; Rønsted et al., 2008b).

Transitioning between life forms. The rate of which the traits tested evolve and transition into each other is an indication of the diversification dynamics of figs and may offer explanations as to how they became so successful. We scored large canopy trees and small understory trees/shrubs as separate life forms, but our results indicate that it is easy to transition between these two states, probably because it does not require large evolutionary sacrifices or changes.

The rate at which figs become hemi-epiphytic is not high but the rate of hemi-epiphytic figs transitioning back towards another life form is very low (Fig. 3). Consequently, the hemi-epiphytic life-form appears to be a more stable state providing an adaptive advantage and explaining why such a high proportion of the figs are hemi-epiphytic (Berg and Corner, 2005).

4.3. What has made figs so successful?

The origin of figs has been dated to the late Cretaceous (75-90 Ma; Cruaud et al., 2012) – which is before the K/Pg boundary (~65 Ma) where a lot of plant diversity is hypothesized to have gone extinct (Wilf and Johnson, 2004). For figs, this might have created opportunities to expand into new niches, as they are generally opportunistic species with low competitive abilities (Berg and Corner, 2005; Harrison, 2005). After having established in such new niches, figs would have kept evolving at a constant rate as competition is low in many (micro)-niches (Frenzke et al., 2016; Gentry and Dodson, 1987). Figs do possess many traits typical of pioneer plants – small seeds, high assimilation and growth rates, high fecundity, and flexible rooting habits – that are advantageous in colonising new areas as also expressed by the genus *Piper* and epiphytic Orchids (Harrison, 2005).

Furthermore, the renowned pollination system of figs with its unique fig-pollinating wasps ensuring high pollination success is beneficial when population densities are low. Even though fig species and their pollinators show high levels of co-diversification and specificity (Cruaud et al., 2012), multiple pollinators for a single fig species are frequently reported (Machado et al., 2005). Hybridization and introgression could also have contributed to the large genetic and species diversity in figs (Bruun-Lund et al., 2016).

5. Conclusion

Overall, the diversification rate analyses show that figs generally follow the museum model of evolution with a gradual accumulation of species over time and with very low extinction rates and no significant evolutionary shifts. However, the trait state-dependent analyses show that monoecy, active pollination, and having a flexible root habit (i.e. hemi-epiphytes) are linked with higher diversification rates and this is further underpinned by these character states being expressed by a subclade with higher diversification rate. These trait states are found throughout most of the phylogenetic tree suggesting that they represent key innovations or syndromes, possibly responsible for making the figs so successful in terms of species diversity and accumulating high

species diversity over an evolutionary time scale.

Contribution of the authors

SBL and NR conceived the ideas. SBL assembled the data, accounted for incomplete sampling, selected and scored the trait states with help from FK and NR. BV designed and performed the diversification rate analyses and helped interpreting the data. SBL wrote the manuscript with contributions from NR. All authors improved and approved the manuscript.

Compliance with ethical standards

This study does not infringe on any bioethical principles and no damage to biodiversity was inflicted while carrying out this study.

Conflicts of interest

The authors declare that they have no conflict of interest.

Acknowledgements

This project has received funding from the Danish Council for Independent Research grant No DFF – 7014-00205 to NR and from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 656853 to NR and BV.

Appendix B. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.actao.2017.11.001.

References

Alfaro, M.E., Santini, F., Brock, C., Alamillo, H., Dornburg, A., Rabosky, D.L., Carnevale, G., Harmon, L.J., 2009. Nine exceptional radiations plus high turnover explain species diversity in jawed vertebrates. Proc. Natl. Acad. Sci. 106, 13410–13414. http://dx.doi.org/10.1073/pnas.0811087106.

Arakaki, M., Christin, P.-A., Nyffeler, R., Lendel, A., Eggli, U., Ogburn, R.M., Spriggs, E., Moore, M.J., Edwards, E.J., 2011. Contemporaneous and recent radiations of the world's major succulent plant lineages. Proc. Natl. Acad. Sci. U. S. A. 108, 8379–8384. http://dx.doi.org/10.1073/pnas.1100628108.

Bapst, D.W., 2012. Paleotree: an R package for paleontological and phylogenetic analyses of evolution. Methods Ecol. Evol. 3, 803–807. http://dx.doi.org/10.1111/j.2041-210X.2012.00223.x.

Berg, C.C., Villavivencio, X., 2004. Taxonomic Studies on *Ficus* (Moraceae) in the West Indies, Extra-Amazonian Brazil and Bolivia, Ilicifolia. Universitetet i Bergen, Bergen, Norway.

Berg, C.C., 2012. Moraceae. In: Davidse, G., Sousa, S.M., Knapp, S., F, C.C. (Eds.), Flora Mesoamericana. Universidad Nacional Autónoma de México, México.

Berg, C.C., Corner, E.J.H., 2005. Moraceae (*Ficus*). In: Nooteboom, H. (Ed.), Flora Malesiana Series I - Seed Plants. vol. 17. Nationaal Herbarium Nederland, Leiden, pp. 1–730 Part 2.

Berg, C.C., Wiebes, J.T., 1992. African fig trees and fig wasps. Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen. (Amsterdam).

Bruun-Lund, S., Clement, W.L., Kjellberg, F., Rønsted, N., 2016. First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). Mol. Phylogenet. Evol. 109, 93–104. http://dx.doi.org/10.1016/j.ympev.2016.12.031.

Burin, G., Kissling, W.D., Guimarães, P.R., Sekercioglu, Ç.H., Quental, T.B., 2016. Omnivory in birds is a macroevolutionary sink. Nat. Commun. 7, 11250. http://dx.doi.org/10.1038/ncomms11250.

Burrows, J., Burrows, S., 2003. Figs of southern & south-central africa. Umdaus press, South Africa.

Clement, W.L., Weiblen, G.D., 2009. Morphological evolution in the mulberry family (Moraceae). Syst. Bot. 34, 530–552. http://dx.doi.org/10.1600/036364409789271155.

Cook, J.M., Rasplus, J.Y., 2003. Mutualists with attitude: coevolving fig wasps and figs. Trends Ecol. Evol. 18, 241–248. http://dx.doi.org/10.1016/S0169-5347(03)00062-4.

Corner, E.J.H., 1961. In Botanical Society of Edinburgh, MacLeod, A.M., & Cobley, L. S., (1961). Contemporary Botanical Thought. Quadrangle Books, Chicago.

Corner, E.J.H., 1967. *Ficus* in the Solomon Islands and its bearing on the post-jurrasic history of Melanesia. Philos. Trans. R. Soc. Lond. B Biol. Sci. 253, 23–159.

Couvreur, T.L.P., Forest, F., Baker, W.J., 2011a. Origin and global diversification patterns of tropical rain forests: inferences from a complete genus-level phylogeny of palms. BMC Biol. 9, 44. http://dx.doi.org/10.1186/1741-7007-9-44.

Couvreur, T.L.P., Pirie, M.D., Chatrou, L.W., Saunders, R.M.K., Su, Y.C.F., Richardson, J.E., Erkens, R.H.J., 2011b. Early evolutionary history of the flowering plant family Annonaceae: steady diversification and boreotropical geodispersal. J. Biogeogr. 38, 664–680. http://dx.doi.org/10.1111/j.1365-2699.2010.02434.x.

Cruaud, A., Rønsted, N., Chantarasuwan, B., Chou, L.S., Clement, W.L., Couloux, A., Cousins, B., Genson, G., Harrison, R.D., Hanson, P.E., Hossaert-McKey, M., Jabbour-Zahab, R., Jousselin, E., Kerdelhué, C., Kjellberg, F., Lopez-Vaamonde, C., Peebles, J., Peng, Y.-Q., Pereira, R.A.S., Schramm, T., Ubaidillah, R., van Noort, S., Weiblen, G.D., Yang, D.-R., Yodpinyanee, A., Libeskind-Hadas, R., Cook, J.M., Rasplus, J.-Y., Savolainen, V., 2012. An extreme case of plant-insect codiversification: figs and fig-pollinating wasps. Syst. Biol. 61, 1029–1047. http://dx.doi.org/10.1093/sysbio/sys068.

Donoghue, M.J., Sanderson, M.J., 2015. Confluence, synnovation, and depauperons in plant diversification. New Phytol. 207, 260–274.

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. http://dx.doi.org/10.1093/sysbio/syr126.

Edwards, E.J., Osborne, C.P., Strömberg, C.A., Smith, S.A.; C4 Grasses Consortium, Bond, W.J., Christin, P.A., Cousins, A.B., Duvall, M.R., Fox, D.L., Freckleton, R.P., Ghannoum, O., Hartwell, J., Huang, Y., Janis, C.M., Keeley, J.E., Kellogg, E.A., Knapp, A.K., Leakey, A.D., Nelson, D.M., Saarela, J.M., Sage, R.F., Sala, O.E., Salamin, N., Still, C.J., Tipple, B., 2010. The origins of C4 grasslands: integrating evolutionary and ecosystem science. Science 328, 587–590. doi:10.1126/science.1177216.

Erkens, R.H.J., Chatrou, L.W., Maas, J.W., van der Niet, T., Savolainen, V., 2007. A rapid diversification of rainforest trees (*Guatteria*; Annonaceae) following dispersal from Central into South America. Mol. Phylogenet. Evol. 44, 399–411. http://dx.doi.org/10.1016/j.ympev.2007.02.017.

Fitzjohn, R.G., 2012. Diversitree: comparative phylogenetic analyses of diversification in R. Methods Ecol. Evol. 3, 1084–1092. http://dx.doi.org/10.1111/j.2041-210X.2012.00234.x.

Frenzke, L., Goetghebeur, P., Neinhuis, C., Samain, M.-S., Wanke, S., 2016. Evolution of epiphytism and fruit traits act unevenly on the diversification of the species-rich genus *Peperomia* (Piperaceae). Front. Plant Sci. 7, 1145. http://dx.doi.org/10.3389/fpls.2016.01145.

Gentry, A.H., Dodson, C.H., 1987. Diversity and biogeography of neotropical vascular epiphytes. Ann. Mo. Bot. Gard. 74, 205–233.

Harrison, R.D., 2005. Figs and the diversity of tropical rainforests. Bioscience 55, 1053. http://dx.doi.org/10.1641/0006-3568(2005)055. [1053:FATDOT]2.0.CO;2.

Harrison, R.D., Hamid, A. a, Kenta, T., LaFrankie, J., Lee, H.-S., Nagamasu, H., Nakashizuka, T., Palmiotto, P., 2003. The diversity of hemi-epiphytic figs in a Bornean lowland rain forest. Biol. J. Linn. Soc. 78, 439–456.

Harrison, R.D., Rasplus, J.-Y., 2006. Dispersal of fig pollinators in Asian tropical rain forests. J. Trop. Ecol. 22, 631. http://dx.doi.org/10.1017/S0266467406003488.

Harrison, R.D., Shanahan, M., 2005. Seventy-seven ways to be a fig: overview of a diverse plant assemblage. Pollinat. Ecol. Rain 174, 111–127. http://dx.doi.org/10.1007/0-387-27161-9_10.

Heard, S.B., Hauser, D.L., 1995. Key evolutionary innovations and their ecological mechanisms. Hist. Biol. 10, 151–173. http://dx.doi.org/10.1080/10292389509380518.

Heilbuth, J.C., 2000. Lower species richness in dioecious clades. Am. Nat. 156, 221–241. http://dx.doi.org/10.1086/303389.

Hodges, S.A., Arnold, M.L., Hodges, S.A., Arnold, M.L., 1995. Spurring plant diversification: are floral nectar spurs a key innovation . Proc. Biol. Sci. 262, 343–348.

Jaramillo, C., Ochoa, D., Contreras, L., Pagani, M., Carvajal-Ortiz, H., Pratt, L.M., Krishnan, S., Cardona, A., Romero, M., Quiroz, L., Rodriguez, G., Rueda, M.J., Parra, F., Morón, S., Green, W., Schouten, S., Bermudez, H., Navarrete, R., Parra, F., Alvarán, M., Osorno, J., Crowley, J.L., Valencia, V., Vervoort, J., 2010. Effects of rapid global warming at the paleocene-eocene boundary on neotropical vegetation. Science 330, 882–886. http://dx.doi.org/10.1126/science.1229223.

Jousselin, E., Hossaert-McKey, M., Herre, E.A., Kjellberg, F., 2003a. Why do fig wasps actively pollinate monoecious figs? Oecologia 134, 381–387. http://dx.doi.org/10.1007/s00442-002-1116-0.

Jousselin, E., Rasplus, J., Kjellberg, F., 2003b. Convergence and coevolution in a mutualism: evidence from a molecular phylogeny of *Ficus*. Evolution 57, 1255–1269. http://dx.doi.org/10.2307/3448849.

Käfer, J., de Boer, H.J., Mousset, S., Kool, a, Dufay, M., Marais, G. a B., 2014. Dioecy is associated with higher diversification rates in flowering plants. J. Evol. Biol. 27, 1478–1490. http://dx.doi.org/10.1111/jeb.12385.

Kay, K.M., Reeves, P.A., Olmstead, R.G., Schemske, D.W., 2005. Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. Am. J. Bot. 92, 1899–1910. http://dx.doi.org/10.3732/ajb.92.11.1899.

Kjellberg, F., Jousselin, E., Bronstein, J.L., Patel, A., Yokoyama, J., Rasplus, J., 2001. Pollination mode in fig wasps: the predictive power of correlated traits. Proc. R. Soc. B Biol. Sci. 268, 1113–1121. http://dx.doi.org/10.1098/rspb.2001.1633.

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

Laenen, B., Machac, A., Gradstein, S.R., Shaw, B., Pati, J., Désamore, A., Goffinet, B., Cox, C.J., Shaw, A.J., Vanderpoorten, A., 2016. Increased diversification rates follow shifts to bisexuality in liverworts. New Phytol. 210, 1121–1129.

Lovette, I.J., Bermingham, E., Ricklefs, R.E., 2002. Clade-specific morphological

diversification and adaptive radiation in Hawaiian songbirds. Proc. R. Soc. B Biol. Sci. 269, 37–42. http://dx.doi.org/10.1098/rspb.2001.1789.
Machado, C.A., Robbins, N., Gilbert, M.T.P., Herre, E.A., 2005. Critical review of host specificity and its coevolutionary implications in the fig fig-wasp mutualism. Proc. Natl. Acad. Sci. U. S. A. 102, 6558–6565.
Maddison, W.P., Midford, P.E., Otto, S.P., 2007. Estimating a binary character's effect on speciation and extinction. Syst. Biol. 56, 701–710. http://dx.doi.org/10.1080/10635150701607033.
Magallon, S., Sanderson, M.J., 2001. Absolute diversification rates in angiosperm clades. Evolution 55, 1762–1780. http://dx.doi.org/10.1111/j.0014-3820.2001.tb00826.x.
Mayhew, P.J., 2007. Why are there so many insect species? Perspectives from fossils and phylogenies. Biol. Rev. 82, 425–454. http://dx.doi.org/10.1111/j.1469-185X.2007.00018.x.
Moore, B.R., Höhna, S., May, M.R., Rannala, B., Huelsenbeck, J.P., 2016. Critically evaluating the theory and performance of bayesian analysis of macroevolutionary mixtures. Proc. Natl. Acad. Sci. 113, 9569–9574. http://dx.doi.org/10.1073/pnas.1518659113.
Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20, 289–290. http://dx.doi.org/10.1093/bioinformatics/btg412.
Pennington, R.T., Hughes, M., Moonlight, P.W., 2015. The origins of tropical rainforest hyperdiversity. Trends Plant Sci. 20, 693–695. http://dx.doi.org/10.1016/j.tplants.2015.10.005.
Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: convergence diagnosis and output analysis for MCMC. R. News 6, 7–11. http://dx.doi.org/10.1159/000323281.
Rabosky, D.L., Donnellan, S.C., Grundler, M., Lovette, I.J., 2014. Analysis and visualization of complex macroevolutionary dynamics: an example from Australian Scincid lizards. Syst. Biol. 63, 610–627. http://dx.doi.org/10.1093/sysbio/syu025.
Richardson, E.J., Pennington, R.T., Pennington, T.D., Hollingsworth, P.M., 2001. Rapid diversifcation of a species-rich genus of Neotropical rain forest trees. Science 293, 2242–2245. http://dx.doi.org/10.1126/science.1061421.
Rønsted, N., Weiblen, G.D., Clement, W.L., Zerega, N.J.C., Savolainen, V., 2008a. Reconstructing the phylogeny of figs (*Ficus*, Moraceae) to reveal the history of the fig pollination mutualism. Symbiosis 45, 1–12.
Rønsted, N., Weiblen, G.D., Cook, J.M., Salamin, N., Machado, C.A., Savolainen, V., 2005. 60 million years of co-divergence in the fig-wasp symbiosis. Proc. R. Soc. B Biol. Sci. 272, 2593–2599. http://dx.doi.org/10.1098/rspb.2005.3249.
Rønsted, N., Weiblen, G.D., Savolainen, V., Cook, J.M., 2008b. Phylogeny, biogeography, and ecology of *Ficus* section *Malvanthera* (Moraceae). Mol. Phylogenet. Evol. 48, 12–22. http://dx.doi.org/10.1016/j.ympev.2008.04.005.
Sanderson, M.J., Donoghue, M.J., 1994. Shifts in diversification rates with the origin of angiosperms. Science 265, 1590–1593.
Särkinen, T.E., Newman, M.F., Maas, P.J.M., Maas, H., Poulsen, A.D., Harris, D.J., Richardson, J.E., Clark, A., Hollingsworth, M., Toby Pennington, R., 2007. Recent oceanic long-distance dispersal and divergence in the amphi-Atlantic rain forest genus *Renealmia* L.f. (Zingiberaceae). Mol. Phylogenet. Evol. 44, 968–980. http://dx.doi.org/10.1016/j.ympev.2007.06.007.
Shanahan, M., Harrison, R.D., Yamuna, R., Boen, W., Thornton, I.W.B., 2001. Colonization of an island volcano, Long Island, Papua New Guinea, and an emergent island, Motmot, in its caldera lake. V. Colonization by figs (*Ficus* spp.), their dispersers and pollinators. J. Biogeogr. http://dx.doi.org/10.1046/j.1365-2699.2001.00642.x.
Silvestro, D., Zizka, G., Schulte, K., 2014. Disentangling the effects of key innovations on the diversification of Bromelioideae (Bromeliaceae). Evolution 68, 163–175. http://dx.doi.org/10.1111/evo.12236.
Smith, S.A., Beaulieu, J.M., Stamatakis, A., Donoghue, M.J., 2011. Understanding angiosperm diversification using small and large phylogenetic trees. Am. J. Bot. 98, 404–414. http://dx.doi.org/10.3732/ajb.1000481.
Stebbins, G.L., 1974. Flowering Plants: Evolution above the Species Level. Harvard University Press, Cambridge, MA, USA.
Vamosi, J.C., Vamosi, S.M., 2011. Factors influencing diversification in angiosperms: at the crossroads of intrinsic and extrinsic traits. Am. J. Bot. 98, 460–471. http://dx.doi.org/10.3732/ajb.1000311.
Wallace, A.R., 1878. Tropical Nature, and Other Essays. Macmillian, London, UK.
Wang, W., Ortiz, R., del, C., Jacques, F.M.B., Xiang, X.G., Li, H.L., Lin, L., Li, R.Q., Liu, Y., Soltis, P.S., Soltis, D.E., Chen, Z.D., 2012. Menispermaceae and the diversification of tropical rainforests near the Cretaceous-Paleogene boundary. New Phytol. 195, 470–478. http://dx.doi.org/10.1111/j.1469-8137.2012.04158.x.
Weiblen, G., 2004. Correlated evolution in fig pollination. Syst. Biol. 53, 128–139. http://dx.doi.org/10.1080/10635150490265012.
Wilf, P., Cúnego, N.R., Johnson, K.R., Hicks, J.F., Wing, S.L., Obradovich, J.D., 2003. High plant diversity in eocene south America: evidence from patagonia. Science 300, 122–125. http://dx.doi.org/10.1126/science.1080475.
Wilf, P., Johnson, K.R., 2004. Land plant extinction at the end of the Cretaceous: a quantitative analysis of the North Dakota megafloral record. Paleobiology 30, 347–368. http://dx.doi.org/10.1666/0094-8373(2004)030. < 0347:LPEATE > 2.0.CO;2.
Xing, Y., Onstein, R.E., Carter, R.J., Stadler, T., Linder, H.P., 2014. Fossils and a large molecular phylogeny show that the evolution of species richness, generic diversity, and turnover rates are disconnected. Evolution 68, 2821–2832. http://dx.doi.org/10.1111/evo.12489.
Zerega, N.J.C., Clement, W.L., Datwyler, S.L., Weiblen, G.D., 2005. Biogeography and divergence times in the mulberry family (Moraceae). Mol. Phylogenet. Evol. 37, 402–416. http://dx.doi.org/10.1016/j.ympev.2005.07.004.
Flora of China. In: Zhengyi, W., Raven, P.H., Deyuan, H. (Eds.), Ulmaceae through Basellaceae. vol. 5 Science press, Beijing, China.

Chapter IV

First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae)

Buddha attained enlightenment while meditating under a tree belonging to *Ficus religiosa* L.

Molecular Phylogenetics and Evolution 109 (2017) 93–104

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae)

Sam Bruun-Lund [a,*], Wendy L. Clement [b], Finn Kjellberg [c], Nina Rønsted [a]

[a] Natural History Museum of Denmark, Faculty of Science, University of Copenhagen, Sølvgade 83S, DK-1309 Copenhagen, Denmark
[b] Department of Biology, The College of New Jersey, 2000 Pennington Road, Ewing, NJ 08628, USA
[c] CEFE UMR 5175, CNRS, Université de Montpellier, Université Paul-Valéry Montpellier, EPHE, F-34293 Montpellier Cédex 5, France

ARTICLE INFO

Article history:
Received 15 June 2016
Revised 22 December 2016
Accepted 23 December 2016
Available online 29 December 2016

Keywords:
Ficus
High throughput sequencing
Cyto-nuclear discordance
Phylogeny
Plastome

ABSTRACT

Standard Sanger chloroplast markers provide limited information to resolve species level relationships within plants, in particular within large genera. Figs (*Ficus* L., Moraceae) compose one of the 50 largest genera of angiosperms with ~750 species occurring in the tropics and subtropics worldwide. Figs, in addition to being a keystone food resource in rainforests, are well-known for the mutualistic interactions with their pollinating wasps. It is regarded as a model system for understanding co-evolution dating back more than 75 million years. However, despite significant taxon sampling, combinations of low copy nuclear, nuclear ribosomal and chloroplast regions have not been able to confidently resolve relationships among major groups of figs.

Using a high throughput sequencing approach we attempted to resolve the major lineages of *Ficus* based on plastome data. In this study, we show that the use of a *de novo* assembled plastome from within the genus provides less ambiguity and higher coverage across the 59 *Ficus* and 6 outgroup plastome assemblies compared to using the nearest available reference plastome outside the genus resulting in improved resolution and higher support of the phylogenetic relationships within *Ficus* inferred from plastome data.

Chloroplast genome data confidently resolved relationships among major groups of figs and largely support current understanding based on nuclear sequence data including passively pollinated Neotropical section *Pharmacosycea* as sister lineage to all other *Ficus*. However, conflicts between the new plastome topology and previous nuclear studies are observed for both individual species as well as relationships among some sections at deeper levels. Conflicts could be caused by lack of resolution in the nuclear data or may indicate potential cyto-nuclear discordance as previously observed in an African lineage of *Ficus*.

1. Introduction

The figs (*Ficus* L., Moraceae) comprise ~750 species and account for more than 80% of the diversity of Moraceae, the family to which it belongs. Figs are distributed globally in tropical and subtropical regions and exhibit tremendous diversity in growth habits including freestanding trees, shrubs, root climbers, stranglers, epiphytes and lithophytes (Berg and Corner, 2005). Figs produce fruit throughout the year and often have large population sizes making them keystone species for many tropical communities (Berg and Corner, 2005; Harrison et al., 2012). However, figs might be best known for their obligate pollination mutualism with fig wasps (Agaonidae, Chalcidoidea, Hymenoptera), an intricate interaction of two lineages that has been retained for at least 75 million years (Cruaud et al., 2012b) and at the center of many coevolutionary and cospeciation studies (Bronstein and McKey, 1989; Cook and Rasplus, 2003; Cruaud et al., 2012a, 2012b; Herre, 1989; Herre et al., 2008; Herre and West, 1997; Jackson et al., 2008; Jousselin et al., 2008; Lopez-Vaamonde et al., 2001; Machado et al., 2005; Marussich and Machado, 2007; McLeish and van Noort, 2012; Rønsted et al., 2005; Silvieus et al., 2007; Weiblen, 2004, 2001; Weiblen and Bush, 2002).

Over the past 20 years, there have been significant advancements in our understanding of the phylogeny of figs (Cruaud et al., 2012b; Herre et al., 1996; Jousselin et al., 2003; Rønsted et al., 2008, 2005; Weiblen, 2000; Xu et al., 2011). This work guides the re-evaluation of the current classification of figs primarily based on morphology (Berg and Corner, 2005) which divides *Ficus* into six subgenera (*Ficus*, *Pharmacosycea*, *Sycidium*, *Sycomorus*,

* Corresponding author.
E-mail address: sam.bruunlund@snm.ku.dk (S. Bruun-Lund).

http://dx.doi.org/10.1016/j.ympev.2016.12.031
1055-7903/

Synoecia, and *Urostigma*). The most recent and comprehensive phylogenetic work on figs samples ~200 species from five rapidly evolving coding and non-coding nuclear markers (nrITS, nrETS, *G3pdH*, *waxy*, *ncpGS*) and does not support the monophyly of three of the six subgenera – *Ficus*, *Pharmacosycea*, and *Urostigma* (Cruaud et al., 2012b). The Cruaud et al. screened for paralogous sequences for *waxy*, which was know to cause trouble from previous studies in Rosales (Cruaud et al., 2012b).

Yet, despite a consistent increase in genetic data and species sampling through the past decade of phylogenetic work on figs, the current status of the fig phylogeny does not provide sufficient resolution or clade support to unambiguously resolve relationships deep within *Ficus*, thus leaving the backbone of the phylogeny uncertain (Fig. 2).

Lack of a well-supported phylogenetic hypothesis for *Ficus* has hindered progress on key research questions regarding diversification, biogeography and species interactions (Cruaud et al., 2012b; Herre et al., 2008). Further, few evolutionary studies of figs have employed plastid markers in phylogenetic reconstruction and none has adequately sampled the phylogenetic diversity of figs. Chloroplast markers are commonly employed in plant phylogenetic studies but often lack variability to confidently resolve relationships at infrageneric levels, especially within large genera (Rønsted et al., 2010; Roy et al., 2010). To date, two studies used chloroplast markers on a small or taxonomically narrow sample of figs. Herre et al. (1996) presented one of the first published molecular phylogenies of figs that sampling 15 species and was based on *trnL-F* and *rbcL* chloroplast markers. More recently, Renoult et al. (2009) sampled five non-coding chloroplast markers for 38 species of African figs in subgenus *Urostigma* section *Galoglychia* and suggested that these markers had potential for resolving deep nodes in the fig phylogeny. Significant conflicts were recovered when the plastid phylogeny (Renoult et al., 2009) was compared to the fig phylogeny based on nuclear ribosomal ITS and ETS (Rønsted et al., 2007). It was hypothesized that the conflicts recovered were due to ancient hybridization (Renoult et al., 2009). Neither study represented the phylogenetic diversity of *Ficus* as we currently understand it and sampled only a maximum of 3604 bp of plastid DNA (Herre et al., 1996; Renoult et al., 2009; Rønsted et al., 2010) As such, the evolutionary history of *Ficus* reconstructed from plastid markers has yet to be fully explored.

A growing number of studies are applying the use of phylogenomic methods whereby a large number of base pairs from the plastome can be sequenced and used in analysis (Clement et al., 2014; Givnish et al., 2010; Henriquez et al., 2014; Jansen et al., 2007; Moore et al., 2007; Parks et al., 2009). With well over 1000 plastomes of land plants (retrieved Dec 2016 from https://www.ncbi.nlm.nih.gov/genome/organelle/), it is by far the most sequenced genome of the plant cell. The number is rapidly growing and chloroplast phylogenomic data hold the potential to resolve deeper nodes in phylogenetic reconstructions (e.g., Henriquez et al., 2014; Jansen et al., 2007; Moore et al., 2007; Steele et al., 2012) and also at lower taxonomic levels between species (Clement et al., 2014; Parks et al., 2009). Furthermore, plastid DNA is not affected by processes, such as paralogous copies or recombination, as nuclear markers can be (Moore et al., 2007), which may compromise biological inferences from nuclear data. However, plastid DNA is maternally inherited and phylogenetic relationships based on the chloroplast genome may obscure processes such as hybridization (Crowl et al., 2014). Comparing nuclear and plastid gene histories can offer valuable insights to evolutionary processes (e.g., ancient hybridization) that may have had significant impacts on the lineage of study (Maddison, 1997; Raamsdonk et al., 1997). It is therefore crucial to obtain data from multiple genome resources, making it possible to uncover and investigate such mentioned processes as those above.

Here, we examine the potential of plastome data to resolve the evolutionary history and relationships within *Ficus*. Figs are among the 50 most speciose plant genera, thus providing the opportunity to evaluate the utility and impact of a plastid phylogenomic approach (Frodin, 2004; Straub et al., 2012). Additionally, this work represents the first significant plastid phylogeny of *Ficus* sampling all major lineages of the genus, and as such, we will use the phylogeny to make informed comparisons with phylogenetic hypotheses based on nuclear data (Cruaud et al., 2012b; Xu et al., 2011).

To improve our understanding of the evolutionary history of figs from plastid data, our goals are to (I) generate the first plastid genome phylogenomic hypothesis for all major clades of *Ficus*, (II) and describe the contribution of plastome phylogenomic data to the understanding of the evolutionary history and relationships within *Ficus*.

2. Material and methods

2.1. Sampling strategy

We sampled 59 species (Table 1) representing all major lineages of *Ficus* as reported in Cruaud et al. (2012b) with several representatives for each clade. We included three additional species, *F. pseudopalma* (subgenus *Sycomorus* section *Adenosperma*), *F. semicordata* (subgenus *Sycomorus* section *Hemicardia*), and *F. tsiangii* (subgenus *Sycidium* section *Sycidium*) with uncertain taxonomic affiliation based on molecular phylogeny. In recent phylogenetic analysis (Cruaud et al., 2012b; Weiblen, 2000; Xu et al., 2011) these species did not fall within the taxonomic group suggested by morphology (Berg and Corner, 2005) nor were they confidently resolved as being closely related to any other clade of figs. The outgroup was comprised of six species of Castilleae, the sister group of figs (Table 1; Clement and Weiblen, 2009).

2.2. DNA extraction and sequencing

DNA was isolated from fresh, silica dried, or herbarium material (Table 1). Whole genomic DNA was either extracted using the cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1987) or Qiagen DNeasy Minikit (Qiagen, Germantown, Maryland, USA) following manufacturer's protocol with two modifications to some samples to increase yield: (1) 50 μL Proteinase K was added and incubated for 1 h at 45 °C following the second step in the manufacturers protocol, and (2) the final elution step was done twice in 50 μL AE buffer and left on the column membrane for 10 min each iteration. See Supplementary Material S11 for discussion about extraction methods. DNA was quantified on a Qubit 2.0 fluorometer (HS; High Sensitivity) following manufacturers instructions (Life Technologies).

In preparation for building libraries for high-throughout sequencing (NGS), DNA was sheared to ca. 400 bp using a Bioruptor® (Diagenode). Illumina-compatible 100 bp paired-end libraries from DNA extracts were prepared with a NEBNext Library building kits for second-generation sequencing (New England Biolabs, Ipswich, MA, catalogue nr. E6070L) following manufacturer's protocol or by in house protocols (e.g., Petersen et al., 2015) by the Danish National High-Throughput DNA Sequencing Center. Libraries were hereafter amplified and multiplexed with custom in-house indices. Libraries were amplified with either AmpliTaq Gold (Life Technologies, Carlsbad, CA) or Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA). Libraries were characterized on a Qubit 2.0 fluorometer (Life Technologies) and quality checked on an Agilent 2100 Bioanalyzer (Santa Clara, CA). The libraries were sequences on several lanes on an Illumina HiSeq2000 platform at the Danish National High-Throughput

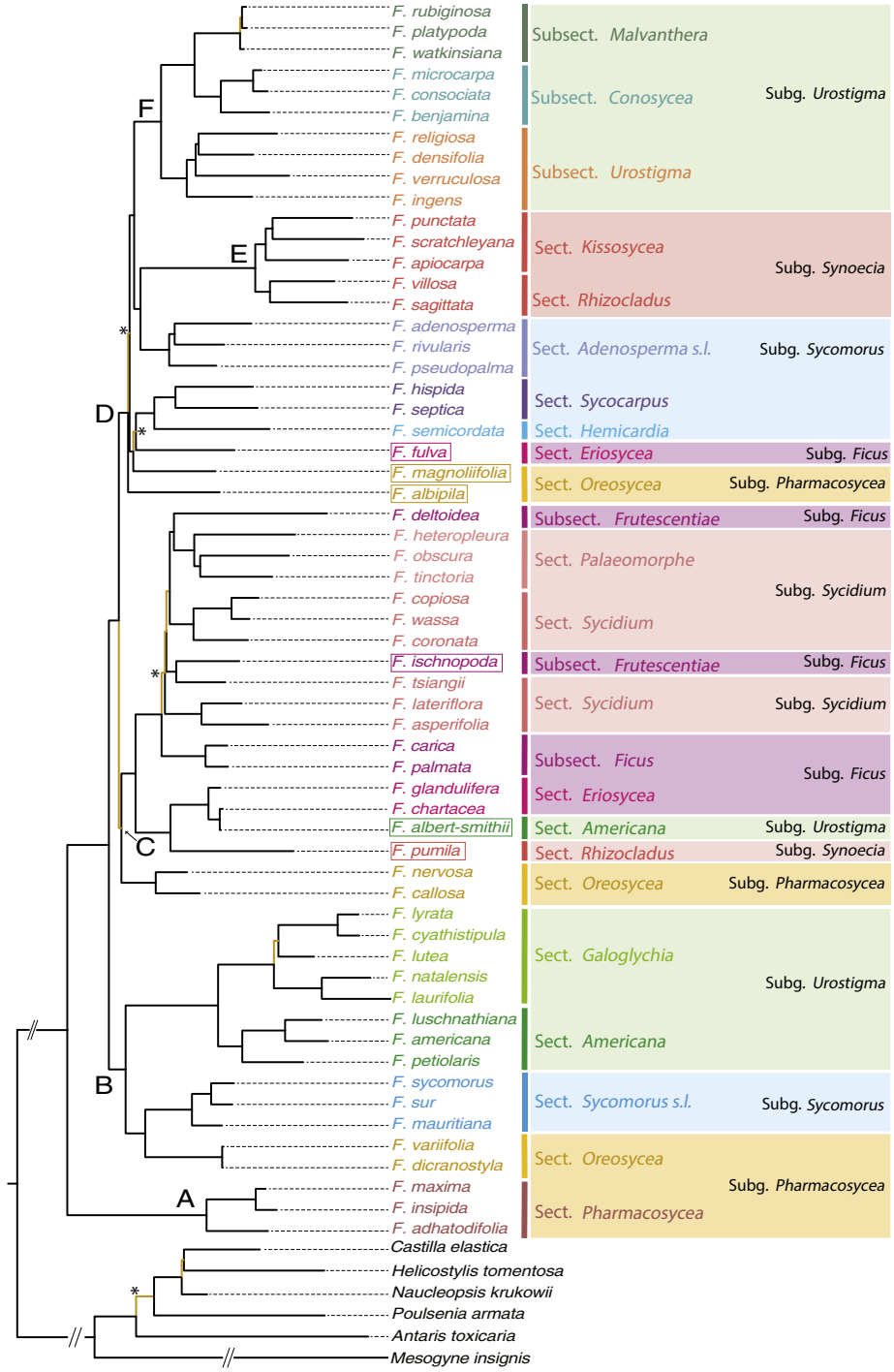


Fig. 1. Bayesian majority-rule consensus tree reconstructed from whole plastome sequences (assembled to *F. religiosa*) for 59 species of *Ficus*. As the majority of relationships are well-supported (ML bootstrap [BS] > 70% and posterior probabilities [PP] > 0.95) only those branches with weak support are noted as follows: branches colored orange indicate PP < 0.95 and BS < 70%; branches marked by an asterisk (*) indicate BS < 70% but PP > 0.95. Classification following Berg and Corner (2005) indicated on the right of taxon names. Displaced taxa boxed and discussed in text.

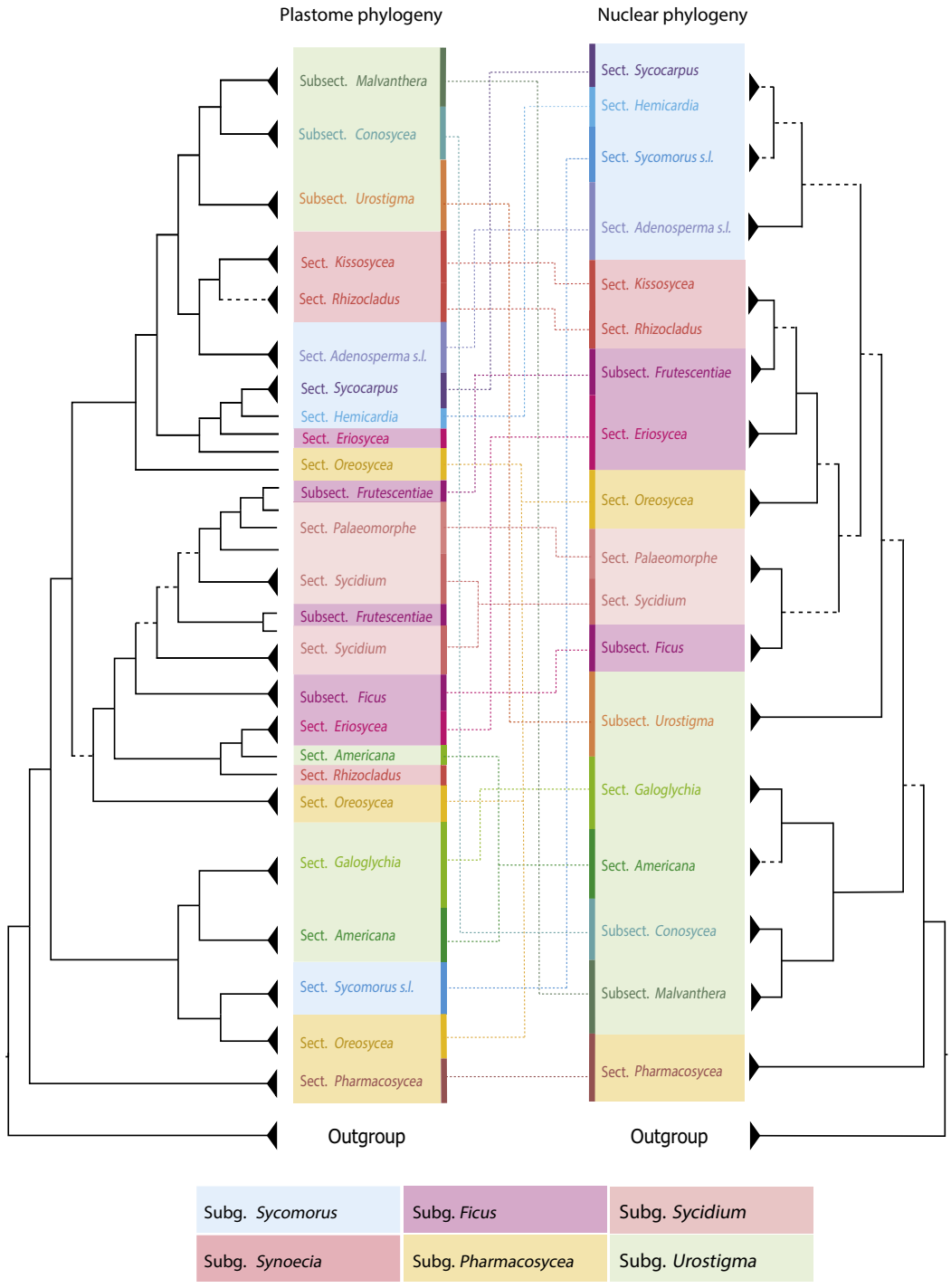


Fig. 2. Comparison of topologies recovered from analysis of the plastome data (left; this study) and nuclear markers (right; modified from Cruaud et al., 2012b). For each phylogeny, terminal branches have been collapsed for simplicity to show relationships among sections and subsections following Berg and Corner (2005). Coloring scheme for boxes indicating subgenera and colored section names follows Cruaud et al. (2012b), which is the most comprehensive phylogenetic study to date based on an average of 60% of 5.5 KB from five nuclear markers. Dotted lines connect the phylogenetic position of the sections and subsections on the plastome (left) and nuclear phylogenies (right). Bootstrap support values below 80% indicated with punctured lines.

Table 1
Voucher information for each specimen included in the present study. Letters in parenthesis indicate herbarium where voucher is housed: A: Harvard University; AAU: Aarhus University; BG: University of Bergen; C: University of Copenhagen; HITBC: Xishuangbanna Tropical Botanical Garden, Academia Sinica; K: Royal Botanical Gardens, Kew; MIN: University of Minnesota; MPU: University of Montpellier; PUH: University of the Philippines; REU: Université de la Réunion; THNHM: Thailand Natural History Museum.

Taxa	Voucher specimen	Origin
Ficus L.		
Sect. *Adenosperma* Corner.		
F. adenosperma Miq.	Takeuchi 14359 (K)	Papua New Guinea
F. rivularis Merr.	Harrison 607 (PUH)	Philippines
F. pseudopalma Blanco	Harrison 610 (PUH)	Philippines
Sect. *Americana* Miq.		
F. albert-smithii Standl.	Rønsted 105 (K)	Cult. BG89-519 (BG)
F. luschnathiana (Miq.) Miq.	Rønsted 78 (C)	Cult. S1947-0243 (C)
F. americana Aubl.	Rønsted 154 (BG)	Cult. BG1994-0678 (BG)
F. petiolaris Kunth.	Rønsted 94 (C)	Cult. P1994-5326 (C)
Subsect. *Conosycea* (Miq.) Corner.		
F. consociata Blume	Rønsted 164 (K/AAU)	Thailand
F. microcarpa L.f.	Rønsted 77 (C)	Cult. P1979-5041 (C)
F. benjamina L.	Rønsted 81 (C)	Cult. P1870-5193 (C)
Subsect. *Eriosycea* Miq.		
F. chartacea King	Chantarasuwan NR246 (THNHM)	Thailand
F. fulva Reinw. Ex Blume	Kjellberg FK1998-79 (MPU)	Brunei
F. glandulifera Wall.	Jousselin 18 (MPU)	Brunei
Subsect. *Ficus*		
F. carica L.	Rønsted 96 (C)	Cult. E1859-0001 (C)
F. palmata Forssk.	C10011606 (C)	Ethiopia
Subsect. *Frutescentiae* Sata.		
F. deltoidea Jack.	Rønsted 73 (C)	Cult. E1859-0015 (C)
F. ischnopoda Miq.	Rønsted 278 (HITBC)	Cult. (HITBC)
Sect. *Galoglychia* (Gasp.) Endl.		
F. cyathistipula Warb.	Rønsted 80 (C)	Cult. E1859-0023 (C)
F. natalensis Hochst.	Rønsted 75 (C)	Cult. P1958-5167 (C)
F. ovata Vahl. (*syn of laurifolia Lam.*)	Rønsted 205 (K)	Cameroon
F. lutea Vahl.	Rønsted 87 (C)	Cult. P1928-5257 (C)
F. lyrata Warb.	Rønsted 85 (C)	Cult. P1909-5032 (C)
Sect. *Hemicardia*		
F. semicordata Buch.-Ham. ex Sm.	Chantarasuwan NR241 (THNHM)	Thailand
Sect. *Kissosycea* Miq.		
F. apiocarpa (Miq.) Miq.	Chantarasuwan C14 (THNHM)	Thailand
F. punctata Thunb.	Kjellberg FK-1997-23 (MPU)	Brunei
F. scratchleyana King.	R. Johns PP10456 (K)	Irian Jaya, Indonesia
Sect. *Malvanthera* Corner.		
F. rubiginosa Desf. Ex Vent.	Rønsted 89 (C)	Cult. E1859-0014 (C)
F. watkinsiana F.M. Bailey	Rønsted 83 (C)	Cult. S1959-1912 (C)
F. platypoda (Miq.) A.Cunn. ex Miq.	Rønsted 142 (K)	Cult. BG-93-220 (BG)
Sect. *Oreosycea* (Miq.) Corner		
F. callosa Willd.	Rønsted 109 (K)	Cult. BG-99-561 (BG)
F. magnoliifolia Blume	Harrison 619 (PUH)	Philippines
F. nervosa Roth	Chantarasuwan NR225 (THNHM)	Thailand
***Albipillae* group**		
F. albipila (Miq.) King	Chantarasuwan NR250 (THNHM)	Thailand
F. dicranostyla Mildbr.	Rønsted 152 (K)	Cult. BG-88-241 (BG)
F. variifolia Warb.	Rønsted 131 (K)	Cult. BG-88-240 (BG)
Sect. *Palaeomorphe* King.		
F. obscura Bl.	Harrison 602 (PUH)	Philippines
F. heteropleura Bl.	Rønsted 70 (C)	Cult. P1972-5164 (C)
F. tinctoria Forst.f.	Rønsted 99 (K)	Cult. BG-89-551 (BG)
Sect. *Pharmacosycea* (Miq.) Benth. & Hook.		
F. adhatodifolia Schott	Rønsted 148 (K)	Cult. BG-2001-0623 (BG)
F. insipida Willd.	Rønsted 119 (K)	Cult. BG-89-523 (BG)
F. maxima Mill.	Rønsted 156 (K)	Cult. (BG)
Sect. *Rhizocladus* Endl.		
F. pumila L.	Rønsted 91 (C)	Cult. P1973-5350 (C)
F. sagittata Vahl.	Rønsted 266 (HITBC)	Cult. (HITBC)
F. villosa Bl.	Chase 19851 (K)	Cult. 1984-2930 (K)
Sect. *Sycidium* Miq.		
F. copiosa Diels	Weiblen D8 (A)	Papua New Guinea
F. wassa Roxb.	Utteridge 101 (K)	Irian Jaya, Indonesia
F. asperifolia Miq.	Rønsted 217 (K)	Cameroon
F. coronata Spin.	Rønsted 71 (C)	Cult. S1953-2381 (C)
F. lateriflora Vahl.	Fournel & Micheneau IF130 (REU)	Reunion
F. tsiangii Merr. Ex Corner	Rønsted 298 (HITBC)	Cult. (HITBC)

(*continued on next page*)

Table 1 (*continued*)

Taxa	Voucher specimen	Origin
Sect. *Sycocarpus* Miq.		
F. hispida L. F.	Rønsted 88 (C)	Cult. E1859-0002 (C)
F. septica Burm.f.	Harrison 618 (PUH)	Philippines
Sect. *Sycomorus* (Gasp.) Miq.		
F. mauritiana Lamarck	Micheneau 20/2/04 (REU)	Reunion
F. sur Forssk.	Rønsted 76 (C)	Cult. S1992-0213 (C)
F. sycomorus L.	Rønsted 72 (C)	Cult. P1965-5118 (C)
Sect. *Urostigma* (Gasp.) Endel.		
F. densifolia Miq.	Fournel & Micheneau IF117 (REU)	Reunion
F. ingens (Miq.) Miq.	Rønsted 106 (K)	Cult. BG-91-839 (BG)
F. religiosa L.	Rønsted 86 (C)	Cult. P1951-5144 (C)
F. verruculosa Warb.	Rønsted 115 (K)	Cult. BG-90-1239 (BG)
Outgroup: Tribe Castillae		
Antaris toxicaria Lesch.	C10011604 (C)	Ethiopia
Castilla elastica Sess.	Bruun-Lund 3 (C)	Cult. S1948-2585 (C)
Helicostylis tomentosa (R. & E.) Rusby	C10011602 (C)	Ethiopia
Mesogyne insignis Engl.	C10011601 (C)	Tanzania
Naucleopsis krukowii (Standl.) Berg	C10011603 (C)	Ecuador
Poulsenia armata (Miq.) Standl.	C10011557 (C)	Costa Rica

DNA Sequencing Center. The first lane consisted of 16 samples, one of which was *F. religiosa*. The second lane consisted of 16 samples (15 *Ficus* for this study). The third lane consisted of 20 samples (19 *Ficus*). The last round consisted of 37 (35 *Ficus*). A total of 70 samples were prepared.

Reads were filtered to remove adaptors and low quality reads using AdaptorRemoval (Lindgreen, 2012) with default settings and a minimum read length set to 30 bp. The quality of the raw data was assessed using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

2.3. De novo assembly of Ficus religiosa plastid genome

We performed additional filtering of our data to improve the efficiency and quality of the *de novo* assembly. Raw data were compared against a custom database consisting of 12 chloroplast genomes (Supplementary Material, S2) across Rosales. Included in the custom database was *Morus indica* L. (Moraceae), which is the closest related species to *Ficus* with a published plastome (DQ22651; Ravi et al., 2006). Filtering was conducted in Geneious R7 (Biomatters Limited, New Zealand) using Geneious' read mapper and all mappings were done using default settings and allowing up to 100 iterations. We compared *de novo* assemblies using filtered reads from *F. religiosa* and Velvet, SOAPdenovo2 and MIRA. *De novo* assembly using Velvet v1.2.10 and VelvetOptimizer (Zerbino and Birney, 2008) was conducted using the implemented version in Geneious with k-mer choice based on n50 and Lbp. Using SOAPdenovo v.2.01 (Luo et al., 2012), we implemented the default settings and k-mer = 63. GapCloser by SOAP (Luo et al., 2012) was used to close gaps and conduct the first round of scaffolding. Lastly, MIRA v.4.0 was implemented in Geneious R7 (Biomatters Limited, New Zealand) with quality level set to 'accurate'. MIRA resulted in many short and overlapping contigs. To improve the assembly, we used the Geneious *de novo* assembler to merge contigs that overlapped with the 'custom sensitivity' option set to 98% alike. With this approach, we obtained much longer contigs (with higher N50) that covered more of the plastome and avoided extensive gap closing.

Contigs and scaffolds from all three *de novo* assemblers they were visually inspected in Geneious and mapped compared with to the annotated plastome of *M. indica* to ensure proper orientation. Further gap closing and scaffolding were hereafter conducted using an iterative approach. First, read mapping using Geneious' read mapper and Bowtie2 (Langmead and Salzberg, 2012) implemented in Geneious were used to repeatedly elongate either side of the gap until the gap was closed. Then, additional mapping using the original quality trimmed but unfiltered pool of reads (including nuclear and mitochondrial data) were conducted to allow for detection of any potential nuclear or mitochondrial insertions in the chloroplast genome of *F. religiosa*, which would not be detected using a database solely including plastome reads. Furthermore, it was assured that reads did not have the wrong orientation, to avoid mis-assembly.

Borders between the single copy regions and inverted repeat (IR) regions were verified using Sanger sequencing (primers and protocols listed in Supplementary Material, S3–S5). The resulting assemblies for *F. religiosa* from Velvet, SOAPdenovo2, and MIRA were aligned, and the differences were evaluated bioinformatically or by Sanger sequencing (Tables S3 and S5). Based on the results a consensus plastome sequence was generated. Quality-filtered reads were mapped back to the resulting assembly to verify (e.g., Eserman et al., 2014).

Through this comparative approach, all base pair differences between the three assemblies and the seven gaps closed with readmapping were verified using Sanger sequencing. Primers for closing seven gaps were designed using Primer3 (Untergasser et al., 2012) as implemented in Geneious (Supplementary Material S5). Annotation of the final *F. religiosa* plastome was done using the webserver DOGMA (Wyman et al., 2004), custom BLAST searches, and comparison our initial annotations with the published annotation of *M. indica* through the *transfer annotations* option to check for consistency. Open reading frames were checked in Geneious and used to aid the final annotation.

2.4. Reference based assembly of Ficus plastid genomes

Phylogenetic distance from a reference genome can significantly affect plastome assembly (Straub et al., 2012) which can pose problems for non-model groups where reference genomes are not available. To evaluate the effect the choice of reference has on assembly, high-throughput sequencing data from 64 additional plastome samples were mapped to both *M. indica* and *F. religiosa* plastomes. Reference based assembly was achieved using Bowtie2 (Langmead and Salzberg, 2012) with default settings. A consensus sequence was produced in Geneious with a threshold set to strict (50%) – if coverage was fewer 10 reads a "?" was called suggesting missing data at those positions in the consensus sequence. The annotated chloroplast genome of *Ficus religiosa* is available from GenBank with accession number KY416513.

2.5. Phylogenetic analyses

Alignment of 59 *Ficus* and six outgroup plastomes was achieved using the MAFFT v7.017 plugin (Katoh and Standley, 2013) in Geneious R7 with default settings and manually reviewed. To avoid artificially increasing the phylogenetic signal from the inverted repeated (IR) region in chloroplast genomes, one IR region was removed prior to analysis. The aligned matrix is available from www.DataDryad.org; http://dx.doi.org/10.5061/dryad.dk34k.

The most appropriate model of evolution (GTR + G) was found using JmodelTest2 (Darriba et al., 2012) according to the AIC criterion as recommended (Posada and Buckley, 2004). We considered the plastid genome as a single inherited unit, and using PartitionFinder V.1.1.1 (Lanfear et al., 2012), we confirmed the same model for both the small and large single copy regions and the included inverted repeat region.

Maximum Likelihood (ML) and Bayesian inference (BI) have been shown to be less susceptible to problems in phylogenetic reconstruction such as long branch attraction (Bergsten, 2005; Moore et al., 2007) and is the preferred choice for high throughput sequencing data (Bergsten, 2005; Eserman et al., 2014; Henriquez et al., 2014; Moore et al., 2007; Parks et al., 2009; Petersen et al., 2015). ML inference was conducted using RAxML v8.2.4 (Stamatakis, 2014) implemented on the CIPRES portal (Miller et al., 2010) using the rapid bootstrapping and search for best-scoring ML tree setting with 1000 Bootstrap replicates.

Similarly, BI was conducted with Mr. Bayes v3.2.6 (Huelsenbeck and Ronquist, 2001) on the CIPRES portal (Miller et al., 2010) with MCMC using two independent runs and four chains, sampling every 2000 generations for up to 20 million generations. Convergence of the run was assessed using Tracer v.1.6 (http://beast.bio.ed.ac.uk/tracer). The first 25% of samples were discarded as burn-in. The Bayesian majority-rule consensus tree is available from www.DataDryad.org; http://dx.doi.org/10.5061/dryad.dk34k.

Many phylogenetic studies using high-throughput sequencing data will extract markers or gene regions prior to phylogenetic analysis (Clement et al., 2014; Moore et al., 2007), but we reconstructed phylogeny from the entire genome. To assess whether noise in the data would generate an artificial signal, and thus, result in an erroneous topology, we extracted three datasets consisting of 11, 20 and 72 gene regions of coding and non-coding DNA. The markers were selected due to being widely used in the plant systematics community (e.g., Saslis-Lagoudakis et al., 2015), previously used in *Ficus* (Herre et al., 1996; Rønsted et al., 2010) or being highly variable (Dong et al., 2012; Shaw et al., 2005). To extract selected markers and gene regions, annotations were transferred from *F. religiosa* to the remaining 64 plastomes. Missing data in plastome sequences were not taken into account when extracting regions. Models of evolution were inferred using JmodelTest2 software (Darriba et al., 2012) and applied to the subsequent analysis of the concatenated data sets. ML was conducted using the RaxML (v. 7.2.8) plugin (Stamatakis, 2006) in Geneious with the GTR + G model of evolution and 200 bootstrap replicates. All trees were visualized and annotated in either Geneious R7 or using FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

To verify that our results are not a product of contamination or other lab errors we returned to our raw high-throughput data and assembled ITS, ETS, *G3pdh*, *waxy* and *ncpGS*, which have been sampled in previous phylogenetic work on *Ficus* (Cruaud et al., 2012b). ITS and ETS was obtained from most of the sequence data, whereas the low copy nuclear regions *G3pdh*, *waxy* and *ncpGS* could only be obtained from a subset of the samples in accordance with previous findings based on Sanger Sequencing (Cruaud et al., 2012b).

These regions were subjected to nBLAST searches in GenBank to confirm section membership. Additionally, we reconstructed a phylogeny of the ITS data using a RAxML analysis with 500 bootstrap replicates. The topology of this phylogeny (data not shown) was congruent with previously published nuclear phylogenies and conflicted with our plastome phylogeny as described in detail above.

3. Results

3.1. Ficus religiosa de novo assembly

A total of 12,960,076 paired end reads were produced on the Illumina HiSeq 2000 platform from *F. religiosa*. Through *de novo* assembly, a complete chloroplast genome with a total of 57× coverage was assembled with a length of 160,627 bp comprising a large single copy region, a small single copy region and two inverted repeats (88,815 bp, 20,106 bp, 25,853 bp respectively). A map of the annotated plastome is included in the Supplementary Material S6 using the OrganellarGenomeDRAW tool (Lohse et al., 2013).

3.2. Effect of choice of reference genome on alignment and phylogenetic analyses

For each of the 65 *Ficus* and outgroup species sequenced, between 217,322 and 94,253,802 reads were generated (Supplementary Material S7). Seven samples failed during library preparation and/or sequencing and was therefore not included. When using *F. religiosa* as a reference, 8× - 2192× coverage of the plastome was recovered (only 5 samples were less than 30× coverage; Supplementary Material S7).

The choice of reference (*M. indica* vs. *F. religiosa*) had a notable impact on coverage and base calls. Reference based assembly using *F. religiosa* resulted in a 154,826 bp alignment after exclusion of one IR. Reference based assembly using *M. indica* resulted in a 178,916 bp alignment again after excluding one IR (phylogenetic trees can be found in Supplementary Material S8). A mean increase of 37× coverage was observed when using *F. religiosa* as reference over *M. indica* (Supplementary Material S7). Additionally, using *F. religiosa* as reference reduced the number of ambiguities (e.g., Y for a C or T) in the consensus sequence and increased coverage across the plastid genome, thus, resulting in a more complete and accurate plastome consensus sequence for the downstream phylogenetic inference. For instance, Bayesian analysis of the *M. indica* based alignment consistently resulted in problems with converging after as many as 20 million generations, whereas analysis of the *F. religiosa* based alignment converged after only ~50,000 generations.

3.3. Impact of noise on phylogenetic inference

Phylogenomic studies often perform tree reconstruction using an extracted set of genes from the sequenced genome. Here, we subjected the entire plastid genome (excluding one copy of the IR) to phylogenetic analysis and compared it to phylogenies reconstructed from concatenated data sets of 11, 20 and 75 gene regions. The concatenated datasets resulted in several minor unsupported differences in relationships within clades among closely related species (see Supplementary Material S9–S11). The majority of clades were recovered as in the full plastome sequence analysis, but internal branches in the tree had significantly lower bootstrap support values (if any support at all) as might be expected from a less comprehensive data set.

3.4. Comprehensive plastid phylogeny of Ficus

ML and BI analyses of the plastid data resulted in no significant differences in topology or clade support (Fig. 1, BI topology shown).

The resulting topology is strongly supported and recovered the monophyly *of Ficus* and many infrageneric groups based on morphology (Fig. 1). We recovered strong support for section *Pharmacosycea* as sister to all other figs (Fig. 1, clade A). Within the sister-clade to section *Pharmacosycea* was a clade comprising subgenus *Urostigma* sections *Galoglychia* and *Americana*, subgenus *Sycomorus* section *Sycomorus*, and a few species of subgenus *Pharmacosycea* section *Oreosycea* (Fig. 1, clade B). The remainder of fig lineages sampled fell among two major clades (Fig 2, clades C and D). Clade C (Fig. 1) included all of subgenera *Sycidium* and *Ficus*, as well as a few species from subgenera *Synoecia*, *Urostigma* and *Pharmacosycea* (Fig. 1 clade C). Clade D (Fig. 1) included subgenus *Synoecia* sections *Kissosycea* and *Rhizocladus* (Fig. 1, Clade E), subgenus *Urostigma* subsections *Conosycea*, *Malvanthera*, and *Urostigma* (Fig. 1, Clade F), subgenus *Sycomorus* sections *Adenosperma* and *Sycocarpus*, as well as a few species from subgenera *Ficus* and *Pharmacosycea*.

Of the six subgenera defined by morphology, we only recovered monophyly of subgenus *Synoecia* (to the exclusion of *F. pumila*, Fig. 1, Clade E). The following sections were paraphyletic due to the exclusion of a single species: subgenus *Ficus* section *Eriosycea*, which excluded *F. fulva*; subgenus *Synoecia* subsection *Rhizocladus*, which excluded *F. pumila*; and subgenus *Urostigma* section *Americana*, which excluded *F. albert-smithii* (Fig. 1). Subgenus *Sycidium* section *Sycidium* was paraphyletic to subgenus *Sycidium* section *Palaeomorphe* and members of subgenus *Ficus* subsection *Frutescentiae* (Clade C, Fig. 1). Subgenus *Pharmacosycea* section *Oreosycea* was polyphyletic and fell out among clades B, C, and D on the phylogeny (Fig. 1).

Of the three included species that have been considered difficult to place based on morphology (Berg and Corner, 2005), *F. tsiangii* was recovered together with members of section *Sycidium* as would be expected from previous studies (Xu et al., 2011), *F. pseudopalma* was recovered in section *Adenosperma* as has been found for its close relative, *F. dammaropsis* (Rønsted et al., 2008), whereas *F. semicordata*, which was also recovered as sister to section *Sycocarpus*, have previously been recovered as member of section *Sycomorus*, which is in better line with expectations from morphology (Berg and Corner, 2005).

4. Discussion

4.1. Impact of reference genome on reference-based assembly

Our assembly and annotation of the plastome of *F. religiosa* (Supplementary Material S6) resulted in a sequence 160,627 bp long with structural arrangements of the small single copy region (SSR) and two inverted repeats (IR) similar to the *Morus indica* plastome (158,484 bp; Ravi et al., 2006). Additionally, our annotated plastome is similar to a recently published *Ficus* plastome of *F. racemosa* (159,473 bp; Mao and Bi, 2015) in structure and also gene content although our plastome is 1154 bp longer. Comparing our plastome annotations to *F. racemosa* and *M. indica* it appears that a few genes vary in length. This is probably due to the different ways of annotation applied and it is not possible to know, which is the more correct version without gene expression experiments. As the majority of annotations from all three plastomes are very similar it indicates a consistent assembly of the *F. religiosa* plastome. Overall, length differences among the annotated plastomes was likely mainly due to insertions and deletions but the methods of data collection and assembly methods were not reported for *F. racemosa* making it difficult to compare our findings. As described in the methods, we verified with Sanger sequencing each junction and area that was not well supported by the original *de novo* assembly.

There are already around 800 annotated plastid genomes of flowering plants that have been published in GenBank. This a vast resource that can be used in future studies that was not available few years ago. However, 800 plastid genomes is only a fraction of the nearly 400,000 extant species of angiosperms. Phylogenomic studies therefore have to sometimes rely on distantly related reference genomes or do a *de novo* assembly. Straub et al. (2012) found that evolutionary divergence of the target taxon from the reference genome can significantly affect plastome assembly (Straub et al., 2012). Our results demonstrated that the choice of reference had a significant effect not only on the quality of the reference-based assemblies but also of downstream phylogenetic analyses. By using *F. religiosa* as opposed to *M. indica* as a reference, our assemblies had fewer ambiguities and higher coverage across all 65 genomes assembled. Additionally, by using a reference genome within the same genus, comparatively more reads could be mapped from each high-throughput data set to less conserved areas that are often difficult to assemble (e.g., introns, intergeneric spacers, IR boundaries). Further, alignment of these data resulted in more informative characters and improved resolution in the phylogenetic analysis and therefore greater clade support for relationships within *Ficus*. It is thus highly recommended for future studies using plastome phylogenomic data to *de novo* assemble a plastome from within a study group to serve as a reference for subsequent reference based assemblies prior to phylogenetic analysis even if there is a closely related published plastome available in GenBank.

4.2. Congruence and conflicts between plastome and nuclear data

Many of the infrageneric relationships corresponded to relationships recovered in previously published phylogenies (Cruaud et al., 2012b; Rønsted et al., 2008). Despite a small sample size representing less than 10% of all figs (59 of ~750 fig species), our sampling reflects the phylogenetic diversity of *Ficus* (Cruaud et al., 2012b) and resulted in a topology that recovered support for well-known sections of *Ficus* as well as infrageneric relationships. Several clades including *Americana + Galoglychia* (within clade B, Fig. 1), *Sycidium-Frutescentiae-Paleomorphe* (clade C, Fig. 1), *Kissosycea + Rhizocladus* (clade E, Fig. 1) and *Conosycea-Malvanthera* (within clade F, Fig. 1) concur with prior work on *Ficus* phylogeny (Fig. 2, Cruaud et al., 2012b). Also, South American section *Pharmacosycea* (clade A, Fig. 1) was strongly supported as sister to all other *Ficus* (ML bootstrap = 100%, posterior probability = 1.00) – a result that has been reported but never confidently supported (Fig. 2; Cruaud et al., 2012b).

However, we observed conflicts between the new plastome topology and previous nuclear studies are observed for both individual species as well as relationships among some sections at deeper levels. Conflicts could be caused by lack of resolution (soft incongruence) in the nuclear data or may indicate potential cyto-nuclear discordance (hard incongruence) as previously observed in an African lineage of *Ficus* (Renoult et al., 2009).

As discussed in the introduction, the currently known nuclear loci based phylogeny does not provide sufficient resolution or clade support to unambiguously resolve relationships within *Ficus*, thus leaving the backbone of the phylogeny uncertain.

At shallow level in the plastid genome analysis, six species were displaced into other sections (Fig. 1) resulting in the paraphyly of subgenera *Urostigma* (including *Conosycea-Malvanthera* and *Americana-Galoglychia-Urostigma* clades) and *Sycomorus* (with section *Sycomorus* separated from the *Adenosperma-Hemicardia-Sycocarpus* clade, although this was not a hard incongruence as the subgenus was resolved but not strongly supported in the nuclear topology by Cruaud et al., 2012b). This stands in conflict with our current understanding of fig evolutionary relationships based on the nuclear sequence data, in which each of the aforementioned sec-

tions and subsections are well-supported monophyletic groups (Cruaud et al., 2012b). At deeper level, we recovered a polyphyletic subgenus *Sycomorus* as sections *Adenosperma*, *Sycocarpus*, *and Sycomorus*, which fell out in three places in the phylogeny (clades B and D, Fig. 1). Also, section *Urostigma*, which has been separated from the remainder of subgenus *Urostigma* (*Americana-Conosycea-Galoglychia-Malvanthera*) in all previous studies based on nuclear markers, is here sister to *Conosycea-Malvanthera* (clade F, Figs. 1 and 2). Lastly, the *Americana-Galoglychia* clade (clade B, Fig. 1) was more closely related to *Sycomorus* + *Oreosycea* as opposed to *Conosycea-Malvanthera-Urostigma* (Fig. 2).

4.3. Introgressive hybridization associated with pollinator shift could explain potential cyto-nuclear discordance

Conflicts between plastome and nuclear data could be caused by lack of support, but could also be the result of hard incongruence. Discordance between nuclear and plastid markers has been reported within African *Ficus* section *Galoglychia* (Renoult et al., 2009) and is a rather common phenomenon in plant systematics (Rieseberg and Soltis, 1991), which is becoming more evident as high quantity phylogenomic data is being increasingly used (Crowl et al., 2014; Jeffroy et al., 2006; Philippe et al., 2005). Renoult et al. (2009) hypothesized that the cyto-nuclear discordance within *Galoglychia* was likely caused by ancient hybridization followed by introgression. Hybridization among closely related species or species complexes within *Ficus* has been reported (Kusumi et al., 2012; Machado et al., 2005; Parrish et al., 2003; Renoult et al., 2009; Wei et al., 2014). Although pollinator specificity in a nursery mutualism is generally thought to prevent hybridization (Moe and Weiblen, 2012; Parrish et al., 2003; Ramirez, 1994), it has been hypothesized that figs may regularly exchange genes during divergence because of nonspecific pollinators or host shifts, which could also explain patterns of conflicts between plant and pollinator phylogenies (Cruaud et al., 2012b; Machado et al., 2005; Renoult et al., 2009). Consequently, cyto-nuclear discordance within *Ficus* mirroring incongruence between plant and pollinator phylogenies would support the host shift hypothesis.

Three of the displaced species *F. albert-smithii*, *F. ischnopoda* and *F. pumila* are all material from living collections in botanical gardens, while *F. albipila*, *F. fulva* and *F. magnoliifolia* were collected in the field. For the taxa collected in gardens, their position recovered by the plastome phylogeny could be due to recent hybridization with other garden material followed by either incomplete lineage sorting or introgression, which has been suggested as explanation for discordance in *Galanthus* and *Armeria* (Fuertes Aguilar et al., 1999; Rønsted et al., 2013). *Ficus albert-smithii* was obtained from the living collection of C.C. Berg in Bergen, Norway (BG). Berg was an authority of *Ficus* and obtained most of his materials from seeds of wild origin. However, he unfortunately passed away a few years ago and no further information about his materials can be obtained (Van Welzen et al., 2013). His materials are most likely introduced as seed directly from their natural origin within his life-time. *Ficus pumila* growing in the botanic garden in Copenhagen (C) was received as a plant in 1973 rather than grown from seed but the origin is not recorded. *Ficus ischnopoda* was obtained from the Xishuangbanna Tropical Botanical Garden (HITBC, Yunnan, China) and is part of a large collection of primarily wild sourced *Ficus* occurring naturally in China.

However, considering the generation time of trees and the likely wild origins of the botanical garden samples, we consider the explanation of hybridization with other garden material unlikely. Alternatively, alien plastid capture could have occurred through horizontal gene transfer during natural grafting as proposed by Stegemann et al. (2012). For example, genetic mosaics, presumably formed by fusion of different individuals, have been reported for several species of strangler figs (Thomson et al., 1991). A possible mechanism could be fertilization of an aneuploid ovule by a non-reduced male gamete, which would instantly lead to cytoplasm exchange. Unreduced gametes are frequently involved in interspecific hybridization and have been suggested to represent a mechanism for evolutionary speciation (Mason and Pires, 2015).

Ficus pumila is considered a member of section *Rhizocladus* based on morphology and its root-climbing habit (Berg and Corner, 2005), but nuclear sequence data has suggested it is more closely related to subsection *Frutescentiae* (Cruaud et al., 2012b; Jousselin et al., 2003; Rønsted et al., 2008), which is also supported by its pollinator *Wiebesia pumilae* being close to the pollinator of *F. deltoidea* (Cruaud et al., 2012b). Both *F. deltoidea*, and *F. ischnopoda* are also members of subsection *Frutescentiae* according to nuclear sequence data (e.g., Cruaud et al., 2012b). Displacement of *F. albert-smithii* outside section *Americana* and *F. fulva* outside section *Eriosycea*, is not supported by any previous studies or morphological features.

The three other displaced samples, *F. albipila*, *F. fulva* and *F. magnoliifolia* are all of wild origin. Both *F. albipila* and *F. magnoliifolia* are members of the Paleotropic section *Oreosycea*, from which the *Albipilae* group has already been separated by nuclear sequence data (e.g., Rønsted et al. (2008), but the *Albipilae* group was not included in Cruaud et al. (2012b) and more comprehensive sampling is likely to recover additional evolutionary linages.

Berg and Corner's (2005) circumscription of subgenus *Urostigma* included about 280 species of banyans and hemi-epiphytic stranglers. Nuclear phylogenetic studies have resolved a clade including *Americana-Galoglychia* and *Conosycea-Malvanthera* as sister clades, but excluding section *Urostigma*, which has been unplaced. Our plastome phylogeny resolves section *Urostigma* together with the *Americana-Galoglychia* clade, supporting the morphological classification and the pollinator phylogeny (Cruaud et al., 2012b, Fig. 1). Furthermore, in the pollinator phylogeny, the pollinators of section *Malvanthera* (genus *Pleistodontes*) are separated from the clade with the pollinators of sections *Americana-Conosycea-Galoglychia* supporting the possibility that ancient introgressive hybridization associated with host-shifting may be a possible explanation for the observed discordance.

Failure to recover monophyly of subgenus *Sycomorus* is in conflict with current understanding based both on nuclear sequence data (Cruaud et al., 2012b; Harrison et al., 2012) and morphology (Berg and Corner, 2005) and pollination of the entire subgenus by the monophyletic wasp genus *Ceratosolen* (Cruaud et al., 2012b). However, whereas *Adenosperma* is sister to *Sycocarpus-Sycomorus* in the nuclear sequence based phylogeny, the pollinators of section *Sycomorus* are sister to the pollinators of *Adenosperma-Sycocarpus* again suggesting a more complex genetic history of this subgenus.

4.4. Impact of plastid phylogenomics on the evolutionary history of figs

This is the first study to confidently recover section *Pharmacosycea* as sister to all other figs (Fig. 1). Prior studies including an outgroup have all recovered this relationship (Herre et al., 1996; Rønsted et al., 2005, 2008; Xu et al., 2011; Cruaud et al., 2012b), but not with the significant bootstrap support (Fig. 2) and posterior probability obtained in this study using plastid genomes. The implication of section *Pharmacosycea* as sister to all other figs supports the hypothesis that *Ficus* originated in Southern Gondwana (Machado et al., 2001; Rønsted et al., 2005) rather than in Eurasia (Cruaud et al., 2012b; Zerega et al., 2005). Additionally, the support of placement of section *Pharmacosycea* has important implications for the co-evolutionary history of figs and their pollinating wasps. In the phylogeny of fig wasp pollinators, *Tetrapus*,

the genus of fig wasp pollinators of section *Pharmacosycea*, was not recovered as the first diverging lineage of wasps (Cruaud et al., 2012b). This conflict does not support an evolutionary history of strict co-speciation among these two lineages and points to more ancient instances of host switching events which has also been suggested by Machado et al. (2005).

5. Conclusions and future perspectives

A complete plastid genome of *F. religiosa* has been assembled *de novo*, annotated, and used to reference assemble additional samples from 64 species to resolve the early diversification with success. These data were used to reconstruct a phylogeny that offered a well-resolved and strongly supported topology that offered insight to relationships among major clades within *Ficus*. This phylogenetic framework represents the first comprehensive investigation of the evolution of figs using plastid data, and largely supported the monophyly of many sections of *Ficus* as well as evolutionary relationships previously recovered by nuclear data.. However, we also detected two types of cyto-nuclear discordance. First, we recovered discordance at deeper levels concerning the relationships of sections between major clades that could possibly be explained by ancient introgressive hybridization associated with pollinator shifts. Second, displacement of individual accessions at shallow level could tentatively be explained by recent unsorted garden hybridization, but we consider this unlikely due to the long generation time of these plants.

Since the evolutionary relationships among figs reconstructed with nuclear sequence data better correspond with morphology, we suspect the discordance is caused by the plastome data reflecting a divergent evolutionary history. Caution should be used when relying on and interpreting plastid phylogenomic data in the future particularly for purposes of classification. However, we are one step closer to resolving the early diversification and understanding the complex evolutionary history of *Ficus*.

To date, commonly used nuclear markers such as the nuclear ribosomal ITS and ETS have been utilized in *Ficus* together with a handful of low-copy nuclear markers. The availability of an expressed sequence tag library (EST) from *Ficus elastica* Roxb. ex Hornem. (Yao et al., 2013) and transcriptomic data from 1 KP (1000 plants) project opens up the possibility of phylogenomic studies of *Ficus* based on more comprehensive sampling of the nuclear genome. With those data in hand, we will be able to continue to compare and study the cyto-nuclear discordance recovered from this work.

Acknowledgements

We would like to thank Marie Lisandra Zepeda Mendoza, Argelia Cuenca Navarro, Aldo Carmona Baez and Tom Gilbert, for helpful discussions and input on bioinformatics pipelines and the employees at the Danish National High-throughput DNA Sequencing Center for sequencing the data for this study and advice on DNA requirements. The Botanic Garden and Herbarium C of the Natural History Museum of Denmark, Xishuangbanna Tropical Botanical Garden, Yunnan, China, Cornelis C. Berg, Bhanumas Chantarasuwan, Rhett Harrison, Emmanuelle Jousselin, Claire Micheneau and George Weiblen are thanked for providing plant material.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2016.12.031.

References

Berg, C.C., Corner, E.J.H., 2005. Flora Malesiana, Series I. Volume 17 part 2. In: Nooteboom, H. (Ed.), Flora Malesiana Series I - Seed Plants, Vol. 17 Part 2. National Herbarium Nederland, Leiden, pp. 1–730.

Bergsten, J., 2005. A review of long-branch attraction. Cladistics. http://dx.doi.org/10.1111/j.1096-0031.2005.00059.x.

Bronstein, J., McKey, D., 1989. The fig-pollinator mutualism: a model system for comparative biology. Experientia 45, 601.

Clement, W.L., Arakaki, M., Sweeney, P.W., Edwards, E.J., Donoghue, M.J., 2014. A chloroplast tree for Viburnum (Adoxaceae) and its implications for phylogenetic classification and character evolution. Am. J. Bot. 101, 1029–1049. http://dx.doi.org/10.3732/ajb.1400015.

Clement, W.L., Weiblen, G.D., 2009. Morphological evolution in the mulberry family (Moraceae). Syst. Bot. 34, 530–552. http://dx.doi.org/10.1600/036364409789271155.

Cook, J.M., Rasplus, J.Y., 2003. Mutualists with attitude: coevolving fig wasps and figs. Trends Ecol. Evol. 18, 241–248. http://dx.doi.org/10.1016/S0169-5347(03)00062-4.

Crowl, A.A., Mavrodiev, E., Mansion, G., Haberle, R., Pistarino, A., Kamari, G., Phitos, D., Borsch, T., Cellinese, N., 2014. Phylogeny of Campanuloideae (Campanulaceae) with emphasis on the utility of nuclear pentatricopeptide repeat (PPR) genes. PLoS ONE 9, e94199. http://dx.doi.org/10.1371/journal.pone.0094199.

Cruaud, A., Jabbour-Zahab, R., Genson, G., Ungricht, S., Rasplus, J.Y., 2012a. Testing the emergence of new caledonia: fig wasp mutualism as a case study and a review of evidence. PLoS ONE 7. http://dx.doi.org/10.1371/journal.pone.0030941.

Cruaud, A., Rønsted, N., Chantarasuwan, B., Chou, L.S., Clement, W.L., Couloux, A., Cousins, B., Genson, G., Harrison, R.D., Hanson, P.E., Hossaert-McKey, M., Jabbour-Zahab, R., Jousselin, E., Kerdelhué, C., Kjellberg, F., Lopez-Vaamonde, C., Peebles, J., Peng, Y.-Q., Pereira, R.A.S., Schramm, T., Ubaidillah, R., van Noort, S., Weiblen, G.D., Yang, D.-R., Yodpinyanee, A., Libeskind-Hadas, R., Cook, J.M., Rasplus, J.-Y., Savolainen, V., 2012b. An extreme case of plant-insect codiversification: figs and fig-pollinating wasps. Syst. Biol. 61, 1029–1047. http://dx.doi.org/10.1093/sysbio/sys068.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772. http://dx.doi.org/10.1038/nmeth.2109.

Dong, W., Liu, J., Yu, J., Wang, L., Zhou, S., 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PLoS ONE 7, e35071. http://dx.doi.org/10.1371/journal.pone.0035071.

Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19, 11–15. http://dx.doi.org/10.2307/4119796.

Eserman, L.A., Tiley, G.P., Jarret, R.L., Leebens-Mack, J.H., Miller, R.E., 2014. Phylogenetics and diversification of morning glories (tribe Ipomoeeae, Convolvulaceae) based on whole plastome sequences. Am. J. Bot. 101, 92–103. http://dx.doi.org/10.3732/ajb.1300207.

Frodin, D.G., 2004. History and concepts of big plant genera. Taxon 53, 753–776.

Fuertes Aguilar, J., Rosselló, J.A., Nieto Feliner, G., 1999. Nuclear ribosomal DNA (nrDNA) concerted evolution in natural and artificial hybrids of Armeria (Plumbaginaceae). Mol. Ecol. 8, 1341–1346. http://dx.doi.org/10.1046/j.1365-294X.1999.00690.x.

Givnish, T.J., Ames, M., McNeal, J.R., McKain, M.R., Steele, P.R., DePamphilis, C.W., Graham, S.W., Pires, J.C., Stevenson, D.W., Zomlefer, W.B., Briggs, B.G., Duvall, M. R., Moore, M.J., Heaney, J.M., Soltis, D.E., Soltis, P.S., Thiele, K., Leebens-Mack, J. H., 2010. Assembling the tree of the monocotyledons: plastome sequence phylogeny and evolution of poales. Ann. Missouri Bot. Gard. 97, 584–616. http://dx.doi.org/10.3417/2010023.

Harrison, R.D., Rønsted, N., Xu, L., Rasplus, J.-Y., Cruaud, A., 2012. Evolution of fruit traits in ficus subgenus Sycomorus (Moraceae): to what extent do frugivores determine seed dispersal mode? PLoS One 7, e38432. http://dx.doi.org/10.1371/journal.pone.0038432.

Henriquez, C.L., Arias, T., Pires, J.C., Croat, T.B., Schaal, B.A., 2014. Phylogenomics of the plant family Araceae. Mol. Phylogenet. Evol. 75, 91–102. http://dx.doi.org/10.1016/j.ympev.2014.02.017.

Herre, E.A., 1989. Coevolution of reproductive characteristics in 12 species of New World figs and their pollinator wasps. Experientia 45, 637–647. http://dx.doi.org/10.1007/BF01975680.

Herre, E.A., Jandér, K.C., Machado, C.A., 2008. Evolutionary ecology of figs and their associates: recent progress and outstanding puzzles. Annu. Rev. Ecol. Evol. Syst. 39, 439–458. http://dx.doi.org/10.1146/annurev.ecolsys.37.091305.110232.

Herre, E.A., Machado, C.A., Bermingham, E., Nason, J.D., Windsor, D.M., McCafferty, S. S., VanHouten, W., Bachmann, K., 1996. Molecular phylogenies of figs and their pollinator wasps. J. Biogeogr. 23, 521–530. http://dx.doi.org/10.1111/j.1365-2699.1996.tb00014.x.

Herre, E.A., West, S.A., 1997. Conflict of interest in a mutualism: documenting the elusive fig wasp-seed trade-off. Proc. Roy. Soc. B: Biol. Sci. 264, 1501–1507. http://dx.doi.org/10.1098/rspb.1997.0208.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Jackson, A.P., Machado, C.A., Robbins, N., Herre, E.A., 2008. Multi-locus phylogenetic analysis of neotropical figs does not support co-speciation with the pollinators:

the importance of systematic scale in fig/wasp cophylogenetic studies. Symbiosis 45.
Jansen, R.K., Cai, Z., Raubeson, L.A., Daniell, H., Depamphilis, C.W., Leebens-Mack, J., Müller, K.F., Guisinger-Bellian, M., Haberle, R.C., Hansen, A.K., Chumley, T.W., Lee, S.-B., Peery, R., McNeal, J.R., Kuehl, J.V, Boore, J.L., 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc. Natl. Acad. Sci. U. S. A. 104, 19369–19374. http://dx.doi.org/10.1073/pnas.0709121104.
Jeffroy, O., Brinkmann, H., Delsuc, F., Philippe, H., 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22, 225–231. http://dx.doi.org/10.1016/j.tig.2006.02.003.
Jousselin, E., Van Noort, S., Berry, V., Rasplus, J.Y., Rønsted, N., Erasmus, J.C., Greeff, J. M., 2008. One fig to bind them all: host conservatism in a fig wasp community unraveled by cospeciation analyses among pollinating and nonpollinating fig wasps. Evolution (NY) 62, 1777–1797. http://dx.doi.org/10.1111/j.1558-5646.2008.00406.x.
Jousselin, E.E., Rasplus, J.J.-Y., Kjellberg, F., 2003. Convergence and coevolution in a mutualism: evidence from a molecular phylogeny of Ficus. Evolution (NY) 57, 1255–1269. http://dx.doi.org/10.2307/3448849.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. http://dx.doi.org/10.1093/molbev/mst010.
Kusumi, J., Azuma, H., Tzeng, H.-Y.Y., Chou, L.-S.S., Peng, Y.-Q.Q., Nakamura, K., Su, Z.-H.H., 2012. Phylogenetic analyses suggest a hybrid origin of the figs (Moraceae: Ficus) that are endemic to the Ogasawara (Bonin) Islands, Japan. Mol. Phylogenet. Evol. 63, 168–179. http://dx.doi.org/10.1016/j.ympev.2012.01.004.
Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S., 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. 29, 1695–1701. http://dx.doi.org/10.1093/molbev/mss020.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods. http://dx.doi.org/10.1038/nmeth.1923.
Lindgreen, S., 2012. AdapterRemoval: easy cleaning of next generation sequencing reads. BMC Res. Notes 5, 337. http://dx.doi.org/10.1186/1756-0500-5-337.
Lohse, M., Drechsel, O., Kahlau, S., Bock, R., 2013. OrganellarGenomeDRAW – a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucl. Acids Res. http://dx.doi.org/10.1093/nar/gkt289.
Lopez-Vaamonde, C., Rasplus, J.Y., Weiblen, G.D., Cook, J.M., 2001. Molecular phylogenies of fig wasps: partial cocladogenesis of pollinators and parasites. Mol. Phylogenet. Evol. 21, 55–71. http://dx.doi.org/10.1006/mpev.2001.0993.
Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y.Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung, D.W., Yiu, S.-M., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J.J., Lam, T.-W., 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience 1, 18. http://dx.doi.org/10.1186/2047-217X-1-18.
Machado, C.A., Jousselin, E., Kjellberg, F., Compton, S.G., Herre, E.A., 2001. Phylogenetic relationships, historical biogeography and character evolution of fig-pollinating wasps. Proc. Biol. Sci. 268, 685–694. http://dx.doi.org/10.1098/rspb.2000.1418.
Machado, C.A., Robbins, N., Gilbert, M.T.P., Herre, E.A., 2005. Critical review of host specificity and its coevolutionary implications in the fig fig-wasp mutualism. Proc. Natl. Acad. Sci. U. S. A. 102.
Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536. http://dx.doi.org/10.1017/CBO9781107415324.004.
Mao, Q., Bi, G., 2015. Complete chloroplast genome of *Ficus racemosa* (Moraceae). Mitochondrial DNA 1736, 1–2. http://dx.doi.org/10.3109/19401736.2015.1106488.
Marussich, W.A., Machado, C.A., 2007. Host-specificity and coevolution among pollinating and nonpollinating New World fig wasps. Mol. Ecol. 16, 1925–1946. http://dx.doi.org/10.1111/j.1365-294X.2007.03278.x.
Mason, A.S., Pires, J.C., 2015. Unreduced gametes: meiotic mishap or evolutionary mechanism? Trends Genet. 31, 5–10. http://dx.doi.org/10.1016/j.tig.2014.09.011.
McLeish, M.J., van Noort, S., 2012. Codivergence and multiple host species use by fig wasp populations of the Ficus pollination mutualism. BMC Evol. Biol. 12, 1. http://dx.doi.org/10.1186/1471-2148-12-1.
Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gatew. Comput. Environ. Work. GCE 2010, 1–8. http://dx.doi.org/10.1109/GCE.2010.5676129.
Moe, A.M., Weiblen, G.D., 2012. Pollinator-mediated reproductive isolation among dioecious fig species (Ficus, Moraceae). Evolution (NY) 66, 3710–3721. http://dx.doi.org/10.1111/j.1558-5646.2012.01727.x.
Moore, M.J., Bell, C.D., Soltis, P.S., Soltis, D.E., 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. Proc. Natl. Acad. Sci. U. S. A. 104, 19363–19368. http://dx.doi.org/10.1073/pnas.0708072104.
Parks, M., Cronn, R., Liston, A., 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. BMC Biol. 7, 1–17. http://dx.doi.org/10.1186/1741-7007-7-84.
Parrish, T.L., Koelewijn, H.P., Van Dijk, P.J., Kruijt, M., 2003. Genetic evidence for natural hybridization between species of dioecious Ficus on island populations. Biotropica 35, 333–343. http://dx.doi.org/10.1111/j.1744-7429.2003.tb00587.x.
Petersen, G., Cuenca, A., Seberg, O., 2015. Plastome evolution in hemiparasitic mistletoes. Genome Biol. Evol. 7, 2520–2532. http://dx.doi.org/10.1093/gbe/evv165.
Philippe, H., Delsuc, F., Brinkmann, H., Lartillot, N., 2005. Phylogenomics. Annu. Rev. Ecol. Evol. Syst. 36, 541–562. http://dx.doi.org/10.1146/annurev.ecolsys.35.112202.130205.
Posada, D., Buckley, T.R., 2004. Model Selection and model averaging in phylogenetics: advantages of Akaike information criterion and bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808. http://dx.doi.org/10.1080/10635150490522304.
Raamsdonk, L.W.D. Van, Smiech, M.P., Sandbrink, J.M., 1997. Introgression explains incongruence between nuclear and chloroplast DNA-based phylogenies in Allium section Cepa. Bot. J. Linn. Soc. 123, 91–108. http://dx.doi.org/10.1111/j.1095-8339.1997.tb01406.x.
Ramirez, W., 1994. Hybridization of *Ficus religiosa* with *F. septica* and *F. aurea* (Moraceae). Rev. Biol. Trop. 42, 339–342.
Ravi, V., Khurana, J.P., Tyagi, A.K., Khurana, P., 2006. The chloroplast genome of mulberry: complete nucleotide sequence, gene organization and comparative analysis. Tree Genet. Genomes 3, 49–59. http://dx.doi.org/10.1007/s11295-006-0051-3.
Renoult, J.P., Kjellberg, F., Grout, C., Santoni, S., Khadari, B., 2009. Cyto-nuclear discordance in the phylogeny of Ficus section Galoglychia and host shifts in plant-pollinator associations. BMC Evol. Biol. 9, 248. http://dx.doi.org/10.1186/1471-2148-9-248.
Rieseberg, L.H., Soltis, D.E., 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evol. Trends Plants 5, 65–84. http://dx.doi.org/10.1007/s00606-006-0485-y.
Rønsted, N., Salvo, G., Savolainen, V., 2007. Biogeographical and phylogenetic origins of African fig species (Ficus section Galoglychia). Mol. Phylogenet. Evol. 43, 190–201. http://dx.doi.org/10.1016/j.ympev.2006.12.010.
Rønsted, N., Weiblen, G.D., Clement, W.L., Zerega, N.J.C., Savolainen, V., 2008. Reconstructing the phylogeny of figs (Ficus, Moraceae) to reveal the history of the fig pollination mutualism. Symbiosis 45, 1–12.
Rønsted, N., Weiblen, G.D., Cook, J.M., Salamin, N., Machado, C.A., Savolainen, V., 2005. 60 Million years of co-divergence in the fig-wasp symbiosis. Proc. Roy. Soc. B: Biol. Sci. 272, 2593–2599. http://dx.doi.org/10.1098/rspb.2005.3249.
Rønsted, N., Yektaei-Karin, E., Truk, K., Clarkson, J.J., Chase, M.W., 2010. Species-level phylogenetics of large genera: prospects of studying coevolution and polyploidy. In: Reconstructing the Tree of Life: Taxonomy and Systematics of Species Rich Taxa. Systematic Association Series, pp. 129–147.
Rønsted, N., Zubov, D., Bruun-Lund, S., Davis, A.P., 2013. Snowdrops falling slowly into place: an improved phylogeny for Galanthus (Amaryllidaceae). Mol. Phylogenet. Evol. 69, 205–217. http://dx.doi.org/10.1016/j.ympev.2013.05.019.
Roy, S., Tyagi, A., Shukla, V., Kumar, A., Singh, U.M., Chaudhary, L.B., Datt, B., Bag, S. K., Singh, P.K., Nair, N.K., Husain, T., Tuli, R., 2010. Universal plant DNA barcode loci may not work in complex groups: a case study with Indian Berberis species. PLoS ONE 5, e13674. http://dx.doi.org/10.1371/journal.pone.0013674.
Saslis-Lagoudakis, C.H., Bruun-Lund, S., Iwanycki, N.E., Seberg, O., Petersen, G., Jäger, A.K., Rønsted, N., 2015. Identification of common horsetail (*Equisetum arvense* L.; Equisetaceae) using thin layer chromatography versus DNA barcoding. Sci. Rep. 5, 11942. http://dx.doi.org/10.1038/srep11942.
Shaw, J., Lickey, E.B., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, K.C., Winder, C. T., Schilling, E.E., Small, R.L., 2005. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am. J. Bot. 92, 142–166. http://dx.doi.org/10.3732/ajb.92.1.142.
Silvieus, S.I., Clement, W.L., Weiblen, G.D., 2007. Cophylogeny of figs, pollinators, gallers, and parasitoids. In: Tilmon, K. (Ed.), Specialization, Speciation and Radiation: The Evolutionary Biology of Herbivorous Insects. University of California Press, California, pp. 225–239.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. http://dx.doi.org/10.1093/bioinformatics/btu033.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690. http://dx.doi.org/10.1093/bioinformatics/btl446.
Steele, P.R., Hertweck, K.L., Mayfield, D., McKain, M.R., Leebens-Mack, J., Pires, J.C., 2012. Quality and quantity of data recovered from massively parallel sequencing: examples in Asparagales and Poaceae. Am. J. Bot. 99, 330–348. http://dx.doi.org/10.3732/ajb.1100491.
Stegemann, S., Keuthe, M., Greiner, S., Bock, R., 2012. Horizontal transfer of chloroplast genomes between plant species. Proc. Natl. Acad. Sci. 109, 2434–2438. http://dx.doi.org/10.1073/pnas.1114076109.
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. Am. J. Bot. 99, 349–364. http://dx.doi.org/10.3732/ajb.1100335.
Thomson, J.D., Herre, E.A., Hamrick, J.L., Stone, J.L., 1991. Genetic mosaics in strangler fig trees: implications for tropical conservation. Science 254 (80), 3–10.
Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G., 2012. Primer3 – new capabilities and interfaces. Nucleic Acids Res. 40, e115. http://dx.doi.org/10.1093/nar/gks596.
van Welzen, P.C., Gadella, T., Maas, P., Daly, D.C., Mori, S.A., Jørgensen, P.M., Obermüller, F.A., Kjellberg, F., Rønsted, N., Chantarasuwan, B., Lut, C.W.J., 2013. In memoriam Cees Berg (2 July 1934–31 August 2012). Blumea 57, 199–206. http://dx.doi.org/10.3767/000651913X662362.
Wei, Z.D., Kobmoo, N., Cruaud, A., Kjellberg, F., 2014. Genetic structure and hybridization in the species group of *Ficus auriculata*: can closely related sympatric *Ficus* species retain their genetic identity while sharing pollinators? Mol. Ecol. 23, 3538–3550. http://dx.doi.org/10.1111/mec.12825.

Weiblen, G., 2004. Correlated evolution in fig pollination. Syst. Biol. 53, 128–139. http://dx.doi.org/10.1080/10635150490265012.
Weiblen, G.D., 2001. Phylogenetic relationships of fig wasps pollinating functionally dioecious Ficus based on mitochondrial DNA sequences and morphology. Syst. Biol. 50, 243–267. http://dx.doi.org/10.1093/sysbio/50.2.243.
Weiblen, G.D., 2000. Phylogenetic relationship of functionally dioecious Ficus (Moraceae) based on ribosomal DNA sequences and morphology. Am. J. Bot. 87, 1342–1357. http://dx.doi.org/10.2307/2656726.
Weiblen, G.D., Bush, G.L., 2002. Speciation in fig pollinators and parasites. Mol. Ecol. 11, 1573–1578. http://dx.doi.org/10.1046/j.1365-294X.2002.01529.x.
Wyman, S.K., Jansen, R.K., Boore, J.L., 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20, 3252–3255. http://dx.doi.org/10.1093/bioinformatics/bth352.
Xu, L., Harrison, R.D., Yang, P., Yang, D.R., 2011. New insight into the phylogenetic and biogeographic history of genus Ficus: vicariance played a relatively minor role compared with ecological opportunity and dispersal. J. Syst. Evol. 49, 546–557. http://dx.doi.org/10.1111/j.1759-6831.2011.00155.x.
Yao, X., Li, C., Dick, C.W., 2013. Exon-primed intron-crossing (EPIC) markers for evolutionary studies of Ficus and other taxa in the fig family (Moraceae). Appl. Plant Sci. 1, 1300037. http://dx.doi.org/10.3732/apps.1300037.
Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829. http://dx.doi.org/10.1101/gr.074492.107.
Zerega, N.J.C., Clement, W.L., Datwyler, S.L., Weiblen, G.D., 2005. Biogeography and divergence times in the mulberry family (Moraceae). Mol. Phylogenet. Evol. 37, 402–416. http://dx.doi.org/10.1016/j.ympev.2005.07.004.

Chapter V

From strangler figs to desert-dwellers – Biome transitioning in *Ficus* sect. *Malvanthera*

Ficus lilliputiana on sandstones in Australia

Data is still being analyzed and only a subset of the data is included, thus this chapter and its results should be considered preliminary

From rainforest strangler figs to desert-dwellers – Biome transitioning in *Ficus* sect. *Malvanthera*

Sam Bruun-Lund[1], Simon Segar[2], Jose Samaniego[3], Vanessa Bieker[4], Peter Jobson[5], Wendy L. Clement[6], Finn Kjellberg[7], Tom Gilbert[3], Michael D. Martin[4], James M. Cook[8], George D. Weiblen[9], Nina Rønsted[1]

[1]Natural History Museum of Denmark, University of Copenhagen, Denmark

[2]Dept of Crop and Environment Sciences, Harper Adams University, UK

[3] University of Copenhagen, Denmark

[4]NTNU University Museum, Norway

[5]Northern Territory Herbarium, Alice Springs, Department of Environment and Natural Resources, NT, Australia

[6]Department of Biology, The College of New Jersey, Ewing, NJ USA

[7]CNRS - UMR 5175 CEFE, Montpellier, France

[8]Western Sydney University, Australia

[9]Bell Museum and Department of Biology, University of Minnesota, Saint Paul, USA

Abstract

Biome shifts are rare and biome shifts from tropical rainforest to arid biomes are hypothesized to be even more rare. Climate change is often viewed as a major driver of plant distributions, but it is not known what factors determine the successful migration and diversification of some lineages. Australia has a long history of geographic isolation and great climatic change, which make its flora unique with many endemic species. Figs (*Ficus* L. ~800 spp.) exhibit an extremely high morphological and adaptive variation with species occurring in very different environments. An Australasian clade of figs, *Ficus* section *Malvanthera* (~23 spp.) is here used as a case study for testing biome shifts from rainforest to more arid biomes. This clade includes two clearly defined growth habits; (i) hemi-epiphytic stranglers in the rainforests of Eastern Australia, New Guinea, and some Australasian islands – and (ii) lithophytic shrubs/smaller trees occurring in the arid parts of Australia. In addition, section *Malvanthera* contains a number of species believed to be transitional as their habit and geographical range comprises both the arid and wet biomes and associated habits. To test for biome shifts and associated adaptive traits, we utilize a high throughput targeted sequence capture approach for reconstructing a phylogenetic hypothesis of the section. A set of probes that can be used across the entire genus of *Ficus* was developed and shown to confidently resolve relationships of major lineages. The probes were used on a preliminary taxon set of the Australasian clade of figs and also provide efficient resolution among closely related species. The final taxon set will be used

to test for biomes shifts and linked traits, between wet and dry regions using relevant diversification analysis.

Keywords (4-6)

Adaptations, biome shifts; dry habitats, niche shifts, transitional species, *Ficus*

1. INTRODUCTION

1.1. Biome shifts and out of the rainforest radiations

Why are some lineages in the tree of life extraordinary successful and species rich? Darwin described the rapid rise and early diversification within the angiosperms as ‘‘an abominable mystery’’ (Darwin & Seward 1903; Davies et al. 2004). Within the group, sister clades can differ in species richness over several orders of magnitude. Darwin attempted to identify a single causal explanation for the rapid diversification of angiosperms but described his own efforts as ‘‘wretchedly poor’’ and the idea of a single key innovation allowing a lineage to diversify rapidly is now abandoned but the mystery is unsolved (Smith et al., 2011).

Climate change is often viewed as a major driver of plant distributions, but it is not known what factors determine the successful

migration and diversification of some lineages (Davies et al., 2004; Donoghue and Edwards, 2014). According to the concept of niche conservation it may be "easier to move than to evolve" (Donoghue, 2008). Consequently, the plant lineages that fail to evolve *in situ* when the environment change will instead track the later environmental distributions, or, contract and develop disjunctions in their geographic ranges. However, many lineages have managed to adapt and transcend biome or niche boundaries as their environments have changed (Donoghue and Edwards, 2014). Novel traits can support the process and success of occupying new niches, but it is still not fully understood how plants sometimes rapidly are able to occupy new niches – so called niche- or biome shifts – and whether these traits are pre-adaptations or 'key innovations' (Donoghue and Edwards, 2014; Toon et al., 2015). In evolutionary biology one of the most difficult problems to explain is how adaptations appear at the optimal time for the next stage of evolution to take place. Pre-adaptations would make it possible for plants to diversify in newly accessible niches that might appear as a result of climate change (Huskins, 1930). Another possibility is the invention of key-adaptations that are suitable for a new opportunity (Heard and Hauser, 1995). However, the concept of key-innovations has more or less been

abandoned for more complex models where synergistic effects of several traits are involved (Donoghue and Sanderson, 2015; Koenen et al., 2015; Smith et al., 2011).

Modern tropical rainforests are one of the most important and species rich biomes today, and it is likewise suggested that tropical rainforests-like biomes were the first biomes on Earth, and that much of our plant-diversity stems from these ancient biomes (Webb et al., 2005). Thus, it is expected that species occurring in more arid areas are later radiations from ancestors living in the tropical rainforest habitat, and that these newly divergent lineages in arid biomes will be more closely related than earlier diverging linages from 'ancient rain forests' (Hughes et al., 2013).

Biome shifts are rare and biome shifts between tropical rainforest to arid biomes are hypothesized to be even more rare (Donoghue and Edwards, 2014) because the differences between wet and dry biomes seems to be a significant barrier for plant groups to cross (Pennington et al., 2009). However, it has for example been shown that grasses (Poaceae) are able to adapt and further speciate in arid biomes due to functional traits that were already present in their evolutionary linages linked to C4 photosynthesis (Heibl and Renner, 2012; Toon et al., 2015).

Plants that are well adapted to these new biomes have also been shown to accelerate their diversification rate due to new opportunities for speciation (Crayn et al., 2006). Plant colonization in arid habitats is often linked to the evolution of traits related to fire tolerance and exploitation of nutrient-poor soils, which may increase growth ability and diversity in these habitats (Donoghue and Edwards, 2014; Edwards et al., 2010). However, little is known about adaptations of larger shrubs and woody plants in dry areas. In for example Bignoniaceae, it has also been hypothesized that plants have radiated from drier (xeric) biomes into wet tropical forests, in a direction much less common (Crisp et al., 2009; Donoghue and Edwards, 2014).

Australian ecosystems are a unique model for investigating evolutionary patterns in plants and biome shifts (Crayn et al., 2006). The high diversity and high numbers of endemic species found in Australia is a product of a long history of geographic isolation and great climatic change during the Tertiary (from about 66 million years ago; Mya). Climatic fluctuations during the Miocene (23-5 Mya) resulted in aridification of the continent and retraction of forests towards the coast, while some groups of plants radiated into the new drier biome (Crayn et

al. 2006). For example, Crayn et al. (2006) documented radiation of dry-adapted shrubs from rainforest tree progenitors in the Elaeocarpaceae.

1.2. Ficus *L. section* Malvanthera *Corner as a model system*

Another potential case of radiation out of the humid rainforest into arid biomes is presented by an Australasian lineage of *Ficus* L. (Moraceae). The genus *Ficus* is showing an impressive array of species diversity with more than 800 known species most famously known for their intricate pollination mutualism with fig-wasps of the family Agaonidae (Cruaud et al. 2012; Clement et al. submitted; Berg & Corner 2005). *Ficus* section *Malvanthera* sensu Corner (1965) with updates from Rønsted et al. (2008b) includes 23 species (**Table 1**) of monoecious figs with a range of habits, such as hemi-epiphyte and lithophytes with an astonishing root system capable of producing aerial- and adventitious roots (Corner, 1965). The section includes species with two clearly defined growth habits; (i) hemi-epiphytic stranglers and free-standing trees in the rainforests of Eastern Australia, New Guinea, and some Australasian islands – and (ii) lithophytic shrubs/smaller trees occurring in the arid parts of Australia (Dixon, 2003).

Table 1
Ficus section Malvanthera *sensu* Corner (1965) with updates from Rønsted et al. (2008) and their habitat use.

LH—Lord Howe Island, NG—New Guinea, NSW—New South Wales, NT—Northern Territory, PAC—Pacific islands, QLD—Queensland, SA—South Australia, SOL—Solomon Islands, WA—Western Australia

Subsection	Series	Taxon	Author	Habitat use
Hesperidiiformes	*Glandiferae*	*Ficus glandifera*	Summerhayes	Transitional, forest, NG, PAC, SOL
Hesperidiiformes	*Glandiferae*	*Ficus baola*	C.C. Berg	Hemi-epiphyte, forest, SOL
Hesperidiiformes	*Glandiferae*	*Ficus rhizophoriphylla*	King	Hemi-epiphyte, forest, NG
Hesperidiiformes	*Hesperidiiformes*	*Ficus hesperidiiformis*	King	Hemi-epiphyte, forest, NG
Hesperidiiformes	*Hesperidiiformes*	*Ficus sterrocarpa*	Diels	Hemi-epiphyte, forest, NG
Hesperidiiformes	*Xylosyciae*	*Ficus augusta*	Corner	Hemi-epiphyte, forest, NG
Hesperidiiformes	*Xylosyciae*	*Ficus heteromeka*	Corner	Hemi-epiphyte, forest, NG
Hesperidiiformes	*Xylosyciae*	*Ficus mafuluensis*	Summerhayes	Hemi-epiphyte, forest, NG
Hesperidiiformes	*Xylosyciae*	*Ficus xylosycia*	Diels	Hemi-epiphyte, forest, NG, SOL
Platypodeae	*Crassipeae*	*Ficus crassipes*	F.M. Bailey	Hemi-epiphyte, forest, QLD
Platypodeae	*Crassipeae*	*Ficus destruens*	C.T. White	Hemi-epiphyte, forest, QLD
Platypodeae	*Eubracteatae*	*Ficus triradiata*	Corner	Hemi-epiphyte, forest, QLD
Platypodeae	*Obliquae*	*Ficus obliqua*	G. Forster	Transitional, QLD, NSW, NG, PAC
Platypodeae	*Obliquae*	*Ficus cerasicarpa*	D.J. Dixon	Lithophyte, arid, QLD, NT, WA
Platypodeae	*Obliquae*	*Ficus lilliputiana*	D.J. Dixon	Lithophyte, arid, NT, WA
Platypodeae	*Obliquae*	*Ficus platypoda*	(Miq.) A. Cunn. ex Miq.	Lithophyte, arid, NT, WA
Platypodeae	*Obliquae*	*Ficus subpuberula*	Corner	Lithophyte, arid, NT, WA
Platypodeae	*Rubiginosae*	*Ficus rubiginosa*	Ventenat	Transitional, QLD, NSW
Platypodeae	*Rubiginosae*	*Ficus atricha*	D.J. Dixon	Lithophyte, arid, NT, WA
Platypodeae	*Rubiginosae*	*Ficus brachypoda*	(Miquel) Miquel	Lithophyte, arid, NT, QLD, SA, WA
Platypodeae	*Rubiginosae*	*Ficus watkinsiana*	F.M. Bailey	Hemi-epihyte, forest, QLD, NSW
Malvantherae		*Ficus macrophylla*	Persoon	Tree, hemi-epiphyte, forest, QLD, NSW, LH
Malvantherae		*Ficus pleurocarpa*	F. Mueller	Hemi-epiphyte, forest, QLD,

In addition, section *Malvanthera* contains a number of species believed to be transitional as their habit and geographical range comprises both the arid and wet biomes and associated habits (Dixon, 2003; Rønsted et al., 2008b). *Ficus rubiginosa* Desf. ex Vent. and *Ficus obliqua* G. Forst. both occur as rainforest hemi-epiphytes and can be found in drier areas occurring as lithophytes and deciduous vine thickets. *Ficus glandifera* Summerhayes is part of a primarily New Guinean rainforest clade but can also be found as transitional.

It has previously been suggested that *Ficus* species with monoecious, actively pollinated figs and a hemi-epiphytic habit foster high diversification rates and possibility of living in niches other plants cannot access (Bruun-Lund et al., 2018). In addition to the diversity in habitat use and associated growth form and morphology, section *Malvanthera* also show an unusual high diversity in other traits compared to *Ficus* in general. A mix of actively and passively pollinating fig-wasps from the genus *Pleistodontes* Saunders (Agaonidae, Hymenoptera) are associated with section *Malvanthera* (Dixon, 2003) and several shifts between active and passive pollination have been proposed (Kjellberg et al., 2001).

An earlier phylogenetic study by Rønsted et al. (2008b) suggested that shifts into more arid biomes may have happened at least twice within section *Malvanthera* (Rønsted et al., 2008b). However, lack of resolution provided by a limited number of nuclear DNA regions, prohibited a deeper understanding of biome shifts in this extraordinary lineage of *Ficus*.

Tremendous work have been performed on *Ficus* using Sanger sequencing focused on rapidly evolving single-copy nuclear markers (e.g. Clement et al. submitted; Rønsted et al. 2005, 2008a; Xu et al. 2011;

Cruaud et al. 2012). However, as for many other large plant genera, the resolution and support obtained is insufficient. Data from the plastid genome has been attempted (Bruun-Lund et al., 2017; Renoult et al., 2009), but found to provide a conflicting evolutionary history compared to the nuclear data, which are aligned with current understanding based on morphology and biogeography (Berg and Corner, 2005). Target sequence capture using RNA probes allows customization of a taxon specific probe-set focusing on genomic regions evolving at a suitable rate and exclusion of organelle DNA (Gnirke et al., 2009; Horn, 2012).

The purpose of the present study therefore is to use a high through-put targeted sequence capture approach to produce a well resolved phylogenetic hypothesis for section *Malvanthera*, which will allow us to (1) test the number and direction of biome shifts between forest and drier habitats, and (2) identify potential drivers of biome shifts, and (3) test if correlated traits are likely pre-adaptations or new key-innovations.

2. MATERIALS AND METHODS

2.1. Taxon sampling

We follow the species concept of section *Malvanthera* sensu Corner (1965) updated by Dixon (2003) for the Australian species and with the addition of *Ficus baola* CC Berg (Berg, 2002). Taxon sampling was thus aimed at including multiple accessions of all species included in *Ficus* subsection *Malvanthera* Berg and Corner (2005), as well as the additional species recognized by Corner (1965). Sampling was conducted through new fieldwork in Queensland and Northern Territory of Australia, as well as in Papua New Guinea. Additional accessions were acquired from collections in the following herbaria, Harvard University (A), Natural History Museum of Denmark, University of Copenhagen (C), James Cook University (JCT), Naturalis (L), and University of Minnesota (MIN). In total, 83 accessions were included representing all 23 currently accepted species of section *Malvanthera* with multiple accessions for the majority of the species. *Ficus cyathistipula* Warb. (sect. *Galoglychia*) and *F. microcarpa* L.f. (sect. *Conosycea*) were used as outgroup to represent other sections of subgenus *Urostigma* (Table 2, Supplementary material).

2.2. *Ficus* target baits development for sequence capture

Baits were designed on two transcriptomes from the 1000 plant genome project (1KP) (Matasci et al., 2014). A transcriptome representing the *Ficus* lineage (*Ficus religiosa* L.) and the outgroup (*Morus nigra* L.) were selected. Due to the huge divergence in the figs we chose to use a representative within the genus and one from the outside to retrieve the loci of the genome which were most conserved and to allow the baits set to be applicable for other studies across the genus. Therefore, the setup is very similar to studies using UCEs (Ultra Conserved Elements) where loci should be shared across large taxonomic ranges (Baca et al., 2017). 120mer baits were designed with 2× tiling density for a total of 57,657 initial probes, which were shared with Arbor Bioscience to manufacture a myBaits Custom Target Capture kit.

To optimize the bait-kit we first sequenced 24 samples covering the phylogenetic diversity of *Ficus* for the initial 57,657 probes. Hereafter we evaluated the performance of the captured sequences and selected the probes that covered most samples with sufficient divergence.

To do this, raw reads were mapped against the initial transcriptome using PALEOMIX (Schubert et al., 2014). The mean coverage of samples varied between 0.016X and 83.566X. For all 24

samples, the mapped reads were used to generate BEDGRAPH files using *bedtools genomecov* (Quinlan and Hall, 2010) with the *–bg* option. *bedtools unionbedg* (Quinlan and Hall, 2010) was then used to combine the BEDGRAPH files into a single multi-sample-file. To select capture regions, only the 12 samples with a mean coverage higher than 6X were used (S44_sagittata, S16_americana, S37_chartacea, S30_punctata, S1_pseudopalma, S5_microcarpa, S53_cyathistipula, S46_tinctoria, S68_septica, S69_tsiangii, S39_magnoliifolia, S20_rubiginosa; Table 2). All sites with coverage higher than ten times the mean coverage and sites with coverage lower than one tenth the mean coverage (or lower than 1) for each of the 12 samples were removed. To combine overlapping or "book-ended" regions in the resulting file, *bedtools merge* (Quinlan & Hall, 2010) was used to create a single bed file. As the capture probes have a length of 120 bp, only sites longer than 120 bp were selected. A custom python script was then used to extract the selected regions from the transcriptome assembly as a fasta file, which was shared by Arbor Bioscience to manufacture a final *Ficus* targeted myBaits Custom Target Capture kit (20K probes). The final *Ficus* probe-set contained 11,951 baits covering 2,276 regions ranging from 120-806bp.

2.3. Molecular work

All work was performed in designated DNA laboratories at the Natural History Museum of Denmark, University of Copenhagen, which are dedicated labs for working with samples of low DNA concentration and strict cleaning procedures to avoid cross contamination.

2.3.1. DNA extraction

Extractions were performed on 12-40mg of silica-dried leaves or herbarium tissue (often using less tissue to avoid disruptive sampling of herbarium collections as much as possible). A TissueLyser II (Qiagen®) was used to rupture cell walls before starting the extraction with KingFisher Duo Prime System (Thermo Scientific) and Thermo Scientific KingFisher Pure DNA Plant kit protocol (Thermo Fisher Scientific, Waltham, Massachussets, USA). In order to increase the yield of DNA recovered, lysis buffer A was supplemented with 2% polyvinylpyrrolidone (PVP) as recommended in the protocol. likewise, 25µL of Proteinase K was added to increase DNA recovery. Samples were incubated at 65°C on a rotor for 1 hour. Finally, samples were eluted in 50µL of EB buffer (instead of 100µL) to yield higher concentration. DNA concentration was estimated using a Qubit® 2.0

fluorometer (high sensitivity) following the manufacturer's protocol. The degree of degradation and general quality of the extraction was evaluated using the Agilent 2200 TapeStation® System. Sheering using Covaris M220 (Covaris, Wobum, MA, USA) was applied to modern samples with long DNA fragments.

2.3.2. Library preparation and indexing

Two library preparation protocols were applied, from here on referred to as 'NEBnext' and 'BEMT'; The NEBnext protocol was used for the initial probe development and BEMT was subsequently used for all samples due to it being optimized for plants and cheaper to use for library preparation in general (Carøe et al., 2017). For the NEBnext protocol Illumina-compatible 100 bp paired-end libraries from DNA extracts using the NEBNext DNA Library Prep Kit for Illumina (New England Biolabs) following manufacturer's protocol or in house protocols as in Bruun-Lund et al. (2017).

For the blut-end single-tube protocol 'BEST' (Carøe et al., 2017) the following modifications where applied: A column purification was done using Monarch® DNA Cleanup Columns after the end-repair

reaction to reduce the presence of enzymatic inhibitors. 450 µl enhanced PB binding buffer (Allentoft et al., 2015) and centrifugation at 6,000 *g*. The column was washed with 800 µl PE buffer and spun at 10,000 *g*, followed by an additional spin for 3 min at 17,000 *g*. DNA was eluted in 34 µl EB buffer, with incubation at 37°C for 10 min before collecting DNA at 17,000 *g*. For the ligation reaction, 2 µl Illumina adapters (20 mM) (Meyer and Kircher, 2010) were mixed with the end-repaired DNA. The final libraries were purified using Solid Phase Reversible Immobilization (SPRI) beads (Cat#: GE45152105050250; Sigma-Aldrich).

Libraries were indexed and amplified using P7 and P5 Illumina primers with a dual indexing approach to minimize tag jumping (Schnell et al., 2015). PCR was performed in 50 µl reactions using 10 µl template,1× AmpliTaq Gold buffer, 2.5 mM MgCl2, 0.8 µg/µl Bovine Serum Albumin (BSA), 0.25 mM dNTP, 0.2 µM forward and reverse indexed primer (specific for each sample), and 0.2 U/µl AmpliTaq Gold polymerase. Libraries were amplified in an Applied Biosystems 2720 Thermal Cycler using the following conditions: 95°C for 12 min, followed by a number of cycles of 95°C for 20 s, 60°C for 30 s, and 72°C for 40 s, followed by 5 min at 72°C. cycle number varied depending on

sample between 15-25 cycles. Quantification and size estimation were performed with an Agilent 2100 Bioanalyzer high sensitivity kit. The amplified library was cleaned using SPRIbeads and eluted in 30 µL of EB buffer.

2.3.3. Capture

All amplified libraries were concentrated using the SpeedVac (Thermo Scientific) until they contained 100-500 ng DNA in 7.14 µl. The libraries were captured following the manufacturer's instructions (http://www.arborbiosci.com/wp-content/uploads/2018/04/myBaits-Manual-v4.pdf or V3 for the initial 24 samples for bait optimization design). The hybridization solution incubated in a Veriti 96-well thermal cycler (Applied Biosystems, Foster City, CA) for approximately 20 hours at 65 °C during which time the targeted DNA sequences hybridized to the biotinylated RNA baits. Post-captured libraries were PCR amplified using KAPA HiFi HotStart ReadyMix (Kapa Biosystems, Wilmington, MA) for 10–20 cycles using primers IS5 (5′-AATGATACG GCGACCACCGA) and IS6 (5′-CAAGCAGAAGACGGCA TACGA). The final purification of the amplified post-capture reactions was done using the Qiagen QIAquick PCR Purification Kit following

manufacturer's protocol and eluting in 30 µl EB buffer incubated for 15 min at 37 °C.

Captured samples were pooled in equimolar amounts into one lane and sequenced (100 bp paired-end reads) for the initial 24 samples on a HiSeq2500 instrument. The second batch of *Malvanthera* species was sequenced (150 bp paired-end reads) on a HiSeq4000 instrument, both batches at the National High-throughput Sequencing Centre, Natural History Museum of Denmark, Copenhagen, Denmark.

2.4. Phylogenomics analyses

For preliminary analysis, sequence read data from each sample was processed using the *PALEOMIX* pipeline (Schubert et al., 2014). Firstly, low quality and missing bases were trimmed from the reads, followed by removal of leftover adaptors using *AdaptorRemoval2* (Lindgreen, 2012). All paired-end reads where there was overlap by more than 10 base pairs were merged into one read. Subsequently, the reads from each of the samples were mapped to the a pseudo-reference, which is the optimized probes using *bwa* (Li et al., 2009). Mapped reads were likewise filtered for PCR duplicates using *Picard*

(https://broadinstitute.github.io/picard), and reads mapping to multiple loci in the genome were excluded from the analysis.

GATK was used for genotype calls (Lee et al., 2015) and hereafter a multi-sample consensus fasta file was created and used for input for *RaxML* (Stamatakis, 2006) using model GTR+GAMMA and the algorithm 'rapid bootstrapping and search for best-scoring ML tree' with 200 Bootstrap replicates.

2.5. Diversification rate analysis, molecular dating etc.

2.5.1. Molecular dating.

On the final dataset, we expect to use *BEAST* (Drummond et al., 2012), a Bayesian method (MB) that estimates model parameters, tree topology and divergence times, to estimate time-calibrated trees of section *Malvanthera*. Due to the poor fossil record of *Ficus,* we will explore both fossil, geological, and secondary dating points.

2.5.2. Ancestral state reconstruction.

Ancestral state analysis will be applied to investigate the dynamics of habitat change and biome shifts throughout the evolutionary history of *Malvanthera*.

For Bayesian ancestral state reconstruction, we will use a reversible jump Markov chain Monte Carlo method (Pagel and Meade, 2006), contained in the *BAYESTRAITS* package, which automatically finds the posterior distribution of models if provided with the appropriate choice of priors (Pagel et al., 2004). This yields posterior distributions of rate coefficients and state probabilities, incorporating phylogenetic uncertainty as well as uncertainty in the ancestral state reconstruction.

2.5.3. Diversification rate analysis.

To test for potential bursts in diversification rates we will explore lineage through time (LTT) plots and *BAMM* (Bayesian Analysis of Macro-evolutionary Mixtures) analysis. *BAMM* tests for changes in diversification rates – if a rise in rate happens when ancestral area change (e.g. from wet to arid habitats) it may indicate novel traits have evolved to exploit the new habit. If no change is seen it could indicate that a pre-adaptation aided the survival in these new environments.

Trees will be pruned to make clades with multiple accessions per species into a single tip, which is needed for the subsequent analyses. A LTT plot will be produced with the R package *paleotree* 2.7 (Bapst, 2012). One thousand random post-burnin trees from the *BEAST* inference

analysis will be selected as input for the analysis. Subsequently, the software *BAMM* 2.5 (Bayesian Analysis of Macro-evolutionary Mixtures) will be used to test if the diversification rate and extinction rate have been equal through time for section *Malvanthera* (Rabosky et al., 2014). The MCMC analysis will be run 1 million generations with sampling every 1000 generations. Check for convergence will be performed with the R package *coda* 0.19–1 (Plummer et al., 2006) and the first 10% will be expected to be discarded as burn-in. Hereafter we will use the R package *BAMMtools* 2.1.6 (Rabosky et al., 2014) and *ape* 4.0 (Paradis et al., 2004) to summarize rates over each branch of the phylogenetic tree (producing the so-called 'phylorate' plot), to plot the 95% credible shift set (CSS) with sampling frequencies of the different shift configurations, and to obtain the shift configuration with the maximum a posteriori (MAP) probability.

2.5.4. Traits and transitions between rainforest and savanna biomes

State-dependent diversification rate tests will be performed with the Binary-State Speciation and Extinction (*BiSSE*) model (Maddison et al., 2007) as implemented in the R package *diversitree* 0.9–9 (Fitzjohn,

2012) with traits linked to adaptions for either arid or wet biomes, which is to be decided upon.

3. RESULTS

3.1. Target sequence capture data

We designed our initial set of probes using transcriptome data from on *Ficus religiosa* and *Morus nigra* as described in the method section. These probes where used on a subset of samples representing the entire genus' phylogenetic diversity. This was done with the aim to optimize the probe set to only contain probes that work for the entire genus. Thus, optimizing sequencing effort and costs. The first round of sequencing using the initial probes worked well for 12 of the 24 samples, producing coverage >6× for all sites. The resulting unrooted phylogenetic hypothesis based on sequence capture can be seen in **Figure 1**. Each of the 12 species are representing clades and the evolutionary relationships between theses clades is agreeing with findings in Cruaud et al. (2012) and Clement et al. (submitted – **Chapter I**). Additionally, the resolution and bootstrap support is much higher for the backbone of the phylogenetic tree, with only one branch below 100% bootstrap (see **Figure 1**) compared to these previous studies. Backbone

resolution has previously been problematic in assessing the global phylogeny using nuclear markers (Cruaud et al. 2012; Clement et al. submitted – **Chapter I**).

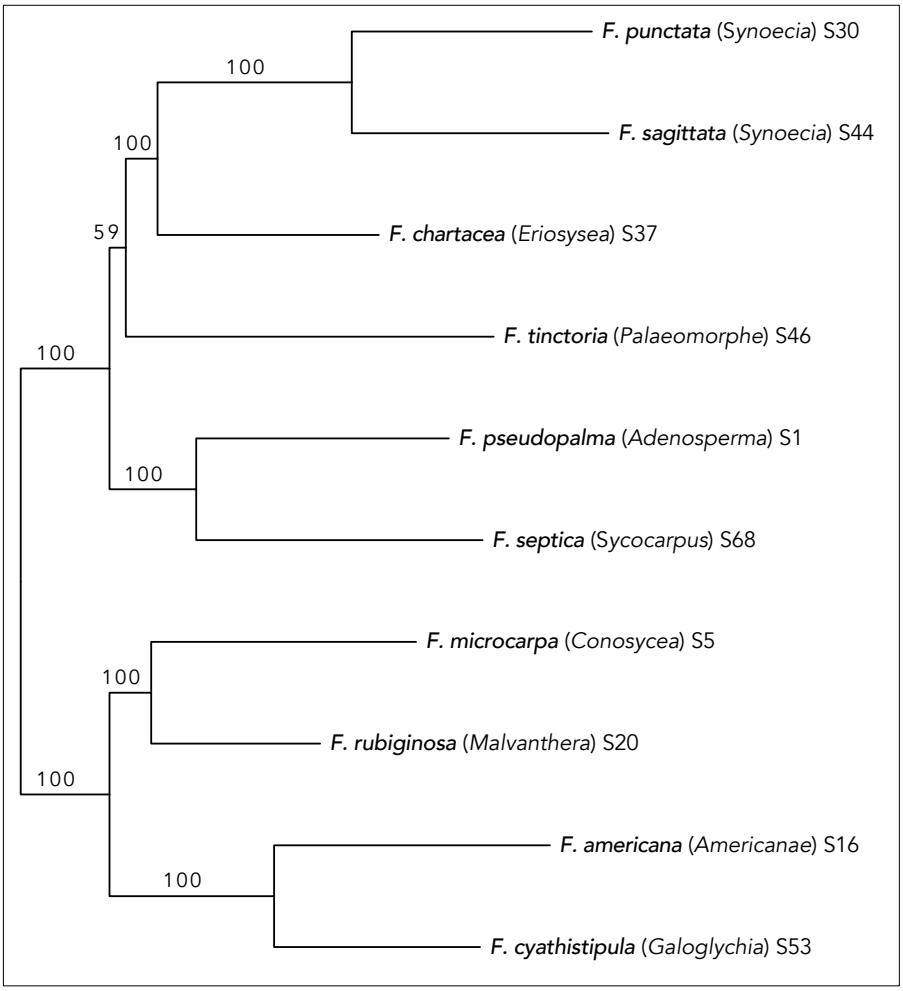


Figure 1. Unrooted phylogenetic hypothesis based on the initial set of probes. Relationships are strongly supported and corresponding with previous findings. Clade names according to Clement et al. (submitted – Chapter I) in parenthesis after each species name.

The optimized probe set was used on section *Malvanthera* for a preliminary subset of the samples – 20 samples were sequenced in the initial run, with 16 samples having coverage higher than >6× (ranging between 0.53×-688×). Three samples with coverage <6× were excluded from later phylogenetic analysis. *RaxML* produced a preliminary phylogenetic hypothesis for the section (**Figure 2**), see below.

3.2. Phylogeny

Preliminary maximum likelihood analysis included 16 taxa of section *Malvanthera* and 2 outgroup taxa. The backbone is well resolved and medium to highly supported (>80% bootstrap support), whereas some internal branches are not well supported. Within *Malvanthera*, the primarily New Guinean rainforest clade of subsection *Hesperidiiformes* (Clade A; *F. augusta, F. hesperidiiformis, F. heteromeka*, and *F. xylosycia*) is sister to the primarily continental Australian *Malvanthera* species (Clade B). Within this major Australian clade, subsection *Malvantherae* consisting of two rainforest species (Clade C; *F. macrophylla* and *F. pleurocarpa*), is sister to the remainder of the Australian taxa (Clade D). Clade D consists of two subclades. Clade E includes four lithophytic accessions (2X *F. brachypoda, F. cerasicarpa*,

and *F. subpuberula*) – Note that *F. brachypoda* might consist of two separate species (Peter Jobson, personal observation). Clade F include a gradient of three rainforest species (*F. crassipes, F. destruens* and *F. watkinsiana*) followed by a clade with a transitional species (*F. rubiginosa*) and two accessions of a lithophytic species (*F. platypoda*).

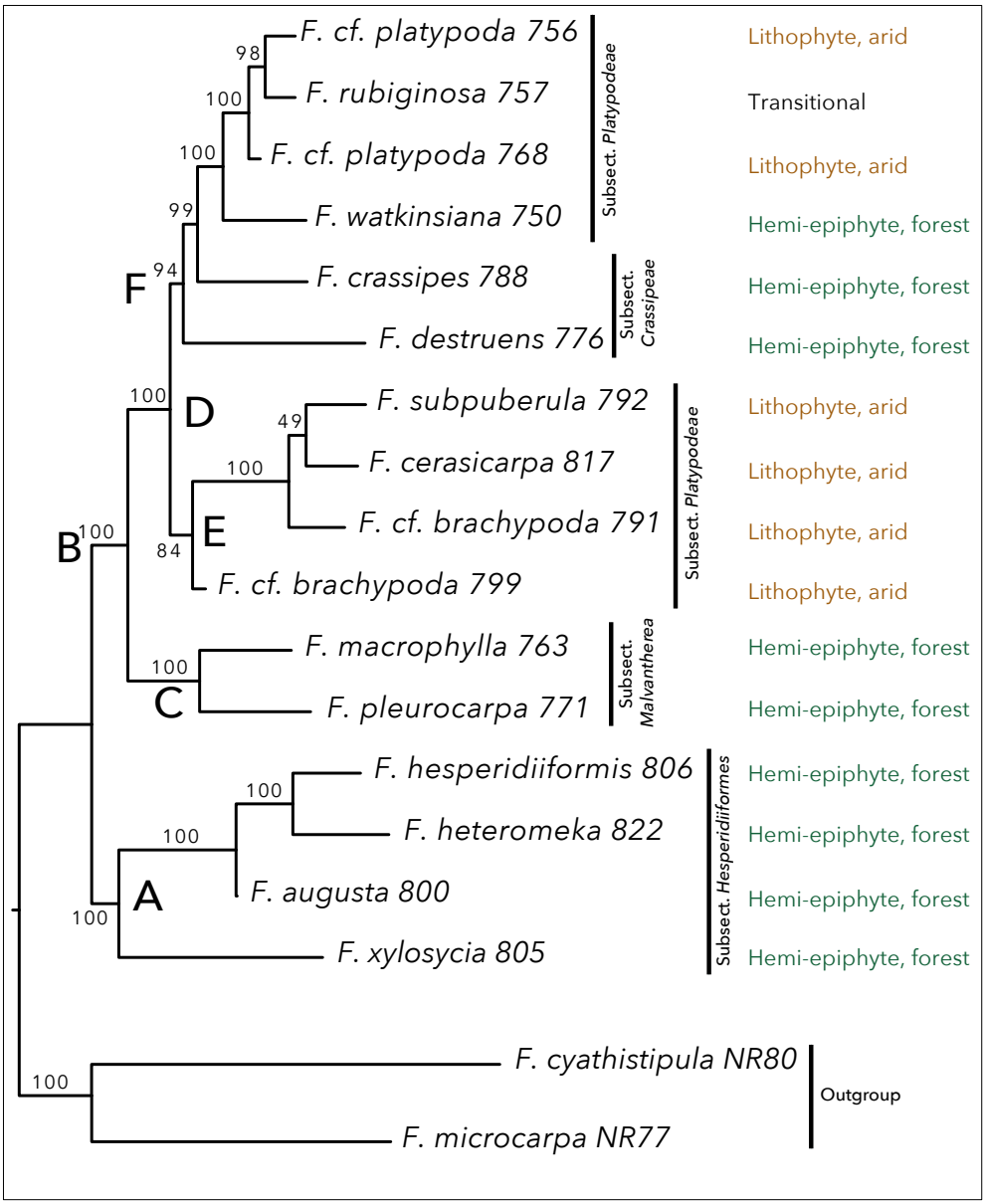


Figure 2. Preliminary Maximum Likelihood inference of section *Malvanthera* with habitat use mapped to each species. Clade A represents a primarily Papua New Guinean clade and clade B represents a primarily Australian clade. Bootstrap support is specified, which is high (100%) on most branches. Biome, habit and sub-sectional classification is indicated on the right.

4. DISCUSSION

4.1. Biome shifts out of the forest into drier habitats

As the understanding of biome shifts and adaptations are closely related to climate changes it is very important to increase our understanding in this area. This is especially true in a diverse group like *Ficus* with a variety of traits and habits. The wide range of traits in a well-known system such as *Ficus* makes the figs an ideal group of plants to focus on in the objective of improving our understanding of the evolutionary patterns of plant diversification.

The preliminary analysis including only 16 taxa of *Malvanthera*, support the initial results of Rønsted et al. (2008b) suggesting a pattern of several biome shifts from rainforest to drier habitats associated with rainforest strangling hemi-epiphytic and lithophytic habit respectively. Only one transitional species, *F. rubiginosa* is included in this preliminary dataset, and it forms a clade with both lithophytic taxa and rainforest species. However, interpretations about both classification and biome shifts will have to await more rigorous analysis of the final dataset including all species of *Malvanthera*, most of which will be represented by multiple accessions.

4.2. Trait dependent diversification – pre-adaptations or key-innovations?

Species with the hemi-epiphytic habit (i) have been shown to diversify faster, and this lifeform could potentially have aided their survival in niches other plants cannot occupy (Bruun-Lund et al., 2018). The hemi-epiphytic habit together with other suggested traits such as small fig size, phenology and active pollination (Bruun-Lund et al., 2018; Harrison, 2005; Machado et al., 2018), might have aided their apparent radiation from humid rainforests to the dry savannas observed by Rønsted et al. (2008b) in the Australasian section *Malvanthera*.

4.3. Targeted gene capture method

Genetic or genomic data sequence data allows us to test previously suggested hypotheses on clades and species relationships, and also improve our understanding of organisms' evolutionary history. Here we applied a target sequence capture approach, using custom myBaits (Arbor Bioscience, MA, USA) designed to work on the highly divergent genus *Ficus* L. (Moraceae), with the aim to resolve deeper nodes that are hampered by uncertainty. This work is a case-study of Australasian *Ficus* section *Malvanthera* Corner*,* however, the probes designed, are usable

for the entire genus and have shown promising results (Bruun-Lund et al., in prep.) for resolving deep nodes with short branch lengths that have otherwise hindered the elucidation of the evolutionary history of this charismatic pan-tropical genus.

Accordingly, much higher coverage can be obtained using our optimized probe set compared to genome skimming approaches (**Chapter IV**), thus making it possible to multiplex and pool more samples per run and hereby lower sequencing costs, while still giving sufficient resolution of the inferred phylogenetic tree (**Figure 1 & 2**).

ACKNOWLEDGEMENTS

Charlotte Hansen is thanked for help with producing DNA libraries. Thomas Kunda, Mentap Sisol and the Binatang Research Centre for assistance during fieldwork in Papua New Guinea. Estrid Huss Eriksson for assistance during fieldwork in Australia. Jason Halford and xxx for help with sampling in Brisbane City Botanic Garden Mt. Cootcha Botanic Garden respectively. Anthony Brach (A), Betsy Jackes (JCT), and Nicolien Sol (L) for samples from the respective herbaria. Thanks to Siavash Mirarab & James H. Leebens-Mack for supplying us with transcriptome data from the 1K-Project for initial probe design.

REFERENCES

Allentoft, M.E., Malaspinas, A.-S., Khartanovich, V., Rasmussen, M., Vicze, M., Jarysz, R., Rasmussen, S., Trifanova, S. V., Harvig, L., Margaryan, A., Shishlina, N., Tóth, G., Merkevicius, A., Merkyte, I., Frei, K., Brunak, S., Kriiska, A., Pokutta, D., Mkrtchyan, R., Ahlström, T., Higham, T., Zhitenev, V., Soenov, V.I., Baron, JustynaPrice, T.D., Sicheritz-Pontén, T., Khokhlov, A., Lynnerup, N., Furmanek, M., Duffy, P.R., Kristiansen, K., Ebel, A. V., Dąbrowski, P., Sablin, M., Stenderup, J., Epimakhov, A., Paja, L., Smrčka, V., Grupe, G., Kiss, V., Schroeder, H., Kolář, J., Hajdu, T., Szeverényi, V., Orlando, L., Nielsen, R., Sjögren, K.-G., Damgaard, P.B., Saag, L., Moiseyev, V., McGlynn, G., Gromov, A., Sikora, M., Vinner, L., Chivall, D., Metspalu, M., Pálfi, G., Longhi, C., Varul, L., Pospieszny, Ł., Yepiskoposyan, L., Gralak, T., Gronkiewicz, S., Lasak, I., Casa, P. Della, Willerslev, E., 2015. Population genomics of Bronze Age Eurasia. *Nature 522*, 167–172. https://doi.org/10.1038/nature14507

Baca, S.M., Alexander, A., Gustafson, G.T., Short, A.E.Z., 2017. Ultraconserved elements show utility in phylogenetic inference of *Adephaga* (Coleoptera) and suggest paraphyly of 'Hydradephaga.' *Systematic Entomology*. https://doi.org/10.1111/syen.12244

Bapst, D.W., 2012. Paleotree: An R package for paleontological and phylogenetic analyses of evolution. *Methods in Ecology and Evolution 3*, 803–807. https://doi.org/10.1111/j.2041-210X.2012.00223.x

Berg, C.C., 2002. *Ficus baola*, a new species of *Ficus* subgenus *Urostigma* section *Malvanthera* (Moraeae) from the Solomon Islands. *Blu 47*, 315–317.

Berg, C.C., Corner, E.J.H., 2005. Flora Malesiana, Series I. Volume 17 part 2., in: Nooteboom, H. (Ed.), Flora Malesiana Series I - Seed Plants Vol. 17 Part 2. Nationaal Herbarium Nederland, Leiden, pp. 1–730.

Bruun-Lund, S., Clement, W.L., Kjellberg, F., Rønsted, N., 2017. First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). *Molecular Phylogenetics and Evolution 109*, 93–104. https://doi.org/10.1016/j.ympev.2016.12.031

Bruun-Lund, S., Verstraete, B., Kjellberg, F., Rønsted, N., 2018. Rush hour at the Museum – Diversification patterns provide new clues for the success of figs (*Ficus* L., Moraceae). *Acta Oecologica 90*, 4–11. https://doi.org/10.1016/j.actao.2017.11.001

Carøe, C., Gopalakrishnan, S., Vinner, L., Mak, S.S.T., Sinding, M.H.S., Samaniego, J.A., Wales, N., Sicheritz-Pontén, T., Gilbert, M.T.P., 2017. Single-tube library preparation for degraded DNA. *Methods in Ecology and Evolution 9*, 1–10. https://doi.org/10.1111/2041-210X.12871

Corner, E.J.H., 1965. Check-list of *Ficus* in Asia and Australasia with keys to identification. Botanic Gardens Singapore, 1965. *Garden Bulletin Singapore 21*, 1–186.

Crayn, D.M., Rossetto, M., Maynard, D.J., 2006. Molecular phylogeny and dating reveals an Oligomiocene radiation of dry-adapted shrubs (former Tremandraceae)

from rainforest tree progenitors (Elaeocarpaceae) in Australia. *American Journal of Botany 93*, 1328–1342. https://doi.org/10.3732/ajb.93.9.1328

Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P., Linder, H.P., 2009. Phylogenetic biome conservatism on a global scale. *Nature 458*, 754–756. https://doi.org/10.1038/nature07764

Cruaud, A., Rønsted, N., Chantarasuwan, B., Chou, L.S., Clement, W.L., Couloux, A., Cousins, B., Genson, G., Harrison, R.D., Hanson, P.E., Hossaert-McKey, M., Jabbour-Zahab, R., Jousselin, E., Kerdelhué, C., Kjellberg, F., Lopez-Vaamonde, C., Peebles, J., Peng, Y.-Q., Pereira, R.A.S., Schramm, T., Ubaidillah, R., van Noort, S., Weiblen, G.D., Yang, D.-R., Yodpinyanee, A., Libeskind-Hadas, R., Cook, J.M., Rasplus, J.-Y., Savolainen, V., 2012. An extreme case of plant-insect codiversification: figs and fig-pollinating wasps. *Systematic Biology 61*, 1029–47. https://doi.org/10.1093/sysbio/sys068

Davies, T.J., Barraclough, T.G., Chase, M.W., Soltis, P.S., Soltis, D.E., Savolainen, V., 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences 101*, 1904–1909. https://doi.org/10.1073/pnas.0308127100

Dixon, D., 2003. A taxonomic revision of the Australian *Ficus* species in the section *Malvanthera* (Ficus subg. *Urostigma*: Moraceae). *Telopea 10*, 125–153. https://doi.org/10.7751/telopea20035611

Donoghue, M.J., 2008. A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.0801962105

Donoghue, M.J., Edwards, E.J., 2014. Biome shifts and niche evolution in plants. *Annual Review of Ecology, Evolution, and Systematics 45*, 547–572. https://doi.org/10.1146/annurev-ecolsys-120213-091905

Donoghue, M.J., Sanderson, M.J., 2015. Confluence, synnovation, and depauperons in plant diversification. *New Phytologist 207*, 260–274.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*. https://doi.org/10.1093/molbev/mss075

Edwards, E.J., Osborne, C.P., Strömberg, C.A.E., Smith, S. a, 2010. The origins of C4 grasslands: Integrating evolutionary and ecosystem science. *Science 328*, 587–590. https://doi.org/10.1126/science.1177216

Fitzjohn, R.G., 2012. Diversitree: Comparative phylogenetic analyses of diversification in R. *Methods in Ecology and Evolution 3*, 1084–1092. https://doi.org/10.1111/j.2041-210X.2012.00234.x

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E.M., Brockman, W., Fennell, T., Giannoukos, G., Fisher, S., Russ, C., Gabriel, S., Jaffe, D.B., Lander, E.S., Nusbaum, C., 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature Biotechnology*. https://doi.org/10.1038/nbt.1523

Harrison, R.D., 2005. Figs and the diversity of tropical rainforests. *BioScience 55*, 1053. https://doi.org/10.1641/0006-3568(2005)055[1053:FATDOT]2.0.CO;2

Heard, S.B., Hauser, D.L., 1995. Key evolutionary innovations and their ecological mechanisms. *Historical Biology 10*, 151–173. https://doi.org/10.1080/10292389509380518

Heibl, C., Renner, S.S., 2012. Distribution models and a dated phylogeny for chilean oxalis species reveal occupation of new habitats by different lineages, not rapid adaptive radiation. *Systematic Biology 61*, 823–834. https://doi.org/10.1093/sysbio/sys034

Horn, S., 2012. Target enrichment via DNA hybridization capture. *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-61779-516-9_21

Hughes, C.E., Pennington, R.T., Antonelli, A., 2013. Neotropical plant evolution: assembling the big picture. *Botanical Journal of the Linnean Society 171*, 1–18. https://doi.org/10.1111/boj.12006

Huskins, C.L., 1930. The origin of *Spartina Townsendii*. *Genetica 12*, 531–538. https://doi.org/10.1007/BF01487665

Kjellberg, F., Jousselin, E., Bronstein, J.L., Patel, A., Yokoyama, J., Rasplus, J., 2001. Pollination mode in fig wasps: The predictive power of correlated traits. *Proceedings of the Royal Society B: Biological Sciences 6*. https://doi.org/10.1098/rspb.2001.1633

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

Lee, T.R.C., Cameron, S.L., Evans, T.A., Ho, S.Y.W., Lo, N., 2015. The origins and radiation of Australian Coptotermes termites: From rainforest to desert dwellers. *Molecular Phylogenetics and Evolution 82*, 234–244. https://doi.org/10.1016/j.ympev.2014.09.026

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 (Oxford, England) 25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lindgreen, S., 2012. AdapterRemoval: easy cleaning of next generation sequencing reads. *BMC Research Notes 5*, 337. https://doi.org/10.1186/1756-0500-5-337

Machado, A.F.P., Rønsted, N., Bruun-Lund, S., Pereira, R.A.S., Paganucci de Queiroz, L., 2018. Atlantic forest to the all Americas: Biogeographical history and divergence times of Neotropical *Ficus* (Moraceae). *Molecular Phylogenetics and Evolution 122*, 46–58. https://doi.org/10.1016/j.ympev.2018.01.015

Maddison, W.P., Midford, P.E., Otto, S.P., 2007. Estimating a binary character's effect on speciation and extinction. *Systematic biology 56*, 701–10. https://doi.org/10.1080/10635150701607033

Matasci, N., Hung, L., Yan, Z., Carpenter, E.J., Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Ayyampalayam, S., Barker, M., Burleigh, J.G., Gitzendanner, M.A.,

Wafula, E., Der, J.P., Claude, W., Roure, B., Shaw, J.A., Degironimo, L., Stevenson, D.W., Villarreal, J.C., Chen, T., Kutchan, T.M., Rolf, M., Baucom, R.S., Deyholos, M.K., Samudrala, R., Tian, Z., Wu, X., Sun, X., Zhang, Y., Wang, J., Leebens-mack, J., Wong, G.K., 2014. Data access for the 1,000 Plants (1KP) project. *GigaScience 3*, 1–10.

Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor Protocols 5*. https://doi.org/10.1101/pdb.prot5448

Pagel, M., Meade, A., 2006. Bayesian Analysis of Correlated Evolution of Discrete Characters by Reversible-Jump Markov Chain Monte Carlo. *The American Naturalist 167*, 808–825. https://doi.org/10.2307/3844739

Pagel, M., Meade, A., Barker, D., 2004. Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology*. https://doi.org/10.1080/10635150490522232

Paradis, E., Claude, J., Strimmer, K., 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics 20*, 289–290. https://doi.org/10.1093/bioinformatics/btg412

Pennington, R.T., Lavin, M., Oliveira-Filho, A., 2009. Woody plant diversity, evolution, and ecology in the tropics: Perspectives from seasonally dry tropical forests. *Annual Review of Ecology, Evolution, and Systematics 40*, 437–457. https://doi.org/10.1146/annurev.ecolsys.110308.120327

Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: convergence diagnosis and output analysis for MCMC. *R News 6*, 7–11. https://doi.org/10.1159/000323281

Quinlan, A.R., Hall, I.M., 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics 26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033

Rabosky, D.L., Donnellan, S.C., Grundler, M., Lovette, I.J., 2014. Analysis and visualization of complex macroevolutionary dynamics: An example from Australian Scincid lizards. *Systematic Biology 63*, 610–627. https://doi.org/10.1093/sysbio/syu025

Renoult, J.P., Kjellberg, F., Grout, C., Santoni, S., Khadari, B., 2009. Cyto-nuclear discordance in the phylogeny of *Ficus* section *Galoglychia* and host shifts in plant-pollinator associations. *BMC Evolutionary Biology 9*, 248. https://doi.org/10.1186/1471-2148-9-248

Rønsted, N., Weiblen, G.D., Clement, W.L., Zerega, N.J.C., Savolainen, V., 2008a. Reconstructing the phylogeny of figs (*Ficus*, Moraceae) to reveal the history of the fig pollination mutualism. *Symbiosis 45*, 1–12.

Rønsted, N., Weiblen, G.D., Cook, J.M., Salamin, N., Machado, C.A., Savolainen, V., 2005. 60 million years of co-divergence in the fig-wasp symbiosis. *Proceedings of the Royal Society B: Biological Sciences 272*, 2593–2599. https://doi.org/10.1098/rspb.2005.3249

Rønsted, N., Weiblen, G.D., Savolainen, V., Cook, J.M., 2008b. Phylogeny,

biogeography, and ecology of *Ficus* section *Malvanthera* (Moraceae). *Molecular phylogenetics and evolution 48*, 12–22. https://doi.org/10.1016/j.ympev.2008.04.005

Schnell, I.B., Bohmann, K., Gilbert, M.T.P., 2015. Tag jumps illuminated - reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources 15*, 1289–1303. https://doi.org/10.1111/1755-0998.12402

Schubert, M., Ermini, L., Sarkissian, C. Der, Jónsson, H., Ginolhac, A., Schaefer, R., Martin, M.D., Fernández, R., Kircher, M., McCue, M., Willerslev, E., Orlando, L., 2014. Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nature Protocols 9*, 1056–1082. https://doi.org/10.1038/nprot.2014.063

Smith, S.A., Beaulieu, J.M., Stamatakis, A., Donoghue, M.J., 2011. Understanding angiosperm diversification using small and large phylogenetic trees. *American Journal of Botany 98*, 404–414. https://doi.org/10.3732/ajb.1000481

Stamatakis, A., 2006. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics 22*, 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

Toon, A., Crisp, M.D., Gamage, H., Mant, J., Morris, D.C., Schmidt, S., Cook, L.G., 2015. Key innovation or adaptive change? A test of leaf traits using Triodiinae in Australia. *Scientific Reports 5*, 1–12. https://doi.org/10.1038/srep12398

Webb, C.O., Donoghue, M.J., Wurdack, K.J., Jaramillo, C.A., Davis, C.C., 2005. Explosive radiation of Malpighiales supports a mid-cretaceous origin of modern tropical rain forests. *The American Naturalist 165*, E36–E65. https://doi.org/10.1086/428296

Xu, L., Harrison, R.D., Yang, P., Yang, D.R., 2011. New insight into the phylogenetic and biogeographic history of genus Ficus: Vicariance played a relatively minor role compared with ecological opportunity and dispersal. *Journal of Systematics and Evolution 49*, 546–557. https://doi.org/10.1111/j.1759-6831.2011.00155.x

FUNDING INFORMATION

The research included in this study was funded by a research grant from the Danish Council for Independent Research - Natural Sciences (#7014-00205) and a Fieldwork grant from the Carlsberg Foundation to NR and (#SBL CF16-0883).

AUTHOR CONTRIBUTIONS

SBL and NR designed the research. SBL performed the research with input from all authors. SBL, NR, and SS conducted fieldwork in Australia and PNG. SBL, JS, VB, and MM, designed the gene capture baits and bioinformatics

pipelines. SBL and NR wrote the paper with contributions from all authors.

DATA ACCESSABILITY

The genomic data and alignments will be available at Dryad and treeBASE.

SUPPORTING INFORMATION

Table 2 (next pages). Voucher information for each specimen to be included in the study. Letters in parenthesis indicate herbarium where voucher is housed: A: Harvard University; BG: University of Bergen; C: University of Copenhagen; HITBC: Xishuangbanna Tropical Botanical Garden, Academia Sinica; JCT: James Cook University; L: Naturalis; MPU: University of Montpellier; PUH: University of the Philippines; K: Royal Botanical Gardens, Kew; THNHM: Thailand Natural History Museum.

Lab-ID	Section	Subsection	Series	Species	Collector	Country.State	Location
750	Malvanthera	Platypodeae	Rubiginosae	*Ficus watkinsiana* F.M. Bailey	N. Rønsted & S. Bruun-Lund 750 (C)	Australia, QLD	Lamington National Park
751	Malvanthera	Platypodeae	Rubiginosae	*Ficus watkinsiana* F.M. Bailey	N. Rønsted & S. Bruun-Lund 751 (C)	Australia, QLD	Lamington National Park
752	Malvanthera	Platypodeae	Rubiginosae	*Ficus watkinsiana* F.M. Bailey	N. Rønsted & S. Bruun-Lund 752 (C)	Australia, QLD	Lamington National Park
753	Malvanthera	Platypodeae	Obliquae	*Ficus obliqua* G.Forster	N. Rønsted & S. Bruun-Lund 753 (C)	Australia, QLD	Brisbane, Mt. Cootcha, Botanic Garden
754	Malvanthera	Platypodeae	Crassipes	*Ficus crassipes* F.M. Bailey	N. Rønsted & S. Bruun-Lund 754 (C)	Australia, QLD	Brisbane, Mt. Cootcha, Botanic Garden
755	Malvanthera			*Ficus sp.*	N. Rønsted & S. Bruun-Lund 755 (C)	Australia, QLD	Brisbane, Mt. Cootcha, Botanic Garden
756	Malvanthera	Platypodeae	Obliquae	*Ficus platypoda* (Miq.) A. Cunn. ex Miq.	N. Rønsted & S. Bruun-Lund 756 (C)	Australia, QLD	Brisbane, Mt. Cootcha, Botanic Garden
757	Malvanthera	Platypodeae	Rubiginosae	*Ficus rubiginosa* Ventenat	N. Rønsted & S. Bruun-Lund 757 (C)	Australia, QLD	Brisbane, Mt. Cootcha, Botanic Garden
762	Malvanthera	Platypodeae	Rubiginosae	*Ficus watkinsiana* F.M. Bailey	N. Rønsted & S. Bruun-Lund 762 (C)	Australia, QLD	Brisbane, Mt. Cootcha, Botanic Garden
763	Malvanthera	Malvantherae		*Ficus macrophylla* Desf. ex. Pers. subsp. *macrophylla*	N. Rønsted & S. Bruun-Lund 763 (C)	Australia, QLD	Brisbane, Mt. Cootcha, Botanic Garden
765	Malvanthera	Malvantherae		*Ficus macrophylla* Desf. ex. Pers. subsp. *Macrophylla*	N. Rønsted & S. Bruun-Lund 765 (C)	Australia, QLD	Brisbane, City Botanic Gardens
768	Malvanthera	Platypodeae	Obliquae	*Ficus platypoda* (Miq.) A. Cunn. ex Miq.	N. Rønsted & S. Bruun-Lund 768 (C)	Australia, QLD	Brisbane, City Botanic Gardens
771	Malvanthera	Malvantherae		*Ficus pleurocarpa* F. Mueller	N. Rønsted & S. Bruun-Lund 771 (C)	Australia, QLD	Lake Echam
772	Malvanthera	Platypodeae	Rubiginosae	*Ficus watkinsiana* F.M. Bailey	N. Rønsted & S. Bruun-Lund 772 (C)	Australia, QLD	Lake Echam
773	Malvanthera	Platypodeae	Obliquae	*Ficus obliqua* G.Forster	N. Rønsted & S. Bruun-Lund 773 (C)	Australia, QLD	Lake Echam
774	Malvanthera	Platypodeae	Rubiginosae	*Ficus rubiginosa* Ventenat	N. Rønsted & S. Bruun-Lund 774 (C)	Australia, QLD	Yungaburra
775	Malvanthera			*Ficus sp.*	N. Rønsted & S. Bruun-Lund 775 (C)	Australia, QLD	Atherton Tables Yungaburra
776	Malvanthera	Platypodeae	Crassipes	*Ficus destruens* C.T. White	N. Rønsted & S. Bruun-Lund 776 (C)	Australia, QLD	Millaa falls
777	Malvanthera	Platypodeae	Obliquae	*Ficus obliqua* G.Forster	N. Rønsted & S. Bruun-Lund 777 (C)	Australia, QLD	Millaa falls, 100 meters down turn off to Millaa Millaa falls on falls circuit, by carpark on roadside

778	Malvanthera			*Ficus sp.*	N. Rønsted & S. Bruun-Lund 778 (C)	Australia, QLD	Millaa falls. Parking spot by Millaa Millaa falls
779	Malvanthera	Malvantherae		*Ficus pleurocarpa* F. Mueller	N. Rønsted & S. Bruun-Lund 779 (C)	Australia, QLD	Millaa falls. On road to Millaa Millaa falls, just where the forrest starts.
780	Malvanthera	Platypodeae	Obliquae	*Ficus obliqua* G.Forster	N. Rønsted & S. Bruun-Lund 780 (C)	Australia, QLD	Palms Cove, Rocks at Port Douglas end of beach, near camping, side parking and bridge.
782	Malvanthera	Platypodeae	Eubracteatae	*Ficus triradiata* Corner	N. Rønsted & S. Bruun-Lund 782 (C)	Australia QLD	Mount Lewis.
783	Malvanthera	Malvantherae		*Ficus pleurocarpa* F. Mueller	N. Rønsted & S. Bruun-Lund 783 (C)	Australia, QLD	Mount Lewis.
784	Malvanthera	Platypodeae	Eubracteatae	*Ficus triradiata* Corner	N. Rønsted & S. Bruun-Lund 784 (C)	Australia, QLD	Mount Lewis.
785	Malvanthera	Platypodeae	Eubracteatae	*Ficus triradiata* Corner	N. Rønsted & S. Bruun-Lund 785 (C)	Australia, QLD	Mount Lewis.
786	Malvanthera	Platypodeae	Crassipes	*Ficus destruens*	N. Rønsted & S. Bruun-Lund 786 (C)	Australia, QLD	Mount Lewis.
787	Malvanthera	Platypodeae	Eubracteatae	*Ficus triradiata* Corner	N. Rønsted & S. Bruun-Lund 787 (C)	Australia, QLD	Mount Lewis. After logging area and exposed soil.
788	Malvanthera	Platypodeae	Crassipes	*Ficus crassipes* F.M. Bailey	N. Rønsted & S. Bruun-Lund 788 (C)	Australia, QLD	Mount Lewis. Brooklyn forrest part.
789	Malvanthera	Platypodeae	Crassipes	*Ficus crassipes* F.M. Bailey	N. Rønsted & S. Bruun-Lund 789 (C)	Australia, QLD	Mount Lewis. Just after entrance over bridge area
790	Malvanthera	Platypodeae	Rubiginosae	*Ficus brachypoda* (Miquel) Miquel	N. Rønsted & S. Bruun-Lund 790 (C)	Australia, NT	Lichtfield National Park
791	Malvanthera	Platypodeae	Rubiginosae	*Ficus brachypoda* (Miquel) Miquel	N. Rønsted & S. Bruun-Lund 791 (C)	Australia, NT	Lichtfield National Park, Circle walk Tolmer falls
792	Malvanthera	Platypodeae	Obliquae	*Ficus subpuberula* Corner	N. Rønsted & S. Bruun-Lund 792 (C)	Australia, NT	Lichtfield National Park, view platform along Tolmer falls.
794	Malvanthera	Platypodeae	Rubiginosae	*Ficus atricha* D.J. Dixon	N. Rønsted & S. Bruun-Lund 794 (C)	Australia, NT	Arnhem Road. A couple of kilometers on east side from the Adeleide river crossing.
797	Malvanthera	Platypodeae	Rubiginosae	*Ficus brachypoda* (Miquel) Miquel	N. Rønsted & S. Bruun-Lund 797 (C)	Australia, NT	Jesse gap, left side from parking area.
798	Malvanthera	Platypodeae	Rubiginosae	*Ficus brachypoda* (Miquel) Miquel	N. Rønsted & S. Bruun-Lund 798 (C)	Australia, NT	Alice springs. Ross Highway. Emily gap, left side from parking area.
799	Malvanthera	Platypodeae	Rubiginosae	*Ficus brachypoda* (Miquel) Miquel	N. Rønsted & S. Bruun-Lund 799 (C)	Australia, NT	Yulara, Desert Garden Hotel, planted in native garden next to reception
800	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus augusta* Corner	T. Kunda TK1 (C)	Papua New Guinea	Memeku -ME Camp

801	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus glandifera* Summerhayes	E.E. Henty, J. Lelean, NGF29474 (A)	Papua New Guinea	Walo, Commodore Bay, West New Britan
802	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus rhizophoriphylla* King	E. E. Henty (L)	Papua New Guinea	Kui, Morobe District, TNG
803	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus augusta* Corner	T. Kunda TK4, Bau 246 (C)	Papua New Guinea	Memeku
804	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus mafuluensis* Summerhayes	G.D. Weiblen 597 (A)	Papua New Guinea	nr. Kaionk village, Madang Province.
805	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus xylosycia* Diels	T. Kunda, TK6 (C)	Papua New Guinea	Numba-Mindik
806	Malvanthera	Hesperidiiformes	Hesperidiiformes	*Ficus hesperidiiformis* King	T. Kunda, TK7 (C)	Papua New Guinea	Numba Camp
807	Malvanthera	Platypodeae	Obliquae	*Ficus lilliputiana* D.J.Dixon	D. Dixon & I. Champion, PHD449 (JCT)	Australia, WA	Lake argyle, off access rad to dam, WA
808	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus augusta* Corner	T. Kunda TK9 (C)	Papua New Guinea	Memeku
809	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus xylosycia* Diels	T. Kunda, TK10 (C)	Papua New Guinea	Numba-Mindik
810	Malvanthera	Platypodeae	Obliquae	*Ficus lilliputiana* D.J.Dixon	D. Dixon & I. Champion, PHD420 (JCT)	Australia, WA	Hills to the E of Hidden Valley Caravan Park, Kununurra, WA
811	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus xylosycia* Diels	T. Kunda, TK12 (C)	Papua New Guinea	Numba-Mindik
812	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus xylosycia* Diels	T. Kunda, TK13, Bau 263 (C)	Papua New Guinea	Numba-Mindik, 589m NE from Numba 700m Station.
813	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus heteromeka* Corner	T. Kunda, TK14 (C)	Papua New Guinea	Memeku
814	Malvanthera	Platypodeae	Obliquae	*Ficus cerasicarpa* D.J.Dixon	D. Dixon & I. Champion, PHD401 (JCT)	Australia, WA	1km E of Cloncurry
815	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus xylosycia* Diels	T. Kunda, TK16 (C)	Papua New Guinea	Numba-Mindik
816	Malvanthera			*Ficus sp.*	T. Kunda, TK17 (C)	Papua New Guinea	Numba Camp
817	Malvanthera	Platypodeae	Obliquae	*Ficus cerasicarpa* D.J.Dixon	D. Dixon & I. Champion, PHD405 (JCT)	Australia, WA	Mount Isa Mines Weather Station
818	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus augusta* Corner	T. Kunda TK19, Bau 248 (C)	Papua New Guinea	Memeku
819	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus heteromeka* Corner	T. Kunda, TK20 (C)	Papua New Guinea	Memeku
820	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus glandifera* Summerhayes	D.B. Foreman, NGF45645(L)	Solomon islands	2,5 miles NE of Tonolei Harbour, Buin sub-district, Bougainville district.
821	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus heteromeka* Corner	T. Kunda, TK22 (C)	Papua New Guinea	Memeku
822	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus heteromeka* Corner	T. Kunda TK23, Bau249 (C)	Papua New Guinea	Memeku

823	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus glandifera* Summerhayes	M.J.S. Sands, 2060 (L)	Papua New Guinea	Namatanai sub-province; coastal region; near the shore. Ca 2 km south of Taron
824	Malvanthera	Hesperidiiformes	Hesperidiiformes	*Ficus sterrocarpa* Diels. *var. pubigemma*	T. Kunda, TK25 (C)	Papua New Guinea	Memeku
825	Malvanthera	Hesperidiiformes	Hesperidiiformes	*Ficus sterrocarpa* Diels. var. *pubigemma*	T. Kunda, TK26 (C)	Papua New Guinea	Numba Camp
826	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus glandifera* Summerhayes	P.F. Stevens (L)	Papua New Guinea	Matafuna Bay, Long Island. Track to crater late. Saidor Sub-dist. Madang Dist. TNG.
827	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus rhizophoriphylla* King	M. Jacobs, 9541 (L)	Papua New Guinea	Southeast of Lae on the coast, opposite Lasange I.
828	Malvanthera	Hesperidiiformes	Hesperidiiformes	*Ficus sterrocarpa* Diels	T. Kunda, TK29, Bau 261 (C)	Papua New Guinea	Numba-Mindik, 106m North of Numba 700m Station near Mindinoru Creek.
829	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus rhizophoriphylla* King	H. Streimann & P. Katik, NGF 28899 (L)	Papua New Guinea	Raba Raba subdist., Milne Bay Dist.
830	Malvanthera	Hesperidiiformes	Hesperidiiformes	*Ficus sterrocarpa* Diels. *var. pubigemma*	T. Kunda, TK31, Bau 251 (C)	Papua New Guinea	Memeku
831	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus rhizophoriphylla* King	E.E. Henty, NGF 28087 (A)	Papua New Guinea	Kui, Morobe dist. TNG
832	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus baola* C.C. Berg	T G Whitemore (L)	Papua New Guinea	N.E. San Crastobal, about 8 miles inland banks of Pegato river near confluence with Warahito
833	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus baola* C.C. Berg	T.G. Whitemore, BSIP2879 (L)	Solomon islands	New Georgia Group, Baga island.
834	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus baola* C.C. Berg	P. Runikera, BSIP12677 (L)	Solomon islands	Onibia area, N.W. San Cristolbal.
835	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus glandifera* Summerhayes	Y Lelean P. F. Stevens (L)	Papua New Guinea	Nuau, logging area, Hoskins Sub-dist. West new britan dist. TNG
836	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus glandifera* Summerhayes	R. Pullen (L)	Papua New Guinea	Nowata, C. 6 miles W. Of Rabaraba, Milne Bay District, E. Papua
837	Malvanthera	Hesperidiiformes	Glandiferae	*Ficus glandifera* Summerhayes	L. Lamothe, MELL03 (L)	Papua New Guinea	Near Manki 15, Bulolo, Morobe Province
838	Malvanthera	Hesperidiiformes	Hesperidiiformes	*Ficus sterrocarpa* Diels	T. Kunda, TK38, Bau 262 (C)	Papua New Guinea	Numba Camp, 758m North of Numba 700m Station.
839	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus xylosycia* Diels	T. Kunda, TK40, Bau 250 (C)	Papua New Guinea	Memeku

842	Malvanthera	Hesperidiiformes	Hesperidiiformes	*Ficus cf. sterrocarpa*	T. Kunda TK61, Bau 257 (C)	Papua New Guinea	Near Numba village
844	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus augusta* Corner	T. Kunda (C)	Papua New Guinea	Sinopas research project area, Numburabanna village
845	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus augusta* Corner	T. Kunda (C)	Papua New Guinea	Sinopas research project area, Mondinugara village
853	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus augusta* Corner	T. Kunda TK71 (C)	Papua New Guinea	Memeku
854	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus augusta* Corner	T. Kunda TK72 (C)	Papua New Guinea	Memeku
855	Malvanthera	Hesperidiiformes	Xylosyciae	*Ficus augusta* Corner	T. Kunda TK73 (C)	Papua New Guinea	Memeku
12 species sequence capture design set							
S1	Adenosperma			*Ficus pseudopalma* Blanco	Harrison 610 (PUH)	Philippines	
S5	Conosycea			*Ficus microcarpa* L.f.	Rønsted 77 (C)	Cult. P1979-5041 (C)	
S16	Americana			*Ficus americana* Aubl.	Rønsted 154 (BG)	Cult. BG1994-0678 (BG)	
S20	Malvanthera	Platypodeae	Rubiginosae	*Ficus rubiginosa* Desf. Ex Vent.	Rønsted 89 (C)	Cult. E1859-0014 (C)	
S30	Kissosycea			*Ficus. punctata* Thunb	Kjellberg FK-1997-23 (MPU)	Brunei	
S37	conosycea			*Ficus chartacea* Wall.	Chantarasuwan NR246 (THNHM)	Thailand	
S39	Oreosycea			*Ficus magnoliifolia* Blume	Harrison 619 (PUH)	Philippines	
S44	Rhizocladus			*Ficus sagittata* Vahl.	Rønsted 266 (HITBC)	Cult. (HITBC)	
S46	Paleomorphe			*Ficus tinctoria* Forst.f.	Rønsted 99 (K)	Cult. BG-89-551 (BG)	
S53	Galoglychia			*Ficus cyathistipula* Warb.	Rønsted 80 (C)	Cult. E1859-0023 (C)	
S68	Sycocarpus			*Ficus septica* Burm. F.	Harrison 618 (K)	Philippines	Palanan Forest Dynamics Plot, NE Luzon
S69	Sinosycidium			*Ficus tsiangii* Merr. ex Corner	Rønsted 298 (HITBC)	Cult. (HITBC)	

Appendix I

Identification of common horsetail (*Equisetum arvense* L.; Equisetaceae) using Thin Layer Chromatography versus DNA barcoding

SCIENTIFIC REPORTS

OPEN

Received: 20 March 2015
Accepted: 11 May 2015
Published: 13 July 2015

Identification of common horsetail (*Equisetum arvense* L.; Equisetaceae) using Thin Layer Chromatography versus DNA barcoding

C. Haris Saslis-Lagoudakis[1], Sam Bruun-Lund[1], Natalie E. Iwanycki[1], Ole Seberg[1], Gitte Petersen[1], Anna K. Jäger[2] & Nina Rønsted[1]

The global herbal products market has grown in recent years, making regulation of these products paramount for public healthcare. For instance, the common horsetail (*Equisetum arvense* L.) is used in numerous herbal products, but it can be adulterated with closely related species, especially *E. palustre* L. that can produce toxic alkaloids. As morphology-based identification is often difficult or impossible, the identification of processed material can be aided by molecular techniques. In this study, we explore two molecular identification techniques as methods of testing the purity of these products: a Thin Layer Chromatography approach (TLC-test) included in the European Pharmacopoeia and a DNA barcoding approach, used in recent years to identify material in herbal products. We test the potential of these methods for distinguishing and identifying these species using material from herbarium collections and commercial herbal products. We find that both methods can discriminate between the two species and positively identify *E. arvense*. The TLC-test is more cost- and time-efficient, but DNA barcoding is more powerful in determining the identity of adulterant species. Our study shows that, although DNA barcoding presents certain advantages, other established laboratory methods can perform as well or even better in confirming species' identity in herbal products.

Tens of thousands of plant species are used medicinally[1] and a substantial portion of the world's population depends on traditional medicine[2]. In recent decades, public interest in herbal products has grown[3–5], but these products are not always regulated. The safety of herbal products can be compromised through accidental adulteration, misidentification and deliberate contamination[6,7], which can lead to severe side effects due to the presence of toxic compounds[8]. This creates a need for authentication of species included in these products. The qualitative and quantitative composition of herbal products is regulated by international and national monographs such as the European Pharmacopoeia[9], which presents a series of monographs for herbal products, including recommended tests for identification and quality. These tests are often based on morphology. However, macroscopic or microscopic identification of plant species requires considerable expertise to differentiate between closely related or similar looking species. Furthermore, morphological characters may be indistinguishable in bulk, pulverised or otherwise processed material[10,11].

[1]Evolutionary Genomics Section, Natural History Museum of Denmark, Sølvgade 83S, Copenhagen, DK-1307, Denmark. [2]Natural Products Research, Department of Drug Design and Pharmacology, University of Copenhagen, Universitetsparken 2, Copenhagen, DK-2100, Denmark. Correspondence and requests for materials should be addressed to C.H.S.L. (email: c.h.saslislagoudakis@gmail.com) or N.R. (email: nronsted@snm.ku.dk)

To circumvent these problems, most monographs define a maximum allowance of foreign matter often based on a Thin Layer Chromatographic (TLC) test using chemical markers allowing distinction between the correct species and other, potentially toxic species[12,13]. However, such chemical markers or fingerprinting analyses have certain drawbacks. First, it is often difficult to find chemical markers that are unique to the target species. Different species can produce the same marker, hindering species' identification. Second, chemical composition can demonstrate considerable intraspecific variability depending on season, growth, storage conditions and harvesting process[14]. Third, herbal products are sometimes spiked with synthesised compounds[15]. In these cases, the TLC-test may lead to false species' identification.

An alternative method that has been used to identify components of herbal products is DNA barcoding[10,11,16–19]. DNA barcoding relies on sequencing of short fragments of the genome, which are unique to the target species[20]. The DNA sequences from the product are compared to a reference database, based on which the identity of the species can be confirmed[21,22]. DNA-based identification methods have often revealed adulteration in traditional medicinal preparations and herbal products. For example, potentially toxic *Ephedra* L. and *Asarum* L. material was found in Traditional Chinese Medicinal products administered in Australia[23], and several adulterant plant species were found in herbal products from North America[17]. Nevertheless, DNA barcoding also has limitations. First, depending on the condition of the plant material, amplification of the target DNA marker may not be possible. Second, DNA barcodes might show low interspecific variability, particularly among closely related species. Finally, because DNA barcoding relies on the presence of a reference database, the absence of a species from the database will impede its identification success[19]. Despite its limitations, DNA barcoding has often been discussed as the primary method of molecular identification of plants in the last decade[11,16,22].

In this study, we explore molecular identification of the genus *Equisetum* L. (Equisetaceae), also known as horsetails. The genus comprises 15 species and has a more or less cosmopolitan distribution[24,25]. *Equisetum arvense* L. is used traditionally against numerous conditions[26] and many *E. arvense* herbal products are sold on the market mainly against urinary and renal conditions[27], as well as skin, hair and nail remedies, potentially due to the species' high silica content[28]. The separation of *E. arvense* from other *Equisetum* species – especially *E. palustre* L. that contains toxic levels of the pyridine alkaloid palustrine – is challenging[29,30], particularly based on microscopic examination of commercial herbal products. Therefore, the European Pharmacopoeia monograph for the common or field horsetail, *E. arvense,* includes a TLC-test (Identification C) that tests for its positive identification, including a test for foreign matter from *E. arvense*. However, it is not clear whether this test can positively identify either *E. arvense* or *E. palustre* among other morphologically similar *Equisetum* species, several of which overlap geographically with *E. arvense*[31]. This is a potential problem because palustrine is not specific to *E. palustre*, but it is found in other horsetail species. An early study detected palustrine in *E. arvense* and *E. hyemale* L.[32]. A later study did not detect it in *E. arvense*, *E. telmateia* Ehrh., and *E. sylvaticum* L.[33], but a more recent compendium of poisonous plants cites palustrine and palustridine alkaloid content for *E. fluviatile* L., *E. hyemale*, *E. palustre*, *E. sylvaticum*, and *E. telmateia*[34].

The main objective of this study was to investigate the resolution power of the European Pharmacopoeia's TLC-test and of the DNA barcoding approach for i) distinguishing between *E. arvense* and *E. palustre* and ii) positively identifying these two species and discriminating them from other *Equisetum* species. In order to perform these investigations, we needed to have a reliable species' delimitation. Therefore, we also reconstructed a molecular phylogeny of *Equisetum* to test currently accepted species boundaries. Our study is based on herbarium collections of wild origin, as well as exemplar herbal products from the market.

Results

Phylogeny of *Equisetum*. We reconstructed a phylogenetic tree of *Equisetum* (Fig. 1) in order to test the monophyly of the species. Previous studies have provided phylogenetic hypotheses for the genus using plastid DNA markers[35–37], but these studies only included one specimen per species. The topology obtained here from nuclear and plastid markers, and including several accessions per species, largely corresponds to the topology found previously[35–37]. *Equisetum bogotense* Kunth is recovered as sister to the rest of the genus and not as a member of subg. *Hippochaete* (Milde) Baker. The remainder of the genus is resolved into two major clades, each comprising seven species and corresponding to the two subgenera *Equisetum* and *Hippochaete* (Fig. 1). With the exception of *E. diffusum* D. Don and *E. sylvaticum*, all species were recovered as monophyletic, including *E. arvense* and *E. palustre*. These two species are resolved in the same clade (subg. *Equisetum*), but not as sister species (Fig. 1).

Distinction between *E. arvense* and *E. palustre*.

Chemistry. The distinction between the two species based on the TLC-test of the European Pharmacopoeia is based on the presence of a combination of marker bands in each species, shown in Fig. 2. The results of the TLC-test recommended by the European Pharmacopoeia are shown in Fig. 3 for the *E. arvense* - *E. palustre* comparison. The two bands at the bottom of the plate that are used for the identification of *E. palustre* are present in all accessions of this species, but not in any of the *E. arvense* accessions. Although some of the marker bands used to identify *E. arvense* can be found in *E. palustre* accessions, the combination of the four marker bands (Fig. 2) is not seen in any *E. palustre* accessions

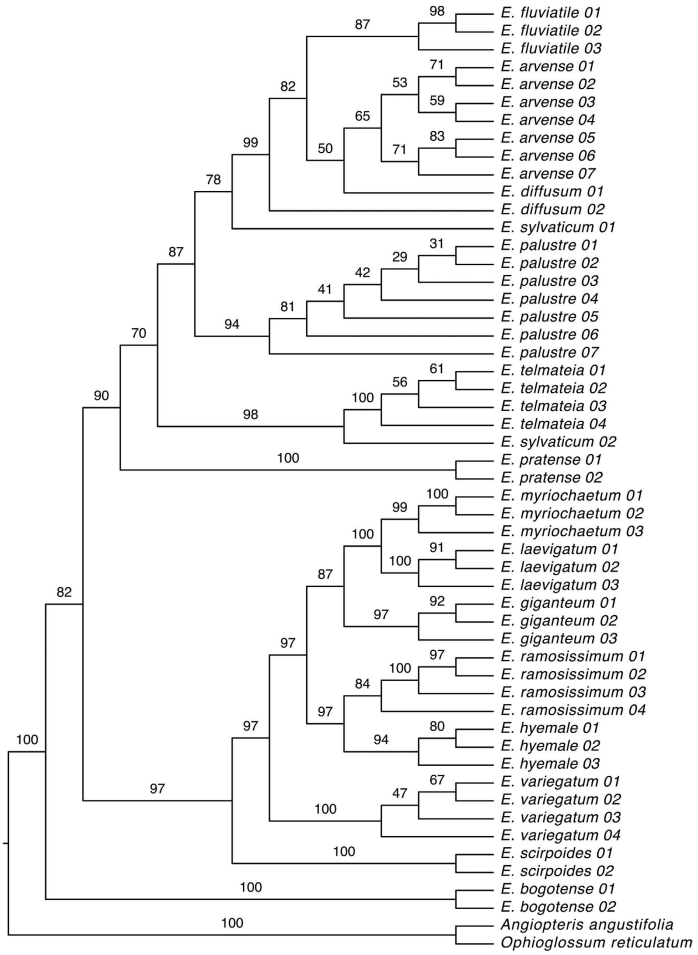


Figure 1. Phylogeny of *Equisetum* reconstructed with a Maximum Likelihood analysis based on five DNA markers (ITS2, *matK, rbcL, rps4, trnH-psbA*). Bootstrap support values are given above respective branches.

(Fig. 3). Therefore, the marker zones used as the distinguishing characters between the two species in the monograph (Fig. 2) could consistently distinguish between *E. arvense* and *E. palustre*. We observed a typical *E. arvense* TLC chromatogram for five out of eight commercial products included in the analysis (Fig. 4). In one product (B – Bulgaria), we observed the marker bands that are used to identify both *E. arvense* and *E. palustre* in the TLC-test of the European Pharmacopoeia, suggesting this product includes a mixture of the two species (Fig. 4). One product (I – UK) seemed to not contain any *Equisetum* material at all and another one (HB – UK) returned no chromatogram (Fig. 4).

DNA barcoding. The two plastid markers we used for the DNA barcoding of *E. arvense* and *E. palustre* resolve the samples into two well supported, monophyletic clades, shown in Fig. 5. We were able to amplify DNA from two of the herbal products, only; one was resolved within the *E. arvense* clade (BP 100). The other, which was shown to be a mixture from the TLC-test (B – Bulgaria), is recovered with the *E. palustre* (BP 77) clade (Fig. 5). For both barcoding regions, we found 36 substitutions (25 for *matK* and 11 for *trnH-psbA*) that can distinguish *E. arvense* and *E. palustre*. Some of them are unique to each species and others are shared with other species, but not between *E. arvense* and *E. palustre* (Table 1).

Positive identification of *E. arvense* and *E. palustre*.

Chemistry. We analysed one exemplar specimen of all *Equisetum* species using the TLC-test recommended by the European Pharmacopoeia (Fig. 6). For *E. diffusum* and *E. sylvaticum*, which were not monophyletic in the DNA analysis, we could test only one sample, as the other sample did not come

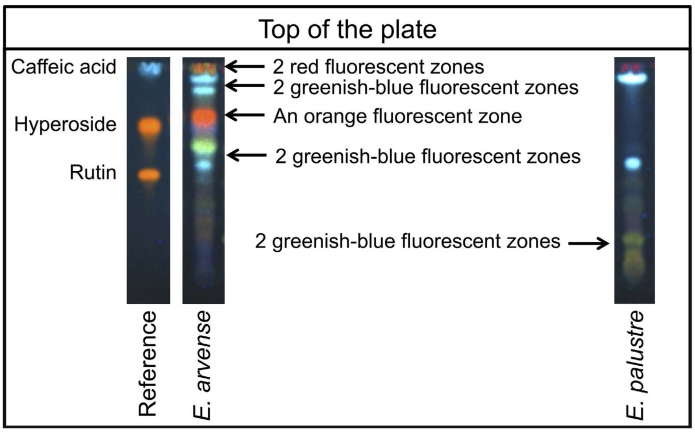


Figure 2. Exemplar chromatograms of *E. arvense* and *E. palustre* pointing out the combination of characters used in the European Pharmacopoeia to identify the species (four for *E. arvense* and one for *E. palustre*). Although some markers are not unique to *E. arvense*, the combination of all four traits serves for its positive identification. The reference solution is also presented.

from our study. The TLC-test (Identification C) of the European Pharmacopoeia can positively identify *E. arvense*. Although some of the marker bands outlined in the TLC identification test for *E. arvense* (Fig. 2) are seen in the chromatograms of other *Equisetum* species, *E. arvense* is the only species with the combination of all these markers bands (Fig. 6). As shown in Fig. 6, one or two of the greenish-blue fluorescent zones used in the TLC-test to detect *E. palustre* were not detected in any other species within subg. *Equisetum*, but were present in all species in subg. *Hippochaete*. Therefore, the TLC-test of the European Pharmacopoeia cannot be used to identify *E. palustre*, because the trait of this species (Fig. 2) is shared with other *Equisetum* species as well (Fig. 6).

DNA barcoding. We investigated whether the two DNA markers we used as barcodes can not only differentiate *E. arvense* from *E. palustre*, but also include a combination of unique traits for these species, which can be used to successfully identify them from all other *Equisetum* species. Table 1 shows that there are unique substitutions in these two markers, the combination of which can positively identify both *E. arvense* and *E. palustre* from other *Equisetum* species. For *matK*, we found no substitutions to be unique for *E. arvense*, but five substitutions were unique for *E. palustre* (Table 1). For *trnH-psbA*, one substitution was unique for *E. arvense* and two for *E. palustre* (Table 1). Regarding the identification of material in herbal products, DNA sequences from one product (F – Germany) show the combination of characters that can identify *E. arvense*. For the other product (B – Bulgaria), we only managed to amplify *matK*. This sequence is actually a chimeric sequence (several doubles peaks are observed in the DNA chromatograms), showing some characters that are characteristic of *E. arvense* and some of *E. palustre*.

Discussion

Some *Equisetum* species are morphologically quite variable and can be difficult to identify based on morphology alone[29,30]. To the untrained eye, *E. arvense* may superficially resemble other species within subgenus *Equisetum*, including *E. palustre*, as well as *E. pratense*, *E. fluviatile*, *E. telmateia* and *E. diffusum*. Positive identification of material lacking strobili, or where information about dimorphism is lacking, may be challenging even for trained botanists, as micro-morphological or anatomical characteristics may be required to separate some species, e.g. *E. arvense* and *E. palustre*[38]. Within their respective ranges, taxa sharing similar morphological characters, such as *E. arvense* and *E. palustre*, may be found co-occurring in the same habitat[31]. A further complication to field-identification is that *E. arvense* is known to form hybrids with *E. palustre* (*E.* × *rothmaleri* C.N. Page) and *E. fluviatile* (*E*×*littorale* Rupr.)[39,40], with morphological and chemical traits that are intermediate between the parent taxa[30,40].

Due to the risk of misidentification or adulteration of *E. arvense* with *E. palustre*, laboratory techniques are needed for the quality control of herbal products of *E. arvense*. The European Pharmacopoeia has devised a simple method using TLC (Identification C) to distinguish the two species[9], and we found this test to be straightforward and consistent. It can confirm that the material is from *E. arvense*, through a combination of marker bands unique to this species (Figs 2,3 and 6). Further, as shown in Fig. 3, the two greenish-blue bands at the bottom are present in all *E. palustre* accessions, but none of the *E. arvense* accessions. The presence of these bands can be used as an indication of adulteration with *E. palustre*, but the identity of the adulterant is not confident, because these bands are also found in other *Equisetum* species besides *E. palustre* (Fig. 6). Also, even in the case of absence of these bands, a partial adulteration

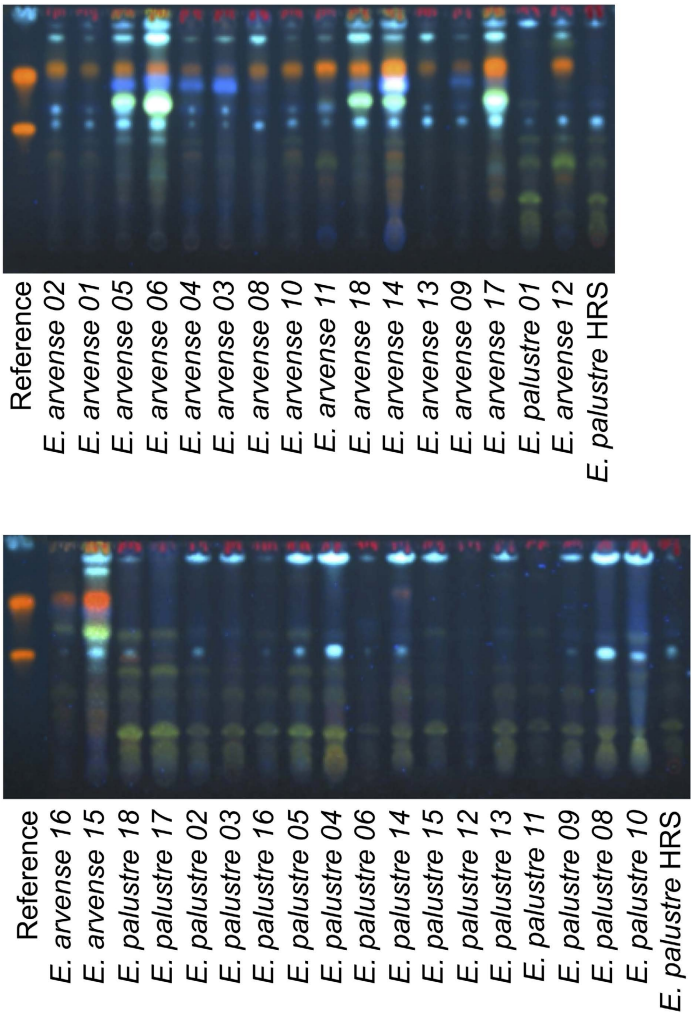


Figure 3. TLC chromatogram of *Equisetum arvense* and *E. palustre* accessions from natural history collections.

with another *Equisetum* species that does not demonstrate them in the chromatogram (Fig. 6) cannot be ruled out.

The current TLC-test is testing for the presence of kaempferol glucosides (flavonoids), instead of directly testing for the presence of alkaloids. We tested for alkaloids using the material which had already been extracted for the flavonoid analysis. This method could only detect alkaloid bands (two bands) present in the reference *E. palustre* HRS and the *E. palustre* accession used on the TLC-test across the genus, whereas possible alkaloids present in other *Equisetum* species were below the detection limit of this method (results not shown). We suggest that a method testing directly for alkaloids be developed and included in the monograph.

DNA barcoding may be an alternative or supplementary method to identify material in herbal products with higher certainty[17–19,41]. We found that two plastid markers can successfully distinguish between *E. arvense* and *E. palustre*. In total, there are 36 characters (25 for *matK* and 11 for *trnH-psbA*) differentiating the two species (Table 1), and the phylogenetic analysis of these two DNA barcoding markers assigns material from these two species to two well-supported clades (Fig. 5). Further, this approach can positively identify the two species, as we found six substitutions (five for *matK* and one for *trnH-psbA*) that are unique to *E. palustre* (Table 1), allowing high confidence in the identification of this species. For *E. arvense*, we only found one unique substitution in *trnH-psbA* and none in *matK* (Table 1), making assignment of material to this species less robust. However, a number of other substitutions are only shared by *E. arvense* and its two closest relatives, *E. fluvatile* and *E. diffusum* (Fig. 1), which co-occur in Asia. Including more DNA barcoding regions that have been proposed by the Consortium for the

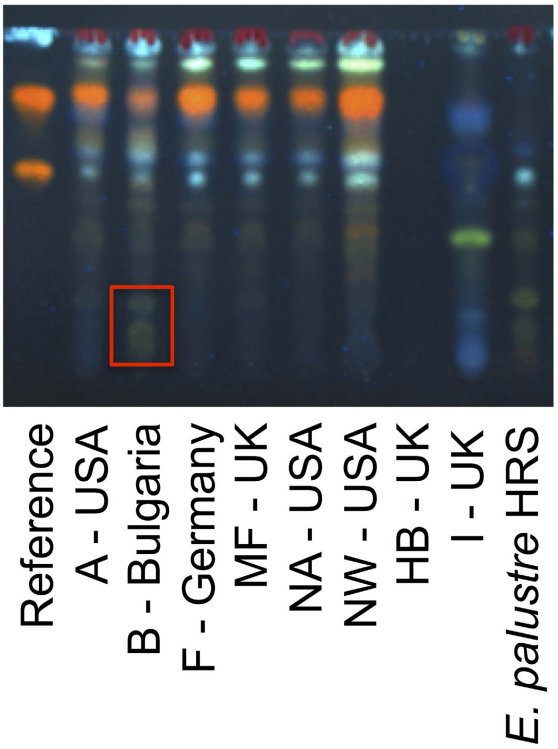


Figure 4. TLC chromatogram of commercial products sold as *Equisetum arvense*. We only provide the acronym of each product and its country of manufacture. The distinctive greenish-blue band area that indicates presence of *E. palustre* material is highlighted inside the red rectangle.

Barcode of Life Plant Working Group[22] could provide further discriminatory power for *E. arvense*. However, we found *rbcL* to show too little interspecific variation, and ITS2, which has been proposed for the DNA barcoding of medicinal plants[16], did not amplify consistently in *Equisetum*. Other DNA barcoding markers that have been shown recently to perform better than the ones we used here [e.g., *ycf1*[42]] could provide more species-specific substitutions in future investigations.

We included eight commercial products claiming to be *E. arvense*, seven of which produced TLC chromatograms allowing assignment of the herbal product to either *E. arvense* or *E. palustre* following the European Pharmacopoeia's TLC-test for foreign matter. Of these seven products, five were assigned to *E. arvense* (Fig. 4). We were only able to gather DNA sequence data for one of these samples (herbal product F - Germany), and it was confirmed to be *E. arvense* (Fig. 5). For one product (herbal product B - Bulgaria), the TLC-test showed the presence of *E. arvense* and potentially *E. palustre* material (Fig. 4). The DNA sequence data confirmed that this product is most likely a mixture, as the resulting sequence was chimeric. Although this product is resolved within the *E. palustre* clade (Fig. 5), the sequence we amplified shows a combination of substitutions characteristic of *E. arvense* and *E. palustre*. It could be adulterated, misidentified or even be of hybrid origin. The product is a tea from south-eastern Europe, an area that is a major source of commercial *E. arvense* products[27], and where the two species co-occur, raising concerns about the risk of contamination with *E. palustre* in commercially available material presumed to be *E. arvense*. Finally, one sample (herbal product I - UK) produced a chromatogram that was different from those characteristic of any *Equisetum* species (Figs 4 and 6), suggesting the botanical material in that sample might not be *Equisetum*. Unfortunately, no DNA sequence data could be gathered from that sample.

Our objective was to explore and compare the power of the European Pharmacopoeia's TLC-test and of the DNA barcoding approach for distinguishing between *E. arvense* and *E. palustre*, as well as for positively identifying the two species. We found both methods to be useful, however with different advantages and shortcomings. In terms of success rate of data collection, the TLC-test approach is more efficient. First and foremost, the laboratory work is less laborious and cheaper than DNA barcoding. Second, the TLC-test had a greater success rate with commercial herbal products: we obtained chromatograms for seven out of eight of these products, while the amplification success of the barcoding regions from these products was limited (only two samples). On the other hand, in terms of resolution and confidence in identification, the DNA barcoding approach is better. Although both methods can

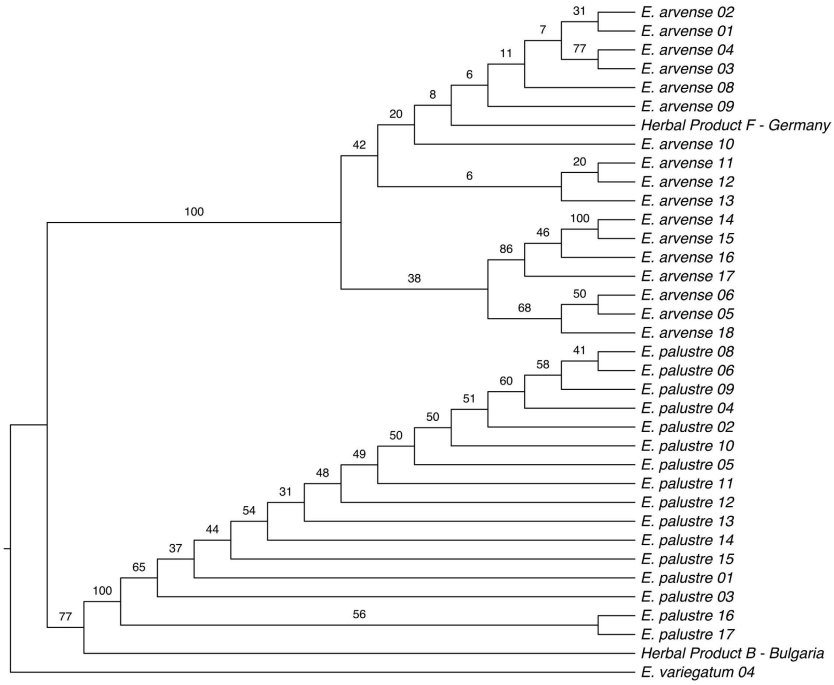


Figure 5. DNA barcoding of *Equisetum arvense* and *E. palustre* based on two markers (*matK* & *trnH-psbA*). The phylogenetic tree was reconstructed with a Maximum Likelihood analysis based, using *E. variegatum* as an outgroup. Bootstrap support values are given above respective branches.

successfully discriminate between *E. arvense* and *E. palustre* and positively identify *E. arvense*, only DNA barcoding provides a combination of traits that is unique to *E. palustre* among horsetail species. However, the amplification of these barcoding markers might prove difficult in processed commercial products. Additionally, an advantage of the TLC-test is that contamination can be quantified based on the level of visibility of the greenish-blue bands on chromatograms, an aspect in which the DNA barcoding approach lacks.

Which method would we recommend as being the best? Given the pros and cons of each method, we believe that it depends on the application. Our results show that, when it comes to confirming whether an herbal product contains *E. arvense,* the TLC-test is the most cost- and time-efficient option. However, the presence of the marker bands described in the TLC-test as characteristic of *E. palustre* can be seen in cases of adulteration with other *Equisetum* species, as these bands are common within the genus (Fig. 6). Similarly, the absence of these marker bands does not guarantee that the product has not been adulterated with other *Equisetum* species, which do not show those bands (Fig. 6). Given that there is uncertainty about which *Equisetum* species produce toxic alkaloids, this could be an important shortcoming of the TLC-test. In these cases, DNA barcoding can be used as a complementary test for quality control, when possible.

Our study also highlights the immense potential of herbarium collections for a wide range of modern approaches to biodiversity research[43,44], and DNA barcoding in particular[45,46]. The majority of the material used in this study was obtained from the collections in herbarium of the Natural History Museum of Denmark (C). The age of the material ranged from 1900–2013. We did not detect any apparent age-related difference in the intensity of the TLC chromatograms or the amplification success of the DNA markers, showing that the chemical profiles and the DNA are not substantially degraded in carefully stored collections[47]. Our findings demonstrate how available collections can be used to set up a modern framework of chemical and molecular identification of economically important species. Without conducting substantial fieldwork, we managed to sample across all *Equisetum* species, as well as within *E. arvense* and *E. palustre*, covering their geographic ranges, hence ensuring that both inter- and intra-specific variation is covered. An incidental advantage of using herbarium material is that the link between the chemical and molecular data and the voucher is established by default. Missing vouchers is a serious problem in many studies[41] which makes replication by future researchers almost impossible[48].

matK	*E. arvense* unique substitutions	none
	E. arvense substitutions shared with other *Equisetum* species but not with *E. palustre*	With *E. fluvatile*: 265 C, 511 G With *E. diffusum, E. fluvatile*: 294 A, 344 C 351 C, 374 C, 397 G, 502 T, 512 C, 567 G, 572 A, 619 T, 636 A, 670 G. With *E. diffusum, E. fluviatile, E. sylvaticum*: 212 T, 462 T. With *E. diffusum, E. fluviatile, E. hyemale, E. sylvaticum*: 372 G.
	E. palustre unique substitutions	273 C, 311 A, 391 C, 575 C, 594 C.
	E. palustre substitutions shared with other *Equisetum* species but not with *E. arvense*	With *E. laevigatum, E. myriochaetum*: 600 G. With *E. sylvaticum, E. telmateia*: 534 C. With *E. hyemale, E. ramosissimum*: 421 A.
trnH-psbA	*E. arvense* unique substitutions	193 A
	E. arvense substitutions shared with other *Equisetum* species but not with *E. palustre*	With *E. diffusum, E. fluviatile*: 141 A, 87 A.
	E. palustre unique substitutions	124 A, 200 A.
	E. palustre substitutions shared with other *Equisetum* species but not with *E. arvense*	With *E. diffusum, E. bogotense, E. sylvaticum*: 121 G, 122 T, 123 A, 125 T, 126 A. With *E. pratense, E. sylvaticum, E. telmateia*: 151 C.

Table 1. Distinguishing characters between *Equisetum arvense* and *E. palustre* in *matK* and *trnH-psbA* barcodes. The numbers and substitutions refer to positions in the alignment presented in the Supplementary Information.

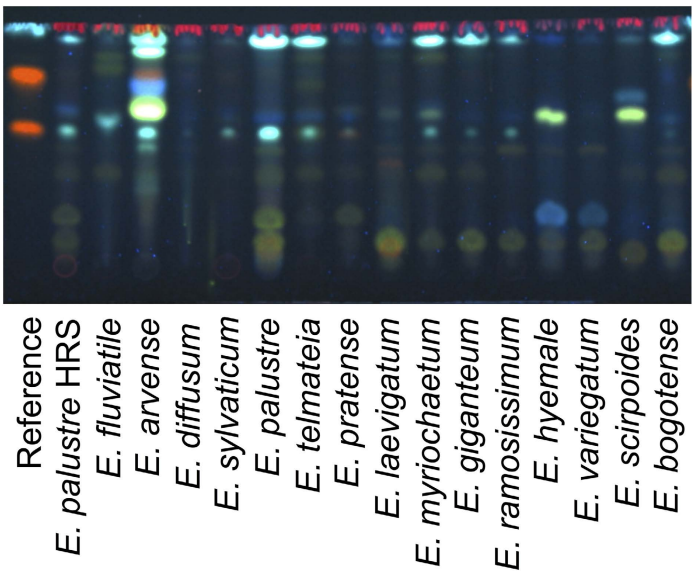


Figure 6. TLC chromatogram of exemplar accessions of all *Equisetum* species.

Conclusions

Given the recent growth of the herbal products market[3–5], efficient methods for regulating these products against accidental adulteration, deliberate contamination and misidentification are more relevant than ever for public healthcare[41]. We tested the European Pharmacopoeia's *E. arvense* TLC-test for foreign matter, particularly from the closely related *E. palustre*. We also tested a DNA barcoding approach to distinguish and identify these species. We found that each method has advantages and disadvantages, but the TLC-test is the most efficient way of confirming that material in herbal products is indeed *E. arvense*.

On the other hand, the DNA barcoding can be used as a complementary test to determine the identity of adulterant species, particularly *E. palustre*.

Future work can focus on systematically studying which *Equisetum* species produce toxic alkaloids, which will assist the quality control of *E. arvense* herbal products. Further, a chemical method that directly tests for the presence of alkaloids in herbal products can circumvent problems in species identification, directly testing for the quality and appropriateness for human consumption of herbal products. Additionally, the steadily dropping price of next generation sequencing techniques – which massively amplify short DNA fragments – may considerably enhance the success rates of DNA barcoding in degraded or processed material. Finally, given the presence of several putative hybrids between *E. arvense* and other *Equisetum* species, further techniques can be applied to investigate the presence of hybrid material in herbal products.

Methods

Plant material. For the phylogenetic reconstruction, we sampled at least one accession of each *Equisetum* species, mostly from material deposited in the herbarium of the Natural History Museum of Denmark (C), in order to produce a well-sampled phylogenetic hypothesis for the genus. From these specimens, we chose one per species for the TLC-test across *Equisetum* species. For the DNA barcoding and TLC-test of *E. arvense* and *E. palustre*, we sampled several accessions of each of the two species covering their distribution ranges to the extent possible. Additionally, we sampled eight herbal products sold on the market as *E. arvense*. Details of plant materials are listed in Supplementary Tables 1 and 2.

DNA sequencing. Complete genomic DNA was extracted using the DNeasy Mini Plant Kit (Qiagen Ltd, Crawley, UK), following the manufacturers protocol. For the DNA barcoding of *E. arvense* and *E. palustre*, we sequenced the *trnH-psbA* spacer and the barcoding fragment of *matK*, which have been used in previous DNA barcoding studies[21,22,49,50]. For the genus wide analysis, we sequenced the plastid regions *rps4*, *rbcL*, the barcoding fragment of *matK*, the *trnH-psbA* spacer, and the nuclear ribosomal *ITS2* region. The *rps4* marker was amplified using primers rps5 (5′-ATG TCC CGT TAT CGA GGA CC T-3) and trnS (5′-TAC CGA GGG TTC GAA TC-3)[51,52] and the *rbcL* marker was amplified with primers rbcL26F (5′-ATG TCA CCA CAA ACA GAA ACT AAA GCA AGT-3′) and rbcL1379R (5′-TCA CAA GCA GCA GCT AGT TCA GAA CTC-3′)[53]. For both these markers, we used the following PCR programme: 3 minutes of initial denaturation at 94 °C, followed by 30 cycles of 45 seconds at 94 °C, 45 seconds at 53 °C, 90 seconds at 72 °C, and a final extension for 10 minutes at 72 °C. For the *trnH-psbA* spacer region a PCR was performed using primers trnHf (5′-CGC GCA TGG TGG ATT CAC AAT CC-3′) and psbA3f (5′-GTT ATG CAT GAA CGT AAT GCT C-3′)[54,55] using the following conditions: 4 minutes at 95 °C, followed by 48 cycles of 30 seconds at 94 °C, 40 seconds at 45 °C, 40 seconds at 72 °C, and a final extension for 5 minutes at 72 °C. For *matK*, *Equisetum* specific primers were used: matK Equisetum F (5′-ATA CCC CAT TTT ATT CAT CC-3′) and matK Equisetum R (5′-GTA CTT TTA TGT TTA CGA GC-3′) [http://www.kew.org/barcoding/update.html] with the following conditions: 4 minutes at 94 °C following 32 cycles of 1 minute at 94 °C, 1 minute at 46 °C, 2:30 minutes at 72 °C and a final extension for 7 minutes at 72 °C. Part of the internal transcribed spacer region (ITS2) was amplified using primers ITS3 (5′-GCA TCG ATG AAG AAC GCA GC-3′) and ITS4 (5′-TCC TCC GCT TAT TGA TAT GC-3′) from White *et al.*[56]. With the following conditions: 4 minutes at 94 °C following 35 cycles of 1 minute at 94 °C, 1 minute at 48 °C, 1 minute at 72 °C and a final extension for 2 minutes at 72 °C.

Reactions of 25 uL were carried out using standard procedures with 1 or 2 uL DNA template. Moreover, for *matK* and ITS DMSO was added to reduce the effects of secondary structure on primer biding. BSA was added to all reactions to enhance polymerase activity. The PCR products were purified using the Qiagen PCR purification kit (Qiagen Inc.) according to the manufacturer's instructions. Direct sequencing of purified PCR products was either performed using BIGDYE v1.1 (Applied Biosystems, Wellesley, Massachusetts, U.S.A.) and purified sequencing products were run on an AB3130 × 1 automated sequencer (Applied Biosystems) or sent to GATC-biotech in Germany (http://www.gatc-biotech.com). Forward and reverse sequences were edited and assembled in Geneious v. 7.1.7 (http://www.biomatters.com). Alignments were conducted using the MAFFT v.7 plugin[57] in Geneious with default options and inspected manually afterwards. Regions that were ambiguously aligned were excluded from the analyses. Genbank accession numbers for all sequences used in the study are shown in Supplementary Table 3

Phylogenetic methods. Two matrices were assembled: (1) a genus wide matrix combining our datasets with DNA sequences from previous phylogenetic studies of *Equisetum*[35–37] to achieve a sampling scheme of multiple accessions per taxon allowing test of species monophyly, and (2) an *Equisetum arvense-E. palustre* dataset for the development of the DNA barcoding methodology. All sequences were aligned with MAFFT[57] and sequence data were analysed under the Maximum Likelihood (ML) criterion, with RAxML[58] using the partitioned model option (five partitions – one per DNA marker) with the GTR+I+G model and running 100 bootstrap replicates[59]. *Angiopteris angustifolia* and *Ophioglossum reticulatum* were used as outgroup for the phylogenetic analysis of the genus, and *E. variegatum* was used as outgroup for the phylogenetic analysis of DNA barcodes.

Thin Layer Chromatography. Reference and test solutions of the plant material was prepared following the European Pharmacopoeia 7.4 monograph for *Equisetum* stem (Equiseti herba) test for foreign matter[9]. Due to the limited availability of material from the herbarium specimens in general, only about 20–50 mg of powdered stem was extracted and the amount of methanol used adjusted accordingly. For the test solutions, powdered *Equisetum* stems were extracted with methanol R (VWR BDH Prolabo Chemicals) in the ratio 100 mg/mL. The mixture was heated in a water-bath at 60 °C for 10 min with occasional shaking, allowed to cool and then filtered. The reference solution (a) of *Equisetum palustre* HRS (European Directorate for the Quality of Medicines) was prepared in the same way as the test solutions. Another reference solution (b) was made by dissolving 1.0 mg of caffeic acid R (Sigma), 2.5 mg of hyperoside R (Roth) and 2.5 mg of rutin R (Sigma) in 20 mL of methanol R. For commercial products, 1 g material was extracted with 10 mL methanol R in the same way as the test extracts.

2 μl bands of 8 mm of each solution were applied with a GAMAC nanomat 4 to HPTLC silica gel plates R (5–6 μm; Merck). HPTLC plates were developed over a path of 6 cm using a mobile phase consisting of anhydrous formic acid R (Emsure), glacial acetic acid R (Merck), water R, and ethyl acetate R (Sigma Aldrich) (7.5:7.5:18:67 V/V/V/V). After development, plates were air-dried for 5 min. Detection was achieved by heating at 100 °C for 3 min followed by treatment of the still warm plate with a 10 g/L solution of diphenylboric acid aminoethyl ester R (Roth) in methanol R, and then treatment with a 50 g/L solution of macrogol 400 R in methanol R. Finally plates were air-dried and examined after 10 min in ultraviolet light at 365 nm. System suitability was observed by the appearance of two greenish-blue fluorescent zones from kaempferol glucosides (flavonoids) characteristic of *E. palustre* L. in the reference solution (a) just above the line of application. In the chromatogram of the test solution any greenish fluorescent zones just above the line of application may not be more intense than the corresponding zones (characteristic for *E. palustre*) in chromatogram of the reference solution.

For alkaloid detection, the dried, powdered stem material, which had already been extracted for flavonoid-analysis, was moistened with 10% (1 μL/μg dry plant material). 1 ml dichloromethane (VWR BDH Prolabo) was added, and the mixture was extracted for 24 h at room temperature. 900 μL of the liquid was taken to dryness. The extract was redissolved in 20 μL dichloromethane and applied to a Merck Silica gel 60 F254 TLC plate and eluted in toluene:ethyl acetate:diethylamine (VWR BDH Prolabo; Sigma; Merck) 7:2:1 over 7 cm. 1 mg/mL brucin was used as positive control. The plate was sprayed with 0.15% chloroplatinic acid hydrate (Sigma Aldrich) in a 3% KI solution.

References

1. Schippmann, U., Leaman, D. J. & Cunningham, A. B. in Biodiversity and the Ecosystem Approach in Agriculture, Forestry and Fisheries (Food and Agriculture Organization, 2002).
2. World Health Organization. *Fact sheet N°* **134,** (2003). Available at: http://www.who.int/mediacentre/factsheets/2003/fs134/en/ (Accessed: 19th February 2015).
3. Kennedy, J. Herb and supplement use in the US adult population. *Clinical Therapeutics* **27,** 1847–1858 doi: http://dx.doi.org/10.1016/j.clinthera.2005.11.004 (2005).
4. Marinac, J. S. *et al.* Herbal Products and Dietary Supplements: A Survey of Use, Attitudes, and Knowledge Among Older Adults. *JAOA: Journal of the American Osteopathic Association* **107,** 13–23 (2007).
5. Jordan, S. A., Cunningham, D. G. & Marles, R. J. Assessment of herbal medicinal products: Challenges, and opportunities to increase the knowledge base for safety assessment. *Toxicology and Applied Pharmacology* **243,** 198–216 doi: http://dx.doi.org/10.1016/j.taap.2009.12.005 (2010).
6. Ernst, E. Risks of herbal medicinal products. *Pharmacoepidemiology and Drug Safety* **13,** 767–771 doi: 10.1002/pds.1014 (2004).
7. Van Breemen, R. B., Fong, H. H. S. & Farnsworth, N. R. Ensuring the safety of botanical dietary supplements. *The American Journal of Clinical Nutrition* **87,** 509S–513S (2008).
8. Gilbert, N. Regulations: Herbal medicine rule book. *Nature* **480,** S98–S99 doi: 10.1038/480S97a (2011).
9. Council of Europe. European Pharmacopoeia *8th Edition*. (Council of Europe, 2015).
10. De Boer, H. J., Ouarghidi, A., Martin, G., Abbad, A. & Kool, A. DNA Barcoding Reveals Limited Accuracy of Identifications Based on Folk Taxonomy. *PLoS ONE* **9,** e84291 doi: 10.1371/journal.pone.0084291 (2014).
11. Kool, A. *et al.* Molecular Identification of Commercialized Medicinal Plants in Southern Morocco. *PLoS ONE* **7,** e39459 doi: 10.1371/journal.pone.0039459 (2012).
12. Li, S. *et al.* Chemical markers for the quality control of herbal medicines: an overview. *Chinese Medicine* **3,** 7 (2008).
13. Techen, N., Crockett, S. L., Khan, I. A. & Scheffle, B. E. Authentication of Medicinal Plants Using Molecular Biology Techniques to Compliment Conventional Methods. *Current Medicinal Chemistry* **11,** 1391–1401 (2004).
14. Bennett, R. N. & Wallsgrove, R. M. Secondary metabolites in plant defence mechanisms. *New Phytologist* **127,** 617–633 doi: 10.1111/j.1469-8137.1994.tb02968.x (1994).
15. Kesting, J. R., Huang, J. F. & Sørensen, D. Identification of adulterants in a Chinese herbal medicine by LC-HRMS and LC-MS-SPE/NMR and comparative *in vivo* study with standards in a hypertensive rat model. (2010).
16. Chen, S. *et al.* Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species. *PLoS ONE* **5,** e8613 doi: 10.1371/journal.pone.0008613 (2010).
17. Newmaster, S., Grguric, M., Shanmughanandhan, D., Ramalingam, S. & Ragupathy, S. DNA barcoding detects contamination and substitution in North American herbal products. *BMC Medicine* **11,** 222 (2013).
18. Little, D. P. & Jeanson, M. L. DNA Barcode Authentication of Saw Palmetto Herbal Dietary Supplements. *Scientific Reports* **3,** 3518 doi: 10.1038/srep03518 (2013).
19. Stoeckle, M. Y. *et al.* Commercial Teas Highlight Plant DNA Barcode Identification Successes and Obstacles. *Sci. Rep.* **1,** doi: http://www.nature.com/srep/2011/110721/srep00042/abs/10.1038-srep00042-unlocked-60×70.html – supplementary-information (2011).
20. Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* **270,** 313–321 doi: 10.1098/rspb.2002.2218 (2003).
21. Kress, W. J., Wurdack, K. J., Zimmer, E. A., Weigt, L. A. & Janzen, D. H. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* **102,** 8369–8374 (2005).

22. Hollingsworth, P. M. *et al.* A DNA barcode for land plants. *Proceedings of the National Academy of Sciences* **106,** 12794–12797 (2009).
23. Coghlan, M. L. *et al.* Deep Sequencing of Plant and Animal DNA Contained within Traditional Chinese Medicines Reveals Legality Issues and Health Safety Concerns. *PLoS Genet* **8,** e1002657 doi: 10.1371/journal.pgen.1002657 (2012).
24. Hauke, R. L. A taxonomic monograph of the genus *Equisetum* subgenus *Hippochaete*. *Nova Hedwigia* **8,** 1–123 (1963).
25. Hauke, R. L. A taxonomic monograph of *Equisetum* subgenus *Equisetum*. *Nova Hedwigia* **30,** 385–455 (1978).
26. Sandhu, N. S., Kaur, S. & Chopra, D. *Equisetum arvense*: pharmacology and phytochemistry – a review. *Asian Journal of Pharmaceutical and Clinical Research* **3,** 146–150 (2010).
27. Committee on Herbal Medicinal Products. in *Evaluation of Medicines for Human Use* (ed _ European Medicines Agency) (2008).
28. Currie, H. A. & Perry, C. C. Silica in Plants: Biological, Biochemical and Chemical Studies. *Annals of Botany* **100,** 1383–1389 (2007).
29. Hauke, R. L. An analysis of a variable population of *Equisetum arvense* and *E. litorale*. *American Fern Journal* **55,** 123–135 (1965).
30. Veit, M. *et al.* Phenolic characters of Bristish hybrid taxa in *Equisetum* subgenus *Equisetum*. *Biochemical Systematics and Ecology* **23,** 79–87 (1995).
31. Missouri Botanical Garden and Harvard University Herbaria. eFloras, (2015). Availabe at: http://www.efloras.org (Accessed: 19th February 2015)
32. Karrer, P., Eugster, C. H. & Patel, D. K. Über Inhaltsstoffe einiger *Equisetum*-Arten. *Helvetica Chimica Acta* **32,** 2397–2399 doi: 10.1002/hlca.19490320724 (1949).
33. Phillipson, J. D. & Melville, C. An investigation of the alkaloids of some British species of *Equisetum*. *Journal of Pharmacy and Pharmacology* **12,** 506–508 doi: 10.1111/j.2042-7158.1960.tb12699.x (1960).
34. Roth, L., Daunderer, M. & Kormann, K. Giftpflanzen – Pflanzengifte: Vorkommen, Wirkung, Therapie – allergische und phototoxische Reaktionen. Mit Sonderteil über Gifttiere. (Nikol Verlag, 2012).
35. Guillon, J.-M. Phylogeny of Horsetails (*Equisetum*) based on the Chloroplast *rps4* Gene and Adjacent Noncoding Sequences. *Systematic Botany* **29,** 251–259 doi: 10.1600/036364404774195467 (2004).
36. Guillon, J.-M. Molecular phylogeny of horsetails (*Equisetum*) including chloroplast *atpB* sequences. *J Plant Res* **120,** 569–574 doi: 10.1007/s10265-007-0088-x (2007).
37. Des Marais, David L., Smith, Alan R., Britton, Donald M. & Kathleen, M. Pryer. Phylogenetic Relationships and Evolution of Extant Horsetails, *Equisetum*, Based on Chloroplast DNA Sequence Data (*rbcL* and *trnL-F*). *International Journal of Plant Sciences* **164,** 737–751 doi: 10.1086/376817 (2003).
38. Page, C. N. An assessment of the inter-specific relationships of *Equisetum* subgenus *Equisetum*. *New Phytologist* **71,** 355–369 (1972).
39. Page, C. N. Two hybrids in *Equisetum* new to the British flora. *Watsonia* **9,** 229–237 (1973).
40. Stace, C. A. *Hybridization and flora of the British Isles*. (Academic Press, 1975).
41. Chen, S. *et al.* A renaissance in herbal medicine identification: From morphology to DNA. *Biotechnology Advances* **32,** 1237–1244 doi: http://dx.doi.org/10.1016/j.biotechadv.2014.07.004 (2014).
42. Dong, W. *et al. ycf1*, the most promising plastid DNA barcode of land plants. *Sci. Rep.* **5,** doi: 10.1038/srep08348 (2015).
43. Kemp, C. The endangered dead. *Nature* **518,** 292–294 (2015).
44. Suarez, A. V. & Tsutsui, N. D. The Value of Museum Collections for Research and Society. *BioScience* **54,** 66–74 (2004).
45. Puillandre, N. *et al.* New taxonomy and old collections: integrating DNA barcoding into the collection curation process. *Molecular Ecology Resources* **12,** 396–402 doi: 10.1111/j.1755-0998.2011.03105.x (2012).
46. Särkinen, T., Staats, M., Richardson, J. E., Cowan, R. S. & Bakker, F. T. How to Open the Treasure Chest? Optimising DNA Extraction from Herbarium Specimens. *PLoS ONE* **7,** e43808 doi: 10.1371/journal.pone.0043808 (2012).
47. Staats, M. *et al.* DNA Damage in Plant Herbarium Tissue. *PLoS ONE* **6,** e28448, doi: 10.1371/journal.pone.0028448 (2011).
48. Kristiansen, K. A. *et al.* DNA Taxonomy—the Riddle of Oxychloë (Juncaceae). *Systematic Botany* **30,** 284–289 doi: 10.1600/0363644054223710 (2005).
49. Seberg, O. & Petersen, G. How Many Loci Does it Take to DNA Barcode a *Crocus*? *PLoS ONE* **4,** e4598 doi: 10.1371/journal.pone.0004598 (2009).
50. Pang, X. *et al.* Utility of the *trnH–psbA* Intergenic Spacer Region and Its Combinations as Plant DNA Barcodes: A Meta-Analysis. *PLoS ONE* **7,** e48833 doi: 10.1371/journal.pone.0048833 (2012).
51. Nadot, S., Bajon, R. & Lejeune, B. The chloroplast gene *rps4* as a tool for the study of Poaceae phylogeny. *Pl Syst Evol* **191,** 27–38 doi: 10.1007/BF00985340 (1994).
52. Souza-Chies, T. T. *et al.* Phylogenetic analysis of Iridaceae with parsimony and distance methods using the plastid gene *rps4*. *Pl Syst Evol* **204,** 109–123 doi: 10.1007/BF00982535 (1997).
53. Little, D. P. & Barrington, D. S. Major evolutionary events in the origin and diversification of the fern genus *Polystichum* (Dryopteridaceae). *American Journal of Botany* **90,** 508–514 (2003).
54. Sang, T., Crawford, D. & Stuessy, T. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84,** 1120–1120 (1997).
55. Tate, J. A. & Simpson, B. B. Paraphyly of *Tarasa* (Malvaceae) and Diverse Origins of the Polyploid Species. *Systematic Botany* **28,** 723–737 doi: 10.1043/02-64.1 (2003).
56. White, T. J., Bruns, T., Lee, S. & Taylor, J. in PCR Protocols (eds M. A. Innis, D. H. Gelfand, J. J. Shinsky & T. J. White) 315–322 (Academic Press, San Diego, California, 1990).
57. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30,** 772–780 (2013).
58. Stamatakis, A. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* (2014).
59. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39,** 783–791 (1985).

Acknowledgements

The authors thank Charlotte Hansen, Corrie Madsen, Cæcilie Ryhl Olsson and Mirnesa Rizvanovic (Natural History Museum of Denmark) for DNA sequences and Katrine Krydsfeldt (Department of Drug Design and Pharmacology, University of Copenhagen) for TLC analysis. Funding: The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Unions 7th Framework programme FP//2007/2013 under REA grant agreements no. PITN-GA-2013-606895 – MedPlant (NEI and NR) and no. PIEF-GA-2012-328637 – BiodiversityAltitude (CHSL).

Author Contributions

N.R. and A.N.K.J. conceived the study. N.R., A.N.K.J., O.S., G.P. and C.H.S.L. designed the study. N.R., N.E.I. and S.B.L. selected the herbarium specimens. N.E.I. identified the specimens. S.B.L. produced and assembled DNA sequences. C.H.S.L. conducted the phylogenetic analyses and interpreted the results. A.N.K.J. and N.R. designed and interpreted the TLC analysis. C.H.S.L. wrote the manuscript with N.R. and N.I. All authors contributed to interpretation of the data and commented on the manuscript. All authors read and approved the final edition of the manuscript.

Additional Information

Supplementary information accompanies this paper at http://www.nature.com/srep

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Saslis-Lagoudakis, C. H. *et al.* Identification of common horsetail (*Equisetum arvense* L.; Equisetaceae) using Thin Layer Chromatography versus DNA barcoding. *Sci. Rep.* **5**, 11942; doi: 10.1038/srep11942 (2015).

Appendix II

The application of high-throughput sequencing for taxonomy: the case of *Plantago* subg. *Plantago* (Plantaginaceae)

Manuscript is under revision in Molecular Phylogenetics and Evolution

The application of high-throughput sequencing for taxonomy: the case of *Plantago* subg. *Plantago* (Plantaginaceae)

Gustavo Hassemer[a,b,*], Sam Bruun-Lund[b], Alexey B. Shipunov[c], Barbara G. Briggs[d], Heidi M. Meudt[e] and Nina Rønsted[b]

[a] Laboratory of Plant Systematics, Federal University of Santa Catarina, CEP 88040-900, Florianópolis, SC, Brazil

[b] Natural History Museum of Denmark, University of Copenhagen, Øster Farimagsgade 5A, 1353 Copenhagen, Denmark

[c] Department of Biology, Minot State University, 500 University Avenue West, Minot, ND 58707, United States of America

[d] National Herbarium of New South Wales, Botanic Gardens Trust, Mrs Macquaries Road, Sydney, New South Wales 2000, Australia

[e] Museum of New Zealand Te Papa Tongarewa, PO Box 467, Cable St, Wellington 6140, New Zealand

[*] Author for correspondence. E-mail: gustavonaha@gmail.com

ABSTRACT

Plantago is a cosmopolitan genus including over 250 species, concentrated in temperate and high-elevation tropical regions. The taxonomy of *Plantago* is very difficult, mainly because of its reduced morphology, which features relatively few characters for species classification. Consequently, the infrageneric classification of the genus remains controversial and inadequate. In this study we applied high-throughput plastid genome skimming to provide powerful phylogenetic resolution to clarify the relationships within subg. *Plantago*, which is the largest, most broadly distributed and poorest understood subgenus of *Plantago*. Ninety-four samples covering 60% of all species and representing all sections of subg. *Plantago* as well as an outgroup were successfully sequenced. The resulting phylogenetic topology was used,

complemented by field and herbarium studies, to revise the sectional classification of subg. *Plantago* and present a complete listing of the accepted species in the subgenus. Our phylogenetic results were also tested for their usefulness in clarifying the taxonomic placement of some taxonomically complicated species in the subgenus. We conclude that a combination of morphological studies and state-of-the art high-throughput DNA data provide a useful toolbox for resolving outstanding taxonomic puzzles exemplified by the genus *Plantago*.

Keywords: classification; Plantagineae; phylogeny; taxonomy

1. Introduction

Determining the genealogy of each great Kingdom of Nature was Charles Darwin's dream (Darwin and Darwin, 1887). Since the emergence of Sanger sequencing techniques in the 1970s (Sanger et al., 1977), DNA sequencing has greatly advanced our understanding of the tree of life and shed new light on previous classifications based on taxonomic studies of primarily morphological characters (Chase and Savolainen, 2003; Morey et al., 2013; van Dijk et al., 2014; Heather and Chain, 2015). Starting in 2004, the so called next-generation sequencing, or high-throughput sequencing (HTS) became available, greatly increasing the speed and the amount of generated data, and hugely decreasing the sequencing cost per base (Morey et al., 2013; van Dijk et al., 2014; Heather and Chain, 2015). The massive amount of data generated by HTS techniques has powerful applications in potentially all fields of the biological sciences (Delseny et al., 2010; Koboldt et al., 2013; Buermans and den Dunnen, 2014), including taxonomy (Harrison and Kidner, 2011; Straub et al., 2012; Soltis et al., 2013). However, the associated costs and necessary infrastructure are still the main restrictors of its use (Delseny et al., 2010), especially in developing countries and low-funding research environments such as

taxonomy. For this reason, and also due to the lack of plant taxonomists in most high-funding institutions (see e.g. Agnarsson and Kuntner, 2007; Ebach et al., 2011; Wägele et al., 2011; Sluys, 2013), the application of HTS to resolve taxonomic problems is still in its infancy (e.g. Gardner et al., 2016; Hou et al., 2016; Uribe-Convers et al., 2017). In order to take full benefit from this new potential, a working connection between phylogeny, morphology and nomenclature is necessary, and without which phylogeny is not translated into advanced systematics.

In this study we demonstrate the utility of applying state-of-the art high-throughput DNA data to test current taxonomic understanding based on morphology and help resolve outstanding taxonomic problems exemplified by the plant genus *Plantago* L. (Plantaginaceae).

The Plantaginaceae had its circumscription radically altered with recent molecular phylogenetic studies (Olmstead et al., 2001; Albach et al., 2005), having been greatly expanded with the inclusion of a large number of species from the former Scrophulariaceae (sensu lato) plus Callitrichaceae, Globulariaceae and Hippuridaceae (Albach et al., 2005). Within the family, tribe Plantagineae (Albach et al., 2005) comprises *Plantago*, *Littorella* P.J.Bergius (Hoggard et al., 2003; Hassemer et al., 2018) and *Aragoa* Kunth (Bello et al., 2002). Molecular phylogenetic analyses have also been used to review the circumscriptions, biogeography and phylogenetic relationships of taxa within Plantagineae (Bello et al., 2002; Rønsted et al., 2002; Hoggard et al., 2003; Cho et al., 2004; Dunbar-Co et al., 2008; Meyers and Liston, 2008; Ishikawa et al., 2009; Tay et al., 2010a; Iwanycki Ahlstrand et al., in press).

Plantago is a cosmopolitan genus which has diversified into over 250 species which are usually anemophilous herbs or rarely subshrubs, perennial or annual, and concentrated in temperate and high-elevation tropical regions (Figs. 1, 2 and 3; Pilger, 1937; Rahn, 1996; Li et al., 2011). Although some species have wide geographic distributions, a few such as *P. major* and *P. lanceolata* L.

being cosmopolitan ruderals, many others have restricted geographic distributions, occurring in more specialised environments, and several of these are endemic to oceanic islands (Dunbar-Co et al., 2008; Meudt, 2012; Hassemer et al., 2016; Iwanycki Ahlstrand et al., in press). A number of *Plantago* species are well-known for their medicinal properties, and also for other traditional uses (Samuelsen, 2000; Weryszko-Chmielewska et al., 2012; Gonçalves and Romano, 2016).

In many areas, *Plantago* species have successfully colonised new habitats and then have undergone consequent rapid and recent diversification, including an extremely high level of mitochondrial DNA evolution often contrasting with low morphological variation (Rønsted et al., 2002; Cho et al., 2004; Meyers and Liston, 2008; Tay et al., 2010a; Ishikawa et al., 2009). Although the genus is one of the most well-studied plant genera from a taxonomic viewpoint, this low morphological variation, reduced morphology, and lack of useful taxonomic characters have precluded a full understanding of the evolution and classification of the genus and its species (Rahn, 1996; Rønsted et al., 2002; Ishikawa et al., 2009; Tay et al., 2010b; Meudt, 2011). Trichomes and seeds are considered the most informative morphological characters (Rahn, 1992, 1996), and both trichomes (e.g. Andrzejewska-Golec, 1991; Rahn, 1992; Andrzejewska-Golec and Świętosławski, 1993) and seeds (e.g. Liu et al., 1992; Shipunov, 1998; Klimko et al., 2004; Shehata and Loutfy, 2006) have been investigated. A series of chemotaxonomic studies have also been conducted (e.g. Andrzejewska-Golec et al., 1993; Jensen et al., 1996; Rønsted et al., 2000, 2003; Taskova et al., 2002). However, none of these characters have enabled a satisfying infrageneric classification of the genus. To complicate things further, there is evidence of polyploidy (Murray et al., 2010; Wong and Murray, 2012, 2014), hybridisation (Rahn, 1974; Wong and Murray, 2014) and reticulate evolution (Ishikawa et al., 2009) in *Plantago*.

According to the classification of Rahn (1996) based on a cladistic analysis of morphological characters, with updates by Rønsted et al. (2002) and Hoggard et al. (2003) using plastid *trn*L–F and nuclear encoded ITS sequence data, the genus *Plantago* is subdivided into four subgenera: *Bougueria* (Decne.) Rahn (once considered to be its own monotypic genus; Rahn, 1996; Rønsted et al., 2002), *Coronopus* (Lam. & DC.) Rahn, *Plantago* and *Psyllium* (Mill.) Harms & Reiche. Also according to this classification, subgenus *Plantago*, the focus of the current study, is cosmopolitan, includes 143 species, and is in turn subdivided into five sections: *Mesembrynia*, *Micropsyllium*, *Plantago*, *Oliganthos* and *Virginica* (see Table 1). Rahn (1996) deemed his sects. *Mesembrynia* and *Plantago* to be not monophyletic, whereas his sects. *Micropsyllium*, *Oliganthos* and *Virginica* were monophyletic according to his analyses.

Table 1. Summary of the sections accepted by Rahn (1996) in *Plantago* subg. *Plantago*, the distribution and the number of species in each section, and the monophyletic status according to his analyses.

Section	Native distribution	No. of species	Monophyletic?
Mesembrynia	Australasia and Eurasia	32	No
Micropsyllium	North America and Eurasia	6	Yes
Oliganthos	Australasia and South America	24	Yes
Plantago	Worldwide except mainland South America and Australasia	53	No
Virginica	The Americas	28	Yes

Previous molecular phylogenetic studies by Rønsted et al. (2002), Hoggard et al. (2003), Ishikawa et al. (2009) and Tay et al. (2010a) based on Sanger sequencing of a limited number of DNA regions from a combined number of ca. 40 species (~28%) of subg. *Plantago* indicated that, although there is strong evidence pointing to the monophyly of subg. *Plantago*, and for sect. *Micropsyllium* being sister to the remainder of the subgenus, most sections

are not monophyletic nor well resolved, demonstrating the need for further investigations to understand the phylogenetic relationships within this group. A recent phylogenetic study (Iwanycki Ahlstrand et al., in press) based on five DNA regions (*nr*ITS and four plastid regions) focusing on the biogeography of the oceanic island endemic species in subg. *Plantago* (30 species included), confirmed the polyphyly of all sections except *Micropsyllium* and *Virginica*. Thus, a revised molecular phylogeny of subg. *Plantago* with increased sampling and more informative markers is therefore needed to better understand the taxonomy, phylogenetic relationships, biogeography and evolutionary history of this group.

The objective of this study is to apply HTS genome skimming techniques to reconstruct phylogenetic relationships within *Plantago* subg. *Plantago* and to propose a new sectional classification of the subgenus.

2. Materials and methods

2.1. Herbarium specimens revision

Due to the difficult taxonomy and identification of *Plantago*, which leads to a considerable proportion of herbarium specimens being misidentified, we found it necessary to conduct an extensive revision of herbarium specimens to ensure that we associate the correct names to the sampled specimens. Special attention was given to type specimens, which fix the application of names and are critical for the correct taxonomic classification. Collections from the following herbaria were studied: AK, ASE, BHCB, C, CEN, CGMS, CHR, CIIDIR, DDMS, EAC, EFC, FI, FLOR, FT, FURB, GB, GH, HAS, HBR, HO, HRB, HURB, IAC, ICN, K, LD, MA, MBM, MVFA, MVJB, MVM, OTA, P, PI, RB, SGO, TANG, TEPB, TUB, UB, UESC, UFMT, UPCB, UPS and WELT (herbarium codes follow Thiers, 2018). Furthermore, images of specimens kept at the following herbaria were studied: A, B, BBF, BM, BR, COI, CONC, CORD, CTES, DD, E, ESA, F, G, GOET, HFLA, IRAI, L, LD,

LE, LINN, M, MO, MPU, PH, PRC, R, RO, S, SP, UC, UEC, US and W. The nomenclature presented here follows the *Shenzhen Code* (Turland et al., 2018). Author names of species and sections included in the taxonomic treatment or in Table 3 are not repeated elsewhere in the text. The terminology and interpretation of morphological characters for *Plantago* follow Rahn (1992, 1996). Unless otherwise informed, all field photographs were taken by the authors.

2.2. Sampling strategy

The samples used in this study were obtained through field work, from cultivated individuals at the Botanical Garden of the Natural History Museum of Denmark, University of Copenhagen, and also from herbarium specimens. Furthermore, nineteen DNA extracts from the Kew DNA Bank (http://apps.kew.org/dnabank/homepage.html) were used. In addition to these, two published reference plastomes were also included in our phylogenies: *P. maritima* L. (GenBank acc. nr. KR297244) and *P. media* (GenBank acc. nr. KR297245) (Zhu et al., 2015).

For this study we successfully sequenced 94 DNA accessions corresponding to 87 *Plantago* species and one *Littorella.* Details of voucher materials are listed in Table 2; these sequences will be submitted to TreeBase during the review process. The five included outgroups were *L. uniflora* (L.) Asch. and two samples each of *Plantago* subgenera *Coronopus* and *Psyllium*. Eleven additional samples (10 *Plantago* and one *Aragoa* species) were sequenced but could not be used for the phylogenies due to highly degraded DNA or evident contamination. Construction of libraries failed twice with samples of *P. nubicola* (Decne.) Rahn, the sole representative of *Plantago* subg. *Bougueria*.

The sampling strategy focused on covering all sections recognised by Rahn (1996) for *Plantago* subg. *Plantago*. We also included species whose

phylogenetic placement had already been considered problematic in the literature, especially the six Eurasian species of Rahn's (1996) sect. *Mesembrynia* (*P. arachnoidea*, *P. camtschatica*, *P. depressa*, *P. komarovii*, *P. perssonii* and *P. schwarzenbergiana*), and as many species as possible of sect. *Plantago,* which Rahn himself deemed to be paraphyletic (Rahn, 1996). Furthermore, we included several species from South America, Australia and New Zealand, as these species-rich areas have been under-sampled in previous molecular studies. Additional samples of some species were included with the purpose of testing the phylogenetic placement of taxonomically problematic subspecies or populations of *P. australis* and *P. lanigera*.

Table 2. List of 94 DNA samples successfully sequenced used for the final phylogenetic analysis. Herbarium codes are listed in parenthesis.

Taxon	**Voucher**	**Sample provenance**
L. uniflora	Chase 2798 (K); Kew DNA Bank 2798	England
P. alismatifolia	Mosquin et al. 6813 (GH)	Northwestern Mexico
P. alpestris	Briggs 10181 (NSW-884676)	Southeastern Australia
P. arachnoidea	Gubanov and Kamelin 2662a (MW)	Mongolia
P. arborescens	Hassemer 918 (C)	Madeira Island, Portugal; cultivated in Copenhagen
P. asiatica	Liu 15395 (C)	Central China
P. aucklandica	Wright s.n. (WELT-SP090482)	Auckland Islands, New Zealand
P. australis subsp. *australis*	Hassemer 738 (FLOR)	Grão Pará, southern Brazil
P. australis subsp. *cumingiana*	Hassemer 917 (C)	Central Chile; cultivated in Copenhagen
P. australis subsp. *hirtella*	Hassemer 768 (C)	Joinville, southern Brazil
P. australis subsp. *leioloma*	Arsène 5422 (US-00137259)	Central Mexico
P. bradei	Hassemer 826 (C)	Alto Caparaó, eastern Brazil

P. camtschatica	Rahn 684 (C); Kew DNA Bank 9402	Origin unknown; cultivated in Copenhagen
P. canescens	Pospelov s.n. (MW)	Northern Russia
P. catharinea 1	Hassemer 706 (FLOR)	Florianópolis, southern Brazil
P. catharinea 2	Hassemer 819 (C)	Santos, southern Brazil
P. cavaleriei	Sino-Brit. exp. Cangshan 935 (K); Kew DNA Bank 31933	Yunnan, China
P. commersoniana	Hassemer 832 (C)	Montevideo, Uruguay
P. cordata	Wagner and Fritsch 90012 (NY)	Michigan, USA
P. cornutii	Rønsted 31 (C); Kew DNA Bank 11180	Origin unknown; cultivated in Copenhagen
P. corvensis	Hassemer 737 (FLOR)	Grão Pará, southern Brazil
P. daltonii	Briggs 9782 (NSW-743874)	Tasmania, Australia
P. debilis	Briggs 10184 (NSW-899215)	Eastern Australia
P. depressa	Yongsok 6295 (F-1535438)	Ulleung Island, South Korea
P. elongata	Bare 1113 (NY)	North Dakota, USA
P. eriopoda	Anonymous s.n. (OKL); Kew DNA Bank 30432	USA
P. euana	Sykes 879/T (US-3121974)	Tonga Islands
P. euryphylla	Briggs 10175 (NSW-884716)	Southeastern Australia
P. fernandezia	Solbrig et al. 3907 (GH)	Juan Fernández Islands, Chile
P. floccosa	Spellman et al. 990 (MO-2898184)	Central Mexico
P. gaudichaudii	Hosking 3286 (NSW-841427)	Eastern Australia
P. gentianoides	Buia et al. s.n. (NY)	Romania
P. glacialis	Briggs 10180 (NSW-884675)	Southeastern Australia
P. guilleminiana	Hassemer 884 (C)	Urubici, southern Brazil
P. hatschbachiana	No voucher; photo: Fig. 3 in Hassemer (2016)	Ponta Grossa, southern Brazil
P. hawaiensis	Dunbar-Co 2002	Hawaii Island, USA

P. hedleyi	Seed 31 (NSW-787790)	Lord Howe Island, Australia
P. himalaica	Stewart 21871 (NY)	Kashmir
P. humboldtiana	Hassemer 766 (C)	Corupá, southern Brazil
P. incisa	Filip H578184-52 (K); Kew DNA Bank 11191	Java, Indonesia
P. komarovii	Petelin 99-546 (MW)	Mongolia
P. lanceolata	Hassemer 364 (FLOR)	Florianópolis, southern Brazil
P. lanigera 1	Meudt 268 (WELT-SP090353)	Rock and Pillar Range, New Zealand
P. lanigera 2	Heenan s.n. (CHR-688758)	Sewell Peak, New Zealand
P. longissima	Glen 1928 (US-3438221)	Northern South Africa
P. macrocarpa	Volkova et al. s.n. (MW-0156805)	Bering Island, northeastern Russia
P. major	Hassemer 760 (C)	Florianópolis, southern Brazil
P. maxima	Rønsted 28 (C); Kew DNA Bank 11181	Origin unknown; cultivated in Copenhagen
P. moorei	Moore 729 (GH)	West Falkland, UK
P. muelleri	Briggs 10179 (NSW-884674)	Southeastern Australia
P. myosuros 1	Hassemer 834 (FURB)	Montevideo, Uruguay
P. myosuros 2	Hassemer 837 (C)	Lavalleja, Uruguay
P. napiformis	Hassemer 809 (C)	Ponta Porã, western Brazil
P. novae-zelandiae	Tay 52 (WELT-SP090356)	Ruahine Range, New Zealand
P. pachyneura	Hassemer 805 (C)	Central Chile; cultivated in Copenhagen
P. pachyphylla	Dunbar-Co 2155 (PTBG)	Oahu Island, USA
P. palmata	Rønsted 9 (C)	Rwanda
P. palustris	Hosking 2486 (NSW-693662)	Eastern Australia
P. paradoxa	Briggs 9781 (NSW-743924)	Tasmania, Australia
P. personii	Qinghai-Xizang exp. 870947 (PE); Kew DNA Bank 20552	Xinjiang, China

P. picta	Atkins s.n. (WELT-SP086772)	Tolaga Bay, New Zealand
P. polysperma	Tsvelev et al. 995 (LE)	Kazakhstan
P. princeps	Dunbar-Co 2341 (PTBG)	Oahu Island, USA
P. pusilla	Cusick and Gardner 36054 (NY)	Indiana, USA
P. rahniana	Hassemer 786 (C)	Bom Jardim da Serra, southern Brazil
P. raoulii	Meudt 281 (WELT-SP086777)	Puketapu, New Zealand
P. rapensis	Motley 2740 (K); Kew DNA Bank 20557	Rapa Iti Island, France
P. reniformis	Rønsted 42 (C); Kew DNA Bank 9446	Origin unknown; cultivated in Copenhagen
P. rhodosperma	No voucher; photo: Fig. S1	Texas, USA
P. rigida	Chase 2767.B (K); Kew DNA Bank 2767.1	Peru
P. rugelii	Rønsted 37 (C); Kew DNA Bank 9447	Ontario, Canada
P. rupicola	Dunbar-Co 2268 (PTBG)	Rapa Iti Island, France
P. schwarzenbergiana	Boros s.n. (GH)	Hungary
P. sparsiflora	LeBlond 5305 (CSU); Kew DNA Bank 30433	Origin unknown
P. spathulata	Garnock-Jones 2629 (WELT-SP090461)	Marfells Beach, New Zealand
P. spathulata × *raoulii* 1	Tay 49 (WELT-SP090387)	Sugarloaf Pass, New Zealand
P. spathulata × *raoulii* 2	Barkla s.n. (WELT-SP087211)	Old Man Range, New Zealand
P. stauntonii	Rahn 706 (C)	St. Paul and New Amsterdam Islands, France
P. subnuda	McClintock and Wheeler s.n. (UC-530075)	California, USA
P. subspathulata	Hassemer 808 (C)	Madeira Island, Portugal; cultivated in Copenhagen

P. subulata	Hassemer 916 (C)	Origin unknown; cultivated in Copenhagen
P. tanalensis	Deroin 260 (MO-5970257)	Madagascar
P. tasmanica	Briggs 9791 (NSW-743928)	Tasmania, Australia
P. tehuelcha	Eyerdam et al. 24025 (GH)	Southern Argentina
P. tomentosa	Hassemer 793 (C)	Santo Antônio das Missões, southern Brazil
P. triandra	Tay 55 (WELT-SP090357)	Manaia, New Zealand
P. trinitatis	Port s.n. (FLOR-49242)	Trindade Island, Brazil
P. tubulosa	Webster 67 (K); Kew DNA Bank 19210	Puno, Peru
P. turficola	Hassemer 621 (FLOR)	Urubici, southern Brazil
P. tweedyi	Hoggard 518 (CSU); Kew DNA Bank 30436	Origin unknown
P. udicola	Sneddon s.n. (WELT-SP090378)	Tablelands, New Zealand
P. unibracteata	Meudt 273 (WELT-SP090464)	Rock and Pillar Range, New Zealand
P. varia	Briggs 10177 (NSW-884666)	Eastern Australia
P. weddelliana	Hjerting et al. 180 (F-1607387)	Northwestern Argentina

2.3. High-Throughput DNA sequencing

Whole genomic DNA was extracted from fresh, silica gel dried or herbarium specimens using the Qiagen DNeasy Minikit (Qiagen, Germany) following the manufacturer's protocol with the three following modifications to increase yield: (1) 50–60 mg dried pulverized tissue was used for each extraction; (2) 50 µl proteinase K was added and incubated for 1 h at 45 °C following the second step in the manufacturers protocol (i.e. add 400 µl of AP1 buffer and 4 µl of RNase A, mix, and incubate for 10 minutes at 65 ^{o}C); and (3) the final elution step was done thrice using 120 µl AE buffer, but re-pipetting the flow-through onto the spin column

each time (instead of adding new AE buffer). DNA was quantified using a Qubit 2.0 fluorometer (Life Technologies, USA) following the manufacturer's instructions for high sensitivity.

Prior to preparing the libraries for sequencing, DNA was fragmented to ca. 300 basepairs (bp) using a Bioruptor (Diagenode, Belgium), running four cycles, with 15 s ON / 90 s OFF. Illumina-compatible 100 bp paired-end libraries from DNA extracts were prepared using NEBNext Library building kits (New England Biolabs, USA, catalogue nr. E6070L) following the manufacturer's protocol. The libraries were amplified using AmpliTaq Gold (Life Technologies, USA), and had their quality checked using a 2200 TapeStation (Agilent Technologies, USA). Subsequently, the libraries were multiplexed and sequenced on three lanes with 32, 40 and 42 samples respectively using an Illumina HiSeq2000 platform at the Danish National High-Throughput DNA Sequencing Centre. For this study we used the Illumina platform because of the large amount of data that it generates, but also because the error-rate in base calling of this method is the lowest, making it advantageous for the purposes of this study compared to other high-throughput sequencing methods available (Bruun-Lund et al., 2017).

2.4. Data analyses

2.4.1. Reference-based plastome assembly

The sequencing resulted in 180.27 gigabytes of reads for 105 samples. Following the analysis pipeline of Bruun-Lund et al. (2017), the sequences were filtered to remove adaptors and low quality reads using AdaptorRemoval v. 2 (Schubert et al. 2016) running with the default settings and a minimum read length set to 30 bp. The data were then tested for quality using FastQC

(http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Reads were then imported into Geneious v. 9.1.8 (Biomatters Ltd., New Zealand). The resulting high-quality reads were then paired and subsequently reference-based assembled to a published plastome of *P. media* (GenBank acc. nr. KR297245; Zhu et al., 2015) using the Bowtie2 v. 2.3.2 (Langmead and Salzberg, 2012) plugin in Geneious, choosing end to end, high-medium sensitivity. This reference was chosen because it was the only published plastome of a species in subg. *Plantago*. Then the consensus sequence of the result was extracted using a 50% (strict) threshold, calling "?" if no coverage or coverage is less than 10 reads. For each of the samples, between 949,457 and 1,049,210 reads were mapped to *P. media* when used as a reference to the mapping process.

Next, successive alignments of the 105 samples sequenced, in addition to the two reference plastomes of *P. media* and *P. maritima*, were made using the MAFFT v. 7.309 (Katoh and Standley, 2013) plugin for Geneious, choosing the default settings. At this point the need for removing eleven of the sequenced samples was verified, either because of lack of high levels of endogenous DNA (two samples) or because of the negative impacts that highly degraded DNA caused to the alignments (nine samples), and thus the final alignment was reduced to 96 sequences including the two reference plastomes. In order to avoid artificially increasing the phylogenetic signal from the inverted repeated region in chloroplast genomes, one of these repeated regions was removed prior to analysis (Bruun-Lund et al., 2017). The final aligned matrix of the 96 sequences included 215,259 bp before removing one of the inverted repeated regions and 164,720 bp after.

2.4.2. Phylogenetic analysis

We considered the plastome as a single heritable unit. The final plastome alignment of 96 taxa and 164,720 bp was uploaded to CIPRES (www.phylo.org), and was tested for the most appropriate model of evolution using jModelTest2 v. 2.1.6) (Darriba et al. 2012) using the default settings.

According to the Akaike information criterion, as recommended by Posada and Buckley (2004), the model was inferred to be GTR + G. Next, a maximum likelihood (ML) analysis was conducted to search for the best tree using RAxML-HPC v. 8.2.9 (Stamatakis 2014) in CIPRES, with the following changes from the default settings: maximum hours to run: 100; model for bootstrapping phase: GTRGAMMA; analysis type: rapid bootstrap analysis / search for best-scoring ML tree; bootstrapping type: rapid bootstrapping; bootstrap iterations: 1000 (the maximum value allowed).

To verify the results of the ML analyses, a Bayesian inference analysis was conducted with MrBayes v. 3.2.6 (Ronquist et al., 2012), also in CIPRES, using two independent runs and four chains, sampling every 500 generations for up to 50 million generations, and capped at 100 hours of analysis resulting in 10,110,000 generations. Chain convergence and effective sample size parameters were inspected with Tracer v. 1.6 (http://tree.bio.ed.ac.uk/software/tracer) and the first 25% of the trees sampled from the posterior were discarded as burn-in. Using the program sumtrees.py from DendroPy v. 4.0.3 (https://github.com/jeetsukumaran/dendropy) (Sukumaran and Holder, 2010) we produced a maximum credibility clade tree. The best tree obtained from both maximum likelihood and Bayesian inference was viewed and annotated using FigTree v. 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree).

3. Results

3.1. Plastome phylogeny

The final alignment of the plastome dataset included 96 samples encompassing 88 species, and allowed the highest coverage to date (83 species; ~60%) of *Plantago* subg. *Plantago*, which is here estimated to include 147 described species. The topology of the trees obtained using the two different analyses (maximum likelihood with RAxML and Bayesian inference with

MrBayes; see above) was identical (Fig. 4; see also Figs. S2 and S3), confirming the robustness of our data. The support for clades at higher levels was generally high (most often 100% bootstrap support and posterior probability = 1.00). However, the support for some more terminal clades, especially in the more species-rich clades (sects. *Mesembrynia* and *Virginica*, see below) was generally much lower.

The infrageneric taxa mentioned here refer to the classification of Rahn (1996), updated by Rønsted et al. (2002) and Hoggard et al. (2003). The RaxML consensus tree is presented in Fig. 4 with clades with low posterior probabilities (PP<1.00 or BS<100%) indicated on the branches. ML and Bayesian phylograms showing branch lengths are included in the Supplementary Material, Figs. S2 and S3. The resulting plastome tree topology (Fig. 4) is in accordance with the topologies from previous phylogenetic studies based on Sanger sequencing of both plastid and nuclear data (Rønsted et al., 2002; Hoggard et al., 2003; Ishikawa et al., 2009; Tay et al., 2010a; and Iwanycki Ahlstrand et al., in press), but providing significantly improved resolution of *Plantago* subg. *Plantago*.

The plastome topology obtained here shows subg. *Plantago* to be monophyletic with strong support (BS=100%; PP=1.00). *Plantago* sect. *Micropsyllium* (clade A; BS=100%; PP=1.00) is sister to the remainder of subg. *Plantago* (BS=100%; PP=1.00). The next dichotomy is between a clade of two species of sect. *Plantago* from southeastern Europe and Asia Minor (clade B; *P. gentianoides* and *P. reniformis*; BS=97; PP=1.00) and the remaining clade, which has lower support (BS=52%; PP=0.95). Within this clade, a well-supported clade (clade C; BS=100%; PP=1.00) consisting of three species from sect. *Mesembrynia* (*P. arachnoidea*, *P. perssonii* and *P. schwarzenbergiana*) and three from sect. *Plantago* (*P. canescens*, *P. maxima* and *P. media*) intermixed, is sister to a clade of the remaining species (BS=100%; PP=1.00).

Within this clade a well-supported clade (clade D; BS=100%; PP=1.00) with three Sub-Saharan African species of sect. *Plantago* (*P. longissima*, *P. palmata* and *P. tanalensis*), is sister to the remaining species (BS=100%; PP=1.00). The next dichotomy consists of a larger well-supported clade (clade E; BS=100%; PP=1.00) with species from Australia, New Zealand and St. Paul and New Amsterdam Islands, coming from sects. *Mesembrynia*, *Oliganthos* and *Plantago*, and a clade of the remaining species (BS=100%; PP=1.00). Within this remaining clade we obtain a clade (BS=100%; PP=1.00) with six species, which splits into two clades: one clade (Clade F; BS=100%; PP=1.00) consisting of three species from sect. *Plantago* (including *P. major*), and another clade (BS=55%; PP=0.73) consisting of *P. cordata* (from sect. *Plantago*), and a clade (Clade G; BS=100%; PP=1.00) with two species from sect. *Oliganthos* (*P. rigida* and *P. tubulosa*). Subsequently, a well-supported clade (clade H; BS=100%; PP=1.00) including 12 species from Asia, North America and oceanic Pacific islands, nine from sect. *Plantago* and three from sect. *Mesembrynia*, is sister to a clade of the remaining species (BS=100%; PP=1.00). The last major dichotomy shows *P. macrocarpa* (sect. *Plantago*), from northwestern North America and northeastern Asia, as sister to a clade (clade I; BS=87%; PP=1.00) of American species from sects. *Oliganthos*, *Plantago* (*P. fernandezia*) and *Virginica*.

4. Discussion

4.1. Revised sectional classification

Supported by our plastome phylogeny (Fig. 4), in addition to the revision of herbarium collections, and a comprehensive revision of the taxonomic and phylogenetic literature on *Plantago*, we propose here a revised sectional classification of *Plantago* subg. *Plantago* (Table 3).

The new sectional classification proposed here follows the following principles:

(1) All accepted sections must be monophyletic;
(2) The recognised sections must be morphologically (especially regarding fruit and flower, the evolutionarily most conservative characters) and secondarily geographically coherent;
(3) The new classification should take into consideration, as much as possible, aspects from the latest classification (Rahn, 1996), but also elements of previous classifications (Barnéoud, 1844, 1845; Decaisne, 1852; Pilger, 1937) when these were proved correct in light of the revised phylogeny.

Table 3. Accepted species in *Plantago* subg. *Plantago* and their native distributions. Their former sectional placement and chromosome numbers are according to Rahn (1996). The corresponding clades highlighted in Fig. 4 are indicated after the section names. Species not included in the present phylogeny, but included in the phylogeny of Iwanycki Ahlstrand et al. (in press) are marked with an exclamation mark (!); species not included in either phylogenies, but whose position we inferred based on the accumulated knowledge (see Rosenberg and Kumar, 2001) are marked with an asterisk (*).

Species	Native distribution	Previously in sect.	2*n* =
Sect. *Carpophorae* Rahn — clade G — 2 species			
P. rigida Kunth	W Bolivia to SW Venezuela	*Oliganthos*	72
P. tubulosa Decne.	NW Argentina to S Mexico	*Oliganthos*	24, 48
Sect. *Eremopsyllium* Pilg. — clade B — 2 species			
P. gentianoides Sibth. & Sm.	SE Europe and SW Asia	*Plantago*	12
P. reniformis Beck	SE Europe	*Plantago*	12
Sect. *Heptaneuron* Decne. — 1 species			
P. cordata Lam.	E North America	*Plantago*	24
Sect. *Holopsyllium* Pilg. — 1 species			
P. macrocarpa Cham. & Schltdl.	NW North America and NE Asia	*Plantago*	24
Sect. *Lamprosantha* Decne. — clade C — 6 species			

	P. arachnoidea Schrenk ex Fisch. & C.A.Mey.	C Asia	*Mesembrynia*	?
	P. canescens Adams	N Asia and NW North America	*Plantago*	12
	P. maxima Juss. ex Jacq.	W Eurasia	*Plantago*	12
	P. media L.	Europe	*Plantago*	12, 24
	P. perssonii Pilg.	NW China	*Mesembrynia*	
	P. schwarzenbergiana Schur	Europe	*Mesembrynia*	12
Sect. *Leptostachys* Decne. — clade D — 7 species				
!	*P. africana* Verdc.	E Africa	*Plantago*	?
!	*P. fischeri* Engl.	E Africa	*Plantago*	?
*	*P. laxiflora* Decne.	S Africa	*Plantago*	?
	P. longissima Decne.	S Africa	*Plantago*	?
	P. palmata Hook.f.	Africa	*Plantago*	24
*	*P. remota* Lam.	S Africa	*Plantago*	?
	P. tanalensis Baker	Madagascar	*Plantago*	?
Sect. *Mesembrynia* Decne. — clade E — 44 species				
	P. alpestris B.G.Briggs et al.	SE Australia	*Mesembrynia*	12
*	*P. aundensis* P.Royen	New Guinea	*Oliganthos*	?
*	*P. antarctica* Decne.	SE Australia	*Mesembrynia*	12
	P. aucklandica Hook.f.	Auckland Islands (New Zealand)	*Plantago*	?
*	*P. bellidioides* Decne.	Tasmania (Australia)	*Mesembrynia*	12
*	*P. cladarophylla* B.G.Briggs et al.	E Australia	*Mesembrynia*	36
*	*P. cunninghamii* Decne.	Australia	*Mesembrynia*	12
	P. daltonii Decne.	Tasmania (Australia)	*Mesembrynia*	?
	P. debilis R.Br.	E Australia	*Mesembrynia*	12
*	*P. depauperata* Merr. & L.M.Perry	New Guinea	*Oliganthos*	?
*	*P. drummondii* Decne.	Australia	*Mesembrynia*	12
	P. euana Hurlim.	Tonga Islands	*Mesembrynia*	?
	P. euryphylla B.G.Briggs et al.	SE Australia	*Mesembrynia*	12
*	*P. exilis* Decne.	W Australia	*Mesembrynia*	?
	P. gaudichaudii Barnéoud	E Australia	*Mesembrynia*	12
*	*P. glabrata* Hook.f.	Tasmania (Australia)	*Mesembrynia*	24
	P. glacialis B.G.Briggs et al.	SE Australia	*Oliganthos*	12
*	*P. gunnii* Hook.f.	Tasmania (Australia)	*Oliganthos*	36
	P. hedleyi Maiden	Lord Howe Island	*Plantago*	24
*	*P. hispida* R.Br.	E Australia	*Mesembrynia*	12
	P. lanigera Hook.f.	New Zealand	*Oliganthos*	12, 24
*	*P. montisdicksonii* P.Royen	New Guinea	*Mesembrynia*	?
	P. muelleri Pilg.	SE Australia	*Oliganthos*	36
*	*P. multiscapa* B.G.Briggs	Australia	*Mesembrynia*	?
	P. novae-zelandiae L.B.Moore	New Zealand	*Oliganthos*	24
*	*P. obconica* Sykes	New Zealand	*Oliganthos*	12
	P. palustris L.R.Fraser & Vickery	E Australia	*Oliganthos*	24
*	*P. papuana* P.Royen	New Guinea	*Mesembrynia*	?

	P. paradoxa Hook.f.	Tasmania (Australia)	*Oliganthos*	24
*	*P. pentasperma* Hemsl.	St. Paul and New Amsterdam Islands (France)	*Mesembrynia*	?
	P. picta Colenso	New Zealand	*Mesembrynia*	48
*	*P. polita* Craven	New Guinea	*Oliganthos*	?
	P. raoulii Decne.	New Zealand	*Mesembrynia*	48
	P. spathulata Hook.f.	New Zealand	*Mesembrynia*	48
	P. stauntonii Reichardt	St. Paul and New Amsterdam Islands (France)	*Mesembrynia*	24
*	*P. stenophylla* Merr. & L.M.Perry	New Guinea	*Oliganthos*	?
	P. tasmanica Hook.f.	Tasmania (Australia)	*Mesembrynia*	12
	P. triandra Berggr.	New Zealand	*Oliganthos*	48
*	*P. triantha* Spreng.	Auckland Islands (New Zealand) and Tasmania (Australia)	*Oliganthos*	12
*	*P. trichophora* Merr. & L.M.Perry	New Guinea	*Mesembrynia*	?
*	*P. turrifera* B.G.Briggs et al.	Australia	*Mesembrynia*	12
	P. udicola Meudt & Garn.-Jones	New Zealand	none (new species)	96
	P. unibracteata Rahn	New Zealand	*Oliganthos*	60, 72
	P. varia R.Br.	E Australia	*Mesembrynia*	12
Sect. *Micropsyllium* Decne. — clade A — 6 species				
!	*P. bigelovii* A.Gray	W North America	*Micropsyllium*	20
	P. elongata Pursh	W North America	*Micropsyllium*	12, 36
*	*P. heterophylla* Nutt.	North America	*Micropsyllium*	12
	P. polysperma Kar. & Kir.	C Asia	*Micropsyllium*	?
	P. pusilla Nutt.	North America	*Micropsyllium*	12
!	*P. tenuiflora* Waldst. & Kit.	W Eurasia	*Micropsyllium*	24
Sect. *Pacifica* Hassemer — clade H — 26 species				
*	*P. alata* Nakai	Jeju (Korea)	*Plantago*	?
	P. asiatica L.	E and SE Asia	*Plantago*	24
	P. camtschatica Link	NE Asia	*Mesembrynia*	12
	P. cavaleriei H.Lév.	S China	*Plantago*	?
*	*P. coreana* H.Lév.	Jeju (Korea)	*Plantago*	?
	P. depressa Willd.	Asia	*Mesembrynia*	12
	P. eriopoda Torr.	North America	*Plantago*	24
*	*P. glabrifolia* (Rock) Pilg.	Hawaiian Archipelago (USA)	*Plantago*	
*	*P. grayana* Pilg.	Hawaiian Archipelago (USA)	*Plantago*	
*	*P. hakusanensis* Koidz.	Japan	*Plantago*	?
*	*P. hasskarlii* Decne.	Java (Indonesia)	*Plantago*	?
	P. hawaiensis (A.Gray) Pilg.	Hawaiian Archipelago (USA)	*Plantago*	?
*	*P. hillebrandii* Pilg.	Hawaiian Archipelago (USA)	*Plantago*	
	P. incisa Hassk.	Java (Indonesia)	*Plantago*	?
	P. komarovii Pavlov	C Asia	*Mesembrynia*	?

*	*P. krajinai* Pilg.	Hawaiian Archipelago (USA)	*Plantago*	
*	*P. melanochrous* Pilg.	Hawaiian Archipelago (USA)	*Plantago*	
*	*P. muscicola* Pilg.	Hawaiian Archipelago (USA)	*Plantago*	
	P. pachyphylla A.Gray	Hawaiian Archipelago (USA)	*Plantago*	24
	P. princeps Cham. & Schltdl.	Hawaiian Archipelago (USA)	*Plantago*	12
	P. rapensis Pilg.	Rapa Iti Island (France)	*Plantago*	?
	P. rugelii Decne.	E North America	*Plantago*	24
	P. rupicola Pilg.	Rapa Iti Island (France)	*Plantago*	?
	P. sparsiflora Michx.	SE North America	*Plantago*	24
!	*P. taquetii* H.Lév.	Jeju (Korea)	*Plantago*	?
	P. tweedyi A.Gray	W North America	*Plantago*	24
Sect. *Plantago* — clade F — 5 species				
	P. cornutii Gouan	S Europe	*Plantago*	12
!	*P. griffithii* Decne.	S Asia	*Plantago* (as a synonym)	?
	P. himalaica Pilg.	S Asia	*Plantago*	?
*	*P. tatarica* Decne.	S Asia	*Plantago* (as a synonym)	?
	P. major L.	Eurasia	*Plantago*	12
Sect. *Virginica* Decne. & Steinh. ex Barnéoud — clade I — 46 species				
	P. alismatifolia Pilg.	C Mexico	*Virginica*	24
*	*P. argentina* Pilg.	NW Argentina	*Virginica*	24, 48
	P. australis Lam.	South America and S North America	*Virginica*	24, 48
!	*P. barbata* G.Forst.	S Argentina and S and C Chile	*Oliganthos*	48, 72
*	*P. berroi* Pilg.	Uruguay and E Argentina	*Virginica*	24
	P. bradei Pilg.	E Brazil	*Virginica* (as a synonym)	?
*	*P. buchtienii* Pilg.	W Bolivia and NW Argentina	*Virginica*	48
	P. catharinea Decne.	S Brazil	*Virginica*	24
	P. commersoniana Decne. & Barnéoud	Uru., S Brazil and SE Par.	*Virginica*	?
*	*P. correae* Rahn	S Argentina and S Chile	*Oliganthos*	96
	P. corvensis Hassemer	S Brazil	none (new species)	?
	P. cumingiana Fisch. & C.A.Mey.	S and C Chile and W Argentina	*Virginica* (as a synonym)	?
*	*P. dielsiana* Pilg.	E Argentina and S Uruguay	*Virginica*	24
	P. fernandezia Bertero ex Barnéoud	Juan Fernández Islands (Chile)	*Plantago*	?
*	*P. firma* Kunze ex Walp.	C Chile	*Virginica*	24
	P. floccosa Decne.	NE Mexico	*Virginica*	24
*	*P. galapagensis* Rahn	Galápagos Islands (Ecuador)	*Virginica*	?
	P. guilleminiana Decne.	S Brazil	*Virginica*	?

	P. hatschbachiana Hassemer	S Brazil	none (new species)	?
	P. humboldtiana Hassemer	S Brazil	none (new species)	?
*	*P. jujuyensis* Rahn	NW Argentina	*Virginica*	24
	P. moorei Rahn	West Falkland (UK)	*Oliganthos*	?
	P. myosuros Lam.	S and W South America	*Virginica*	24
	P. napiformis (Rahn) Hassemer	NE Arg., Par. and S Brazil	*Virginica* (as a subspecies)	?
*	*P. orbignyana* Steinh. ex Decne.	Ecu., Peru, Bol. and NW Arg.	*Virginica*	24, 48
	P. pachyneura Steud.	N and C Chile	*Virginica*	24
	P. penantha Griseb.	Uru., NE Arg. and S Brazil	*Virginica*	24
	P. pretoana (Rahn) Hassemer	SE Brazil	*Virginica* (as a subspecies)	?
	P. pulvinata Speg.	S Argentina and S Chile	*Oliganthos*	24
	P. pyrophila Villarroel & J.R.I.Wood	E Bolivia	none (new species)	?
*	*P. rahniana* Hassemer & R.Trevis.	S Brazil	none (new species)	?
*	*P. rhodosperma* Decne.	SW USA and NE Mexico	*Virginica*	48
*	*P. sempervivoides* Dusén	S Argentina and S Chile	*Oliganthos*	?
*	*P. subnuda* Pilg.	W USA	*Virginica*	48
	P. tehuelcha Speg.	S Argentina and S Chile	*Oliganthos*	24
	P. tenuipala (Rahn) Rahn	C Colombia	*Virginica*	?
*	*P. tomentosa* Lam.	Arg., Bol., Par., Uru. and S Brazil	*Virginica*	24
	P. trinitatis Rahn	Trindade Island (Brazil)	*Virginica*	?
	P. truncata Cham. & Schltdl.	C Chile	*Virginica*	?
*	*P. turficola* Rahn	S Brazil	*Virginica*	?
	P. uniglumis Wallr. ex Walp.	S Argentina and S and C Chile	*Oliganthos*	48, 72
	P. veadeirensis Hassemer	C Brazil	none (new species)	?
	P. ventanensis Pilg.	E Argentina	*Virginica*	24
*	*P. venturii* Pilg.	W Argentina	*Virginica*	24
	P. virginica L.	North America	*Virginica*	24
*	*P. weddelliana* Decne.	S Bolivia and NW Argentina	*Virginica*	24
Incertae sedis — 1 species				
*	*P. robusta* Roxb.	Saint Helena Island (UK)	*Plantago*	?

Plantago subg. *Plantago*

Lectotype (designated by Britton and Brown 1913: 245): *P. major* L.

Our results corroborate, with a strong support (BS=100%; PP=1.00), the already well-established perception (Rønsted et al., 2002; Hoggard et al., 2003;

Ishikawa et al., 2009) that subg. *Plantago* is monophyletic. We estimate that there are 147 species in the subgenus (Table 3), although this number certainly will change in the future, with new species being described, species being re-established and names being synonymised as taxonomic knowledge of the group advances.

***Plantago* sect. *Micropsyllium* Decne.** in A.DC., Prodr. 13(1): 696. 1852
Lectotype (designated by Dietrich 1980: 563): *P. tenuiflora* Waldst. & Kit.

= *Plantago* sect. *Diandra* H.Dietr., Wiss. Z. Friedrich-Schiller-Univ. Jena, Math.-Naturwiss. Reihe 29(4): 563. 1980
Holotype: *P. elongata* Pursh

Plantago sect. *Micropsyllium* comprises six species from North America and Eurasia (Table 3), is morphologically well defined (Bassett 1966; Rahn 1996) and is the only section within subg. *Plantago* that did not undergo any changes in this study. The three sampled species form a well-supported clade (BS=100%; PP=1.00; Fig. 4, clade A).

This section appears to have originated in Eurasia and subsequently colonised North America. In terms of morphology, sect. *Micropsyllium* is characterised by the following apomorphic characters: diminutive annual plants; annual root; hairs on scapes antrorse; corolla lobes shorter than 1 mm; anthers less than 1 mm long; seeds shorter than 2 mm. One important plesiomorphic character shared by all species is the linear leaves. Reported chromosome numbers are variable, including $2n = 12$, 24, 36 and 20 (this last one, reported by Bassett (1966) for *P. bigelovii*, is the only record of $x = 5$ in the subg. *Plantago*, and needs to be confirmed).

Selected taxonomic references: Pilger (1937), Bassett (1966) and Moore et al. (1976).

***Plantago* sect. *Eremopsyllium* Pilg.**, Pflanzenr. 102: 283. 1937
Holotype: *P. reniformis* Beck

= *Plantago* sect. *Gentianoides* Pilg., Pflanzenr. 102: 306. 1937, ***syn. nov.***
Holotype: *P. gentianoides* Sibth. & Sm.

The two species included in the hereby re-established sect. *Eremopsyllium* (Table 3) were both sampled here and form a well-supported clade (BS=97%; PP=1.00; Fig. 4, clade B) that is sister to all the rest of subg. *Plantago* except for sect. *Micropsyllium*. These species are distributed in southeastern Europe and Asia Minor. Here, we do not accept *P. griffithii* as a subspecies of *P. gentianoides* (see Hassemer 2018). In terms of morphology, sect. *Eremopsyllium* is characterised by the following apomorphic characters: adventitious roots; spike less than 1/3 the length of the scape; anthers white both when fresh and when dried. Reported chromosome numbers are $2n = 12$.

Selected taxonomic references: Pilger (1937), Moore et al. (1976) and Tutel (1982).

***Plantago* sect. *Lamprosantha* Decne.** in A.DC., Prodr. 13(1): 697. 1852
Lectotype (designated by Rahn 1996): *P. media* L.

The hereby re-established sect. *Lamprosantha* (Table 3) includes six species, all of which were sampled here: three species from Rahn's (1996) sect. *Plantago* (*P. canescens*, *P. maxima* and *P. media*) and three from sect. *Mesembrynia* (*P. arachnoidea*, *P. perssonii* and *P. schwarzenbergiana*). This section is distributed in temperate Eurasia, with the exception of *P. arachnoidea* which occurs in northern Asia and northwestern North America. This section forms a well-supported clade (BS=100%; PP=1.00; Fig. 4, clade C) which

includes the rare and threatened *P. maxima* as sister to a clade (BS=100%; PP=1.00) which includes the remainder of the species included in the section.

It should be noted that the support for the clade that includes sect. *Lamprosantha* plus all subsequent groups within subg. *Plantago* is relatively low (BS=52%; PP=0.95). It is possible that further investigations could indicate that sects. *Eremopsyllium* and *Lamprosantha* are sisters, in which case their merger into an enlarged sect. *Lamprosantha* would be desirable due to morphology and biogeography. However, our plastome phylogenies do not support this (Fig. 4), and therefore we recognise these two sections as distinct.

This section apparently has a temperate Eurasian origin. In terms of morphology, sect. *Lamprosantha* is characterised by the following apomorphic characters: spike less than 1/3 the length of the scape; seeds shorter than 2 mm. Reported chromosome numbers are $2n = 12$, with the exception of a few populations of *P. media*, which have $2n = 24$.

Selected taxonomic references: Pilger (1937), Grigoriev (1958), Moore et al. (1976) and Li et al. (2011).

***Plantago* sect. *Leptostachys* Decne.** in A.DC., Prodr. 13(1): 720. 1852
Holotype: *P. leptostachys* E.Mey. ex Decne., *nom. illeg.*, *non* Hook.f. (1847), *nec* Ledeb. (1849) — = *P. laxiflora* Decne.

= *Plantago* sect. *Palaeopsyllium* Pilg., Pflanzenr. 102: 75. 1937, ***syn. nov.***
Lectotype (**designated here**): *P. palmata* Hook.f.

The seven species recognised in the hereby re-established sect. *Leptostachys* (Table 3) occur in sub-Saharan Africa and Madagascar, comprise the most tropical among the species within subg. *Plantago,* and comprise species included in Rahn's (1996) sect. *Plantago*. The three sampled species (*P. longissima*, *P. palmata* and *P. tanalensis*) form a clade which has a high support

(BS=100%; PP=1.00; Fig. 4, clade D). Taxonomic knowledge of this section is the poorest among subg. *Plantago* and a taxonomic revision is critically needed, especially to clarify questions regarding *P. leptostachys* E.Mey. ex Decne. *nom. illeg.*, which is the type species of the section, and its supposedly accepted name, *P. laxiflora* (fide Pilger, 1937), which was not sampled here.

This section appears to have originated in continental sub-Saharan Africa and subsequently expanded to Madagascar. In terms of morphology, sect. *Leptostachys* is characterised by the following apomorphic characters: adventitious roots; corolla lobes shorter than 1.5 mm. One important plesiomorphic character shared by all species is the spikes normally equalling the length of the scape. Chromosome numbers are unknown for all species except for *P. palmata*, which has $2n = 24$.

Selected taxonomic references: Pilger (1937) and Verdcourt (1971).

***Plantago* sect. *Mesembrynia* Decne.** in A.DC., Prodr. 13(1): 701. 1852
Lectotype (designated by Rahn 1996): *P. debilis* R.Br.

= *Plantago* sect. *Microcalyx* Pilg., Pflanzenr. 102: 122. 1937, ***syn. nov.***
Lectotype (designated by Rahn 1996): *P. triandra* Berggr.

Plantago sect. *Mesembrynia* is hereby accepted as including 44 species (Table 3), although this number will probably increase in the future with the discovery of new species. Of these 44 species, 23 were sampled here which form a well-supported clade (BS=100%; PP=1.00; Fig. 4, clade E). This section encompasses all *Plantago* species native to Australia, New Zealand and New Guinea, and also species from some neighbouring islands (Auckland Islands, Lord Howe Island and Tonga), in addition to two species from St. Paul and New Amsterdam Islands in the southern Indian Ocean. Our expanded sect. *Mesembrynia* includes all species in Rahn's (1996) homonymous section except

for the six Eurasian species (*P. arachnoidea*, *P. camtschatica*, *P. depressa*, *P. komarovii*, *P. perssonii* and *P. schwarzenbergiana*), 15 species from his sect. *Oliganthos*, plus two island endemics (*P. aucklandica* and *P. hedleyi*) from his sect. *Plantago*.

Our cpDNA tree topology is similar but not identical to a previous phylogenetic study of the Australasian species, which analysed ITS, cpDNA and mtDNA sequences (Tay et al., 2010a). Our results indicate that at least four long-distance dispersal events occurred for the New Zealand *Plantago* species, likely originating from Australian ancestors, thus corroborating the conclusions of Tay et al. (2010a) who indicated three such events. One of these dispersal events comprises the clade of *P. lanigera* and *P. novae-zelandiae*, the second dispersal event comprises *P. unibracteata*, the third comprises *P. triandra*, and the fourth comprises the clade of *P. picta*, *P. raoulii*, *P. spathulata* and *P. udicola*. The small oceanic island species *P. euana*, *P. hedleyi* and *P. stauntonii* are spread across sect. *Mesembrynia*, again likely originating from Australian ancestors.

Although we attempted sequencing of samples of two species from New Guinea, *P. aundensis* and *P. papuana*, these were not able to be included here due to highly degraded DNA. In accordance with the accumulated biogeographic and morphological knowledge of these species (Craven, 1976; van Royen, 1983; Rahn, 1996), the seven *Plantago* species from New Guinea are here grouped together with the other Australasian species. However, we advocate that future phylogenetic works should focus on sampling New Guinean species, in order to confirm their phylogenetic placement and fully understand the biogeography of this Southern Hemisphere section.

Our results show that the two sampled individuals of *P. lanigera*, although in the same lineage, are paraphyletic relative to the closely related species, *P. novae-zelandiae* (BS=100%; PP=1.00). Because our study used cpDNA only on very few (1–2) individuals per species, it is not possible to

speculate further on the taxonomic implications of this finding due to the complex polyploid evolutionary history of the New Zealand species (Meudt, 2011, 2012; Ishikawa et al., 2009). Nevertheless, the plants from Sewell Peak, from which the sample "*P. lanigera* 2" was collected, will be further investigated in regards to their morphology and will be compared with the types of other New Zealand *Plantago*, as it is possible that this sample could correspond to a still undescribed species.

This section appears to have originated in Australia and subsequently spread to New Zealand and other neighbouring islands in several separate dispersal events (Tay et al., 2010a). However, it should be noted that the lack of sampling of New Guinean species precludes the inference of the centre of origin of this section. In terms of morphology, apparently there is not a single apomorphic character that is shared by all species in sect. *Mesembrynia*. This is certainly the reason why the species included in the section have never before been all placed under the same section—there are considerable morphological differences between the species formerly placed in sects. *Mesembrynia* and *Oliganthos*, which is not reflected in the phylogeny. Some plesiomorphic characters seem to be shared by all the species in the section, such as: lamina with attenuate base or not distinguishable from the petiole; apex of the leaves without a colourless acumen; pedicel absent; corolla lobes always patent; anthers never white; ovary with more than 4 ovules; carpophore absent. Reported chromosome numbers are $2n$ = 12, 24, 36, 48, 60, 72 and 96, with 12 being the most common and apparently the ancestral condition.

Selected taxonomic references: Briggs et al. (1973, 1977), Craven (1976), Briggs (1980), van Royen (1983) and Meudt (2012).

Plantago* sect. *Plantago

Type species: *P. major* L.

= *Plantago* sect. *Major* Barnéoud, Rech. Plantagin. Plumbagin.: 17. 1844
Holotype: *P. major* L.

= *Plantago* sect. *Polyneuron* Decne. in A.DC., Prodr. 13(1): 694. 1852
Lectotype (designated by Rahn 1996): *P. major* L.

The hereby much reduced sect. *Plantago* (Table 3), with only five species, is without doubt the most unexpected and drastic change departing from Rahn's (1996) classification, who recognised 53 species (although with much doubt) in his admittedly paraphyletic sect. *Plantago*. The three species sampled here (*P. cornutii*, *P. himalaica* and *P. major*) form a clade, which is well supported (BS=100%; PP=1.00; Fig. 4, clade F). The five species in this section are all originally distributed in temperate Eurasia. The type species of the genus, subg. and sect. *Plantago*, *P. major*, is now a cosmopolitan species with a Eurasian origin. The morphologically similar *P. cornutii*, from southern Europe, and *P. griffithii*, *P. himalaica* and *P. tatarica*, from southern Asia (see Hassemer, 2018), are also included in this section.

This section appears to have its origin in southwestern Eurasia. In terms of morphology, sect. *Plantago* is characterised by the following apomorphic characters: adventitious roots; lamina more than four times as wide as the petiole; leaves remaining green on drying; hairs on scape antrorse; spur-like elongation on lowermost cell of non-glandular hairs on the scape; corolla lobes shorter than 2 mm; pyxidia globose, not conspicuously elongated. Chromosome numbers are known for two species (*P. cornutii* and *P. major*), $2n = 12$.

Selected taxonomic references: Pilger (1937), Moore et al. (1976) and Hassemer (2018).

***Plantago* sect. *Carpophorae* Rahn**, Nordic J. Bot. 5: 144. 1985
Holotype: *P. rigida* Kunth

The hereby re-established sect. *Carpophorae* (Table 3) comprises two species from mountains in Central America and western South America (Rahn, 1985). Both species sampled here form a clade that is well supported (BS=100%; PP=1.00; Fig. 4, clade G) and is sister to *P. cordata* (sect. *Heptaneuron*)—although with rather low support (BS=55%; PP=0.73). *Plantago* sect. *Carpophorae* is notable among *Plantago* for producing a carpophore (Rahn, 1985). The section was created with this exact circumscription by Rahn (1985), who later (Rahn 1996) changed this taxon to the series level, *Plantago* ser. *Carpophorae* (Rahn) Rahn, under *Plantago* sect. *Oliganthos*, because of the numerical results of his morphological phylogeny.

From a biogeographic perspective, it would be expected that the species in sect. *Carpophorae* would be closely related to the predominantly South American sect. *Virginica* (see below). From a morphological perspective, however, the two species in sect. *Carpophorae* could well be placed in a separate subgenus, given their unique, very distinct fruit morphology (Rahn, 1985). The inflorescences with very few flowers (normally 1, rarely 2–3) led to the two species in sect. *Carpophorae* having been included in sect. *Oliganthos* in most classification systems (e.g. Pilger, 1937; Rahn, 1996).

It would be extremely undesirable to unite the species in sects. *Carpophorae* (*P. rigida* and *P. tubulosa*) and *Heptaneuron* (*P. cordata*) in a single section, due to the very distinctive morphology of the two species in sect. *Carpophorae*, which is unique in *Plantago* and not coherent with that of *P. cordata*. Of less importance is the fact that *P. cordata* is a semi-aquatic plant endemic to eastern North America, whereas the species in sect. *Carpophorae* are endemic to mountains in western South America and southern North America (southern Mexico southwards). Although our phylogeny would allow the merger of these two sections, it would also allow their recognition as distinct, and we believe that classification systems based on molecular

phylogenies should nevertheless be coherent from a morphological point of view.

It would have been possible, based on our phylogeny, to include the clade that encompasses sects. *Heptaneuron* and *Carpophorae* in a more broadly defined sect. *Plantago*—this clade of seven species is strongly supported (BS=100%; PP=1.00). Nevertheless, as explained above, this enlarged circumscription would be extremely undesirable because of considerable morphological differences of the two species in sect. *Carpophorae*. Therefore, we opted to recognise sect. *Carpophorae*, and as a consequence it was also necessary to recognise the monotypic sect. *Heptaneuron* (see below). However, due to the low support (BS=55%; PP=0.7306) of the clade containing sects. *Carpophorae* + *Heptaneuron*, further phylogenetic investigation is critically needed to elucidate this group of sections. If improved phylogenies would in the future show that sects. *Heptaneuron* and *Plantago* are sister, they should probably be merged. In any case, because of morphology, the species in these two sections should not be merged with sect. *Carpophorae*.

This section appears to have originated in the Andes and subsequently colonised southern North America. In terms of morphology, sect. *Carpophorae* is well-characterised by the following apomorphic characters: adventitious roots; scape very short, less than a quarter of the supporting leaf; hairs on scape antrorse; small, three-celled, glandular hairs placed in cavities; flower solitary, only one bract present; anthers longer than 2 mm long; carpophore present. Reported chromosome numbers are $2n = 24$, 48 and 72.

Selected taxonomic reference: Rahn (1985).

***Plantago* sect. *Heptaneuron* Decne.** in A.DC., Prodr. 13(1): 698. 1852
Lectotype (designated by Rahn 1996): *P. cordata* Lam.

The hereby re-established sect. *Heptaneuron* (Table 3) is monotypic, including only the semi-aquatic eastern North American *P. cordata*, previously in Rahn's (1996) sect. *Plantago*. *Plantago* sect. *Heptaneuron* is sister to sect. *Carpophorae* (BS=55%; PP=0.73) and, due to considerable morphological differences, should not be united with the latter (see explanation above). *Plantago cordata* is unique in the genus in that the fruits are still green and alive at the time of dehiscence, when the lid of the pyxidia readily falls off, and the seeds with the entire fleshy placenta fall out as a unit; this structure is buoyant and may represent an adaptation to dispersal by water (Tessene, 1969; Rosatti, 1984). In terms of morphology, sect. *Heptaneuron* is characterised by the following apomorphic characters: lamina less than 1.9 times as long as wide; lamina more than 4 times as wide as the petiole; base of lamina truncate; corolla lobes shorter than 1.5 mm; anthers longer than 2 mm long; ovary with four ovules. The chromosome number of *P. cordata* is $2n = 24$.

Selected taxonomic reference: Pilger (1937) and Bassett (1973).

Plantago* sect. *Pacifica* Hassemer**, ***sect. nov.

Diagnosis: plants perennial; apex of the leaves without a colourless acumen; scape length at least more than a quarter of the supporting leaf; scape not elongating conspicuously after anthesis; trichomes on leaves up to 2 mm long; trichomes on leaves more than 0.04 mm wide; absence of small, three-celled, glandular hairs placed in cavities; normal spike with 12 flowers or more; sepals glabrous on the back; corolla lobes always patent; corolla lobes up to 3 mm long; stamens 4; anthers never white; anthers longer than 0.5 mm; carpophore absent; ovary with more than 4 ovules; mature pyxidia pyriform, elongated; seeds shorter than 3 mm.

Holotype: *P. princeps* Cham. & Schltdl.

This new section, which we estimate to include 26 species, 15 of which were sampled here (Table 3), corresponds to a well-supported clade (BS=100%; PP=1.00; Fig. 4, clade H) sister to sects. *Holopsyllium* + *Virginica*. All species were formerly placed in Rahn's (1996) sect. *Plantago*, except for three species from his sect. *Mesembrynia* (*P. camtschatica*, *P. depressa* and *P. komarovii*). Because this lineage does not include *P. major*, the type species of sect. *Plantago*, it required a new name at section level. A thorough sampling, especially of the Eurasian species in subg. *Plantago*, is needed to confirm the circumscriptions of sects. *Pacifica* and *Plantago*. The name of the new section is a reference to the distribution of its species in Asia and North America, i.e. at both sides of the Pacific Ocean, and also in some Pacific oceanic islands (Hawaiian Archipelago and Rapa Iti Island).

The central Asian *P. komarovii* is sister to the remainder of the species in the section (BS=100%; PP=1.00). This clade then splits into an Asian clade (BS=100%; PP=1.00) including *P. asiatica*, *P. camtschatica*, *P. cavaleriei*, *P. depressa* and *P. incisa*, and another clade (BS=100%; PP=1.00) comprising North American and the oceanic species mentioned above. Our phylogeny indicates that the Hawaiian (*P. glabrifolia*, *P. grayana*, *P. hawaiensis*, *P. hillebrandii*, *P. krajinai*, *P. melanochrous*, *P. muscicola*, *P. pachyphylla* and *P. princeps*) and Rapa Iti Island (*P. rapensis* and *P. rupicola*) species in this section originated from North American ancestors, as all these island species are included in the clade that also includes the North American *P. eriopoda*, *P. rugelii*, *P. sparsiflora* and *P. tweedyi*. The closest living relatives of the Hawaiian and Rapa Iti *Plantago* were indicated, in a strongly-supported clade (BS=100%; PP=1.00), to be *P. rugelii* and *P. sparsiflora*.

Based on morphology and biogeography, it would appear that *P. alata*, *P. coreana* and *P. taquetii*, all described from Jeju Island, could be synonyms of *P. asiatica*—however, a comprehensive taxonomic revision of the Korean *Plantago* is needed to confirm this. A comprehensive taxonomic treatment of

the Hawaiian *Plantago* is also urgently needed, as the synonymisation of *P. glabrifolia*, *P. grayana*, *P. hillebrandii*, *P. krajinai*, *P. melanochrous* and *P. muscicola* under *P. pachyphylla* done by Wagner et al. (1990) appears weakly supported from a morphological point of view (G. Hassemer, pers. obs.), which may have been why Rahn (1996) decided to keep these six species in his phylogenetic study. Furthermore, there is phylogenetic and morphological evidence (Dunbar-Co et al., 2008, 2009) that there are more species in the Hawaiian Archipelago than currently recognised. We are here following the treatment of Pilger (1937) regarding the Hawaiian *Plantago*, because it seems to better reflect the specific diversity in this group than the treatment of Wagner et al. (1990). Such as occurred with some other plant groups such as Asteraceae (Baldwin and Sanderson, 1998; Knope et al., 2012) and Campanulaceae (Givnish et al., 2009), we believe it possible that a great diversification occurred upon the arrival of *Plantago* from North America to Hawaii, due to the abundance of unoccupied niches. The taxonomic resolution of Hawaiian *Plantago* is critical because of implications it would have for the conservation of narrowly endemic species, but also for allowing a better understanding of sect. *Pacifica*.

This section appears to have its origin in central and eastern Eurasia, and subsequently colonised North America, and from there it spread to the Hawaiian archipelago and Rapa Iti Island. In terms of morphology, apparently there is not a single apomorphic character that is shared by all species in sect. *Pacifica*. Some plesiomorphic characters seem to be shared by all the species in the section, such as: plants perennial; apex of the leaves without a colourless acumen; scape length at least more than a quarter of the supporting leaf; scape not elongating conspicuously after anthesis; trichomes on leaves up to 2 mm long; trichomes on leaves more than 0.04 mm wide; absence of small, three-celled, glandular hairs placed in cavities; normal spike with 12 flowers or more; sepals glabrous on the back; corolla lobes always patent; corolla lobes up to 3

mm long; stamens 4; anthers never white; anthers longer than 0.5 mm; carpophore absent; ovary with more than 4 ovules; mature pyxidia pyriform, elongated; seeds shorter than 3 mm. Reported chromosome numbers are $2n$ = 12 and 24.

Selected taxonomic references: Pilger (1937), Grigoriev (1958), Bassett (1973), Wagner et al. (1990), Yamazaki (1993) and Li et al. (2011).

***Plantago* sect. *Holopsyllium* Pilg.**, Pflanzenr. 102: 101. 1937
Holotype: *P. macrocarpa* Cham. & Schltdl.

This monotypic section (Table 3), hereby re-established, is sister to the predominantly South American sect. *Virginica* (BS=100%; PP=1.00). Its only species, *P. macrocarpa*, occurs on the coast of northwestern North America, the Aleutian archipelago and the Commander Islands (Russia). The uniqueness of several morphological characters of *P. macrocarpa* has already been evidenced by Pilger (1937), of which the most prominent are the indehiscent pyxidia. Because of pronounced morphological differences, it would be undesirable to merge sect. *Holopsyllium* with its sister, sect. *Virginica*, and for this reason both sections are accepted. In terms of distribution and phylogeny, sect. *Holopsyllium* could perhaps be a testimony of the crossing of subg. *Plantago* from Eurasia to the Americas. In terms of morphology, sect. *Holopsyllium* is characterised by the following apomorphic characters: anterior sepals distinctly narrower than the posterior, and differently shaped; corolla lobes shorter than 2 mm; ovary with two ovules, and no rudiment of an upper compartment; fruit an indehiscent pyxidium; seeds longer than 3 mm. The chromosome number of *P. macrocarpa* is $2n$ = 24.

Selected taxonomic references: Pilger (1937), Grigoriev (1958) and Bassett (1973).

***Plantago* sect. *Virginica* Decne. & Steinh. ex Barnéoud**, Rech. Plantagin. Plumbagin.: 17. 1844
Holotype: *P. virginica* L.

= *Plantago* sect. *Cleiosantha* Decne. in A.DC., Prodr. 13(1): 721. 1852
Lectotype (designated by Rahn 1996): *P. veratrifolia* Decne. — = *P. australis* subsp. *hirtella* (Kunth) Rahn

= *Plantago* sect. *Dendriopsyllium* Decne. in A.DC., Prodr. 13(1): 704. 1852, ***syn. nov.***
Lectotype (designated by Rahn 1996): *P. fernandezia* Bertero ex Barnéoud

= *Plantago* sect. *Fernandezia* Barnéoud, Rech. Plantagin. Plumbagin.: 19. 1844, ***syn. nov.***
Holotype: *P. fernandezia* Bertero ex Barnéoud

= *Plantago* sect. *Novorbis* Decne. in A.DC., Prodr. 13(1): 724. 1852
Lectotype (designated by Rahn 1996): *P. tomentosa* Lam.

= *Plantago* sect. *Oliganthos* Barnéoud, Rech. Plantagin. Plumbagin.: 17. 1844, ***syn. nov.***
Lectotype (designated by Rahn 1984): *P. pauciflora* Lam. — = *P. barbata* G.Forst.

= *Plantago* sect. *Oreophytum* Decne. in A.DC., Prodr. 13(1): 704. 1852
Holotype: *P. orbignyana* Steinh. ex Decne.

= *Plantago* sect. *Plantaginella* Decne. in A.DC., Prodr. 13(1): 727. 1852, ***syn. nov.***

Lectotype (designated by Rahn 1984): *P. barbata* G.Forst.

With 46 recognised species, one of which (*P. cumingiana*) hereby re-established (see below), our enlarged sect. *Virginica* (Table 3) is sister to the monotypic sect. *Holopsyllium*, which has distinct fruit morphology and distribution (see above). The clade of sect. *Virginica* is well-supported in our phylogeny (BS=87%; PP=1.00; Fig. 4, clade I), and includes all species in Rahn's (1996) homonymous section, in addition to the seven American species in series *Oliganthos*, and *P. fernandezia*, which was previously placed in sect. *Plantago*. This predominantly South American clade has two centres of diversity: one in high-elevation grasslands and open coastal vegetation of central-eastern South America, including Uruguay, southern Brazil and eastern Argentina, and another in moist rocky environments of southern South America, which includes the American species in Rahn's (1996) series *Oliganthos*.

In *Plantago* sect. *Virginica*, our phylogeny indicated an early split between a clade (BS=100%; PP=1.00) including the southern South American *P. fernandezia* and *P. tehuelcha*, and another clade (BS=100%; PP=1.00) including the remainder of species sampled, including the West Falkland endemic *P. moorei*. Based on our phylogeny it is impossible to infer the position of the five unsampled American species in Rahn's (1996) series *Oliganthos* (*P. barbata*, *P. correae*, *P. pulvinata*, *P. sempervivoides* and *P. uniglumis*) between these two possible early branches within sect. *Virginica*. Similarly to the case of another large section, i.e. sect. *Mesembrynia*, we consider that sect. *Virginica* constitutes a phylogenetically and biogeographically coherent unit whose splitting would be undesirable because the species previously recognised in sect. *Oliganthos* (due to distinct morphology) are spread in multiple branches through the phylogeny of the section.

This section appears to have its origin in southern South America, and subsequently expanded to the rest of South America and also to North America. In terms of morphology, apparently there is not a single apomorphic character that is shared by all species in sect. *Virginica*. This is certainly the reason why the species included in the section have never before been all placed under the same section—there are considerable morphological differences between the species formerly placed in sects. *Virginica* and *Oliganthos*, which is not reflected in the phylogeny. Some plesiomorphic characters seem to be shared by all the species in the section, such as: nerves of dead leaf never remaining on the plant as long bristles; lamina with attenuate base or not distinguishable from the petiole; scape not elongating conspicuously after anthesis; spike open and cylindrical, the rachis visible between the flowers; pedicel absent; corolla lobes longer than 1 mm; stamens 4; anthers never white; carpophore absent; mature pyxidia pyriform, elongated. Reported chromosome numbers are $2n = 24$, 48, 72 and 96, with 24 being the most common and apparently the ancestral condition.

Our phylogenetic results indicated that the current concept of *P. catharinea* is polyphyletic; this taxonomic problem, which is caused by some populations of *P. napiformis* being misidentified as *P. catharinea*, was discussed in detail in Hassemer (2019). We also highlight that our sampling of *P. australis*, albeit limited considering its continental distribution with eight subspecies currently recognised, clearly indicated that one of its subspecies, i.e. *P. australis* subsp. *cumingiana*, is polyphyletic in relation to the rest of the sampled subspecies, which formed a monophyletic clade. The monophyly of the remainder of the sampled subspecies of *P. australis* does not disagree with the current taxonomic treatment of the species (Rahn, 1974; Hassemer et al., 2015), but also does not necessarily agree with the recognition of these taxa at the subspecies rank. The resolution of the *P. australis* complex will require an extensive sampling of populations encompassing all subspecies and preferably

all environmentally distinct regions where it occurs, coupled with comprehensive morphological and nomenclatural knowledge of the group.

Selected taxonomic references: Rahn (1974, 1984), complemented with novelties in Villarroel and Wood (2011), Hassemer and Baumann (2014), Hassemer et al. (2014, 2015); Hassemer (2016, 2017, 2019) and Hassemer and Rønsted (2016).

Species *incertae sedis*:

P. robusta Roxb.

Based on our results, it is not possible to ascertain the phylogenetic position of *P. robusta*—unfortunately, the sample of this species that we sequenced was contaminated with a species belonging to *Plantago* subg. *Coronopus*. This species is endemic to Saint Helena, a small (122 km^2) South Atlantic oceanic island more than 2,000 kilometres from the nearest major landmass (Africa). This species has aerial woody stems like other oceanic island endemics such as *P. fernandezia* and *P. trinitatis*. However, based on morphology we cannot infer its phylogenetic placement. Therefore, new sampling of this species, preferably from living specimens, is necessary.

Revalidation of *Plantago cumingiana*

***Plantago cumingiana* Fisch. & C.A.Mey.**, Index Seminum [St. Petersburg] 3: 44–45. 1837

≡ *Plantago australis* subsp. *cumingiana* (Fisch. & C.A.Mey.) Rahn, Bot. Tidsskr. 60: 48–49. 1964

Lectotype (or maybe neotype, designated by Rahn 1964): CHILE. S.d., *H. Cuming s.n.* (LE-00016458! [Fig. S4]).

In all plastome trees *P. australis* subsp. *cumingiana* did not form a clade with the three other sampled subspecies of *P. australis* (*P. australis* subsp. *australis*, *P. australis* subsp. *hirtella* and *P. australis* subsp. *leioloma*). In the plastome tree, *P. australis* subsp. *cumingiana* is sister (BS=52%, PP=0.98) to a clade (BS=49%, PP=0.77) which includes, among other species, *P. bradei* and *P. tomentosa* in addition to the other *P. australis* samples. Although the inclusion of *P. bradei* is not strongly supported, the next clade excluding it is very well supported (BS=100%, PP=1.00). This phylogenetic evidence, in addition to the study of several hundred specimens of *P. australis* from all over its distribution, has convinced us of the need for re-establishing this species, whose geographic dispersion does not overlap with the huge extent of occurrence of *P. australis*, the most common and widespread species in *Plantago* sect. *Virginica* (see Rahn, 1974).

Plantago cumingiana occurs in central (Valparaíso) to southern (Tierra del Fuego) Chile, and also in southwestern Argentina, in the western parts of the provinces of Chubut, Neuquén and Río Negro (see Rahn, 1974; Murillo, 2012). Some morphological differences from the other subspecies of *P. australis* have been observed during the revisions of herbarium collections, and also with cultivation experiments: a taproot is often present among cord-like secondary roots, the leaves have a slightly thicker consistency, the corollas are slightly longer, and the length/breadth ratio of the seeds is slightly less (slightly more globose-like than ellipsoid). However, *P. australis* is a morphologically very variable species, and outlier specimens exist for most of its subspecies, what makes us conclude that morphology alone is not enough to resolve the *P. australis* species complex. This could explain why Rahn (1964, 1974) decided to lump together over a dozen previously-accepted species in his enlarged concept of *P. australis*. Our results show that the three sampled subspecies (*P. australis* subsp. *australis*, *P. australis* subsp. *hirtella* and *P. australis* subsp. *leioloma*) cluster together in a single clade, indicating that it is possible that

they are conspecific, whereas *P. cumingiana* clearly constitutes a separate phylogenetic branch. Our results suggest that *P. australis* does not occur in Chile—all records of this species in this country are *P. cumingiana* instead. Inclusion of multiple accessions and the other synonymised previously recognised species, as well as nuclear molecular markers, is necessary to resolve the *P. australis* complex in the future.

Notes on the new classification

The hereby-proposed classification system for subg. *Plantago*, with 11 accepted sections, recognises considerably more sections than that of Rahn (1996), which accepted five sections (Table 1), but slightly less sections than Pilger (1937), who accepted 13 sections (i.e. sects. *Eremopsyllium*, *Gentianoides*, *Holopsyllium*, *Lamprosantha*, *Mesembrynia*, *Microcalyx*, *Micropsyllium*, *Novorbis*, *Oliganthos*, *Oreophytum*, *Palaeopsyllium* and *Polyneuron*) for the subgenus as we understand it. Compared to Rahn's (1996) classification, the most important changes were the transfer of the majority of the species of his admittedly non-monophyletic sect. *Plantago* to six other sections (i.e. sects. *Eremopsyllium*, *Heptaneuron*, *Holopsyllium*, *Lamprosantha*, *Leptostachys* and *Pacifica*) and the disintegration of his sect. *Oliganthos*, the species of which were transferred to sects. *Mesembrynia* and *Virginica* following a geographically coherent pattern.

From our revised classification it is evident that some morphological characters that have been used for classifying the species in subg. *Plantago* are not appropriate to this end, as they overlap across different clades. Examples of such are number of flowers in the inflorescences, and trichome and seed morphology. Other characters are more conserved across the phylogeny and therefore significant for infrageneric classification, namely flower (flower symmetry, number of stamens, and the flowers being hermaphroditic or not)

and fruit (fruit shape and dehiscence, and number of seeds) characters. Based on our results, we argue that morphology remains the most adequate tool for the discovery of new species of flowering plants, as molecular phylogenetic techniques are still very far from being universally available, and are not helpful when exploring the biodiversity out in the field or during herbarium revisions. Furthermore, it should be highlighted that the correct identification of specimens relies on morphology, and phylogenies based on misidentified specimens are very detrimental to science. Regarding chromosome numbers, some sections are relatively homogenous, while others present wider variation (e.g. sects. *Carpophorae*, *Mesembrynia*, *Micropsyllium* and *Virginica*).

The considerably reduced morphology within subg. *Plantago*, and the fact that molecular phylogeny evidenced that most characters formerly used to distinguish sections are variable within and overlap between different sections, has convinced us that an attempt to produce an identification key to the sections of subg. *Plantago* would most probably result in an impractical and unusable key, thus thwarting the purpose of an identification key which is to facilitate the identification of specimens by non-specialists. For this reason, we do not provide such a key here. The identification of specimens of *Plantago* requires the consultation to specialised taxonomic works and regional floras, which have paramount importance for the advancement of the taxonomic knowledge.

4.2. The application of molecular phylogeny to classification

The final alignment of the plastome dataset included 96 samples encompassing 88 species providing a significant improvement compared to previous studies (from 40 species, ~28% of the subgenus previous, to 83 species, ~60% here) of *Plantago* subg. *Plantago*. Ten additional samples were sequenced, but could not be included due to too much degradation of DNA or contamination. Although HTS approaches have proven very efficient in

obtaining DNA from even highly degraded species in general compared to Sanger sequencing, difficulties in obtaining samples of sufficient DNA quality of rare and rarely collected species remains a problem. However, overall our phylogenetic results, combined with insights from the extensive herbarium and literature revision, has evidenced that HTS is a very promising tool to support the resolution of taxonomic problems and we have here been able to propose a new sectional classification of the taxonomically difficult subg. *Plantago*.

Our newly proposed classification departs considerably, in many aspects, from all previous major classification systems for *Plantago*, all based on morphology: Barnéoud (1844, 1845), Decaisne (1852), Pilger (1937) and Rahn (1978, 1996). Some of the proposed changes to the most recent and currently accepted system, Rahn (1996), re-established aspects from previous classifications, including Rahn's previous ideas, as in the case of sect. *Carpophorae.* This had already happened before, when Rønsted et al. (2002) indicated that *Plantago* subg. *Albicans* Rahn was paraphyletic to subg. *Psyllium* and argued for its merging with subg. *Psyllium* as discussed above. *Plantago* subg. *Albicans* was described by Rahn (1996) as result of his phylogeny based on morphology, and was a departure from his previous proposal, based on his taxonomic experience and insight, to unite several of Pilger's (1937) sections into a much enlarged subg. *Psyllium* (Rahn 1978).

Despite the great value of morphology for the classification and identification of species, the ineffectiveness of morphological phylogeny (Rahn 1996) to infer relationships within *Plantago* becomes evident with the results of our molecular phylogeny. This is probably aggravated by the general morphological reduction of most reproductive structures in *Plantago*, and possibly also the parallel evolution of similar characteristics in similar habitats. One illustrative example is the trichomes on scapes, which have repeatedly been reported as one of the most important taxonomic characters for *Plantago* (Rahn, 1974, 1992, 1996; Andrzejewska-Golec, 1991; Andrzejewska-Golec and

Świętosławski, 1993; Hassemer et al., 2014, 2015; Hassemer, 2016, 2017). Our phylogenetic results indicate that, although this character is very useful to classify and identify species, even closely-related species can differ considerably.

Even so, the morphology-based phylogeny of Rahn (1996) is more similar to our findings than molecular phylogeny based on the nuclear marker SUC1 (Ishikawa et al. 2009), emphasising the need for inclusion of multiple markers and interpretation in the light of current taxonomic understanding based on morphology, biogeography, and other evidence. A molecular phylogeny is only reliable when the samples used are correctly identified, which requires morphological knowledge of the taxa studied, and also nomenclatural knowledge, otherwise it is impossible to link morphologies to names. Furthermore, errors committed during the laboratory work, especially contamination, can also compromise the reliability of the resulting phylogenies and lead to erroneous conclusions in the worst case.

5. Future perspectives

Future phylogenetic research on *Plantago* should include the species indicated here as *incertae sedis* (*P. robusta*), as well as *P. nubicola* and the New Guinean species, which unfortunately could not be included in this study. Taxonomic revisions are critically necessary for the African, southern Asian and Hawaiian species of subg. *Plantago*. Furthermore, intensified taxonomic work is necessary to discover and present to science the still undescribed narrowly endemic species in sects. *Mesembrynia* and *Virginica*, whose species numbers are certainly underestimated.

A species level phylogeny including multiple accessions of all species and using nuclear as well as chloroplast DNA markers would greatly contribute towards the necessary knowledge for the appropriate development and application of conservation efforts and strategies for the narrowly endemic,

endangered *Plantago* species (e.g. Hassemer and Baumann, 2014; Hassemer, 2016, 2017; Hassemer and Rønsted, 2016), including the still little-understood cryptic species (Rahn, 1974; Hassemer et al., 2015). Conservation biologists should rely on the most reliable information available on the species, i.e. the most updated taxonomic treatments, and consider the combined knowledge accumulated by taxonomists and the results of new tools and techniques.

6. Acknowledgement

We would like to express our gratitude to Ricardo Segovia for providing seeds of *Plantago* species from central Chile; to Polina Volkova, Nikolaj Tikhomirov and Yuri Kopylov-Guskov for collecting *P. macrocarpa* samples from Bering Island, northeastern Russia; Yutao Wang and Zhengmei Lin for providing samples from China; and to Stephanie Dunbar-Co for providing samples from Hawaii and Rapa Iti. Gustavo Hassemer also acknowledges *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES) for a doctoral scholarship (process BEX 13765/13-2); *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) for a post-doctoral fellowship (process 153430/2018-4); the Linnean Society of London and the Systematics Association for a Systematics Research Fund (SRF) award; the Society of Systematic Biologists for a Graduate Student Research Award (GSRA); and three SYNTHESYS grants for visits to K, MA and P herbaria (processes GB-TAF-5317, ES-TAF-6610 and FR-TAF-6529). Alexey Shipunov was supported by North Dakota INBRE. Heidi M. Meudt was supported by core funding for Crown Research Institutes from the New Zealand Ministry of Business, Innovation and Employment's Science and Innovation Group.

7. References

Agnarsson, I., Kuntner, M., 2007. Taxonomy in a changing world: seeking solutions for a science in crisis. Systematic Biology 56, 531–539.

Albach, D.C., Meudt, H.M., Oxelman, B., 2005. Piecing together the ‘‘new’’ Plantaginaceae. American Journal of Botany 92, 297–315.

Andrzejewska-Golec, E., 1991. Ontogeny of trichomes in taxa of genus *Plantago* L. subgenus *Plantago*. Acta Societatis Botanicorum Poloniae 60, 249–258.

Andrzejewska-Golec, E., Świętosławski, J., 1993. Hair anatomy in *Plantago* subg. *Psyllium* (Plantaginaceae). Plant Systematics and Evolution 184, 113–123.

Andrzejewska-Golec, E., Ofterdinger-Daegel, S., Calis, I., Świątek, L., 1993. Chemotaxonomic aspects of iridoids occurring in *Plantago* subg. *Psyllium* (Plantaginaceae). Plant Systematics and Evolution 185, 85–89.

Baldwin, B.G., Sanderson, M.J., 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proceedings of the National Academy of Sciences 95, 9402–9406.

Barnéoud, F.-M., 1844. Mémoire de Botanique—Recherches sur le développement, la structure générale et la classification des Plantaginées et des Plumbaginées; Mémoire de Géologie—De l’origine des lacs. Schneider et Langrand, Paris, 44 pp., 2 tabs.

Barnéoud, F.-M., 1845. Monographie Générale de la Famille des Plantaginées. Fortin, Masson et Cie, Paris, 52 pp.

Bassett, I.J., 1966. Taxonomy of North American *Plantago* L., section *Micropsyllium* Decne. Canadian Journal of Botany 44, 467–479.

Bassett, I.J., 1973. The Plantains of Canada. Information Canada, Ottawa, 47 pp.

Bello, M.A., Chase, M.W., Olmstead, R.J., Rønsted, N., Albach, D., 2002. The páramo endemic *Aragoa* is the sister genus of *Plantago* (Plantaginaceae; Lamiales): evidence from plastid *rbc*L and nuclear ribosomal ITS sequence data. Kew Bulletin 57, 585–597.

Briggs, B.G., Carolin, R.C., Pulley, J.M., 1973. New species and lectotypification in Australian *Plantago*. Contributions from the N.S.W. National Herbarium 4, 395–398.

Briggs, B.G., Carolin, R.C., Pulley, J.M., 1977. Plantaginaceae. Flora of New South Wales 181, 1–35.

Briggs, B.G., 1980. *Plantago multiscapa*, a new species from Eremaean Australia, and notes on *Plantago* in western Australia. Telopea 2, 77–81.

Britton, N.L., Brown, A., 1913. An Illustrated Flora of the Northern United States, Canada and the British Possessions, 2nd ed., vol. 3. C. Scribner's Sons, New York, 637 pp.

Bruun-Lund, S., Clement, W.L., Kjellberg, F., Rønsted, N., 2017. First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). Molecular Phylogenetics and Evolution 109, 93–104.

Buermans, H.P.J., den Dunnen, J.T., 2014. Next generation sequencing technology: advances and applications. Biochimica et Biophysica Acta 1842, 1932–1941.

Cho, Y., Mower, J.P., Qiu, Y.L., Palmer, J.D., 2004. Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. Proceedings of the National Academy of Sciences 101, 17741–17746.

Craven, L.A., 1976. A review of the genus *Plantago* L. in New Guinea. Contributions from Herbarium Australiense 1976(13), 1–7.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9, 772.

Darwin, C., Darwin, F., 1887. The Life and Letters of Charles Darwin, Including an Autobiographical Chapter, vol. 2. J. Murray, London, 393 pp.

Decaisne, J., 1852. Plantaginaceæ. In: de Candolle, A.L.P.P. (Ed.), Prodromus Systematis Naturalis Regni Vegetabilis, vol. 13(1), V. Masson, Paris, pp. 693–737.

Delseny, M., Han, B., Hsing, Y.I., 2010. High throughput DNA sequencing: the new sequencing revolution. Plant Science 179, 407–422.

Dietrich, H., 1980. Cytologische Untersuchungen innerhalb der Familie der Plantaginaceae III. Cytotaxonomische Ergebnisse. Wissenschaftliche Zeitschrift der Friedrich-Schiller-Universität Jena, mathematisch-naturwissenschaftliche Reihe 29, 559–587.

van Dijk, E.L., Auger, H., Jaszczyszyn, Y., Thermes, C., 2014. Ten years of next-generation sequencing technology. Trends in Genetics 30, 418–426.

Dunbar-Co, S., Wieczorek, A.M., Morden, C.W., 2008. Molecular phylogeny and adaptive radiation of the endemic Hawaiian *Plantago* species (Plantaginaceae). American Journal of Botany 95, 1177–1188.

Dunbar-Co, S., Sporck, M.J., Sack, L., 2009. Leaf trait diversification and design in seven rare taxa of the Hawaiian *Plantago* radiation. International Journal of Plant Sciences 170, 61–75.

Ebach, M.C., Valdecasas, A.G., Wheeler, Q.D., 2011. Impediments to taxonomy and users of taxonomy: accessibility and impact evaluation. Cladistics 27, 550–557.

Freeland, J.R., Ciotir, C., Wensink, L., Dorken, M., 2017. Widespread cytonuclear discordance in narrow-leaved cattail (*Typha angustifolia*) does not explain the dominance of its invasive hybrid (*Typha* × *glauca*). Hydrobiologia 792, 53–65.

Gardner, A.G., Sessa, E.B., Michener, P., Johnson, E., Shepherd, K.A., Howarth, D.G., Jabaily, R.S., 2016. Utilizing next-generation sequencing to resolve the backbone of the Core Goodeniaceae and inform future taxonomic and floral form studies. Molecular Phylogenetics and Evolution 94, 605–617.

Givnish, T.J., Millam, K.C., Mast, A.R., Paterson, T.B., Theim, T.J., Hipp, A.L., Henss, J.M., Smith, J.F., Wood, K.R., Sytsma, K.J., 2009. Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). Proceedings of the Royal Society of London B: Biological Sciences 276, 407–416.

Gonçalves, S., Romano, A., 2016. The medicinal potential of plants from the genus *Plantago* (Plantaginaceae). Industrial Crops and Products 83, 213–226.

Grigoriev, Y.S., 1958. Plantaginaceae. In: Shishkin, B.K. (Ed.), Флора СССР, vol. 23, Izdatel'stvo Akademii Nauk SSSR, Moscow and Saint Petersburg, pp. 133–164.

Harrison, N., Kidner, C.A., 2011. Next-generation sequencing and systematics: what can a billion base pairs of DNA sequence data do for you? *Taxon* 60: 1552–1566.

Hassemer, G., Baumann, M.C., 2014. *Plantago corvensis* (Plantaginaceae): a new narrowly endemic species from rocky cliffs in southern Brazil. Journal of the Torrey Botanical Society 141, 181–185.

Hassemer, G., Baumann, M.C., Trevisan, R., 2014. *Plantago rahniana* (Plantaginaceae): a narrow endemic, new species from southern Brazil. Systematic Botany 39, 637–643.

Hassemer, G., Trevisan, R., Meudt, H.M., Rønsted, N., 2015. Taxonomic novelties in *Plantago* section *Virginica* (Plantaginaceae) and an updated identification key. Phytotaxa 221, 226–246.

Hassemer, G., 2016. *Plantago hatschbachiana* (Plantaginaceae), a critically-endangered new species from sandstone grasslands in Brazil, and an updated identification key to *Plantago* in Brazil and Paraguay. Phytotaxa 278, 141–152.

Hassemer, G., Rønsted, N., 2016. Yet another new species from one of the best-studied neotropical areas: *Plantago humboldtiana* (Plantaginaceae), an extremely narrow endemic new species from a waterfall in southern Brazil. PeerJ 4, e2050.

Hassemer, G., De Giovanni, R., Trevisan, R., 2016. The use of potential distribution models in the study of the distribution and conservation status of plants: the case of *Plantago* L. (Plantaginaceae) in Brazil. Journal of the Torrey Botanical Society 143, 38–49.

Hassemer, G., 2017. Reestablishment of *Plantago bradei* (Plantaginaceae), an overlooked narrowly endemic species from Serra do Caparaó, eastern Brazil, and range extension of *P. guilleminiana*. Phytotaxa 296, 253–264.

Hassemer, G., 2018. Notes on the montane Indo-Iranian species in *Plantago* subgenus *Plantago* (Plantaginaceae). Phytotaxa 336, 59–68.

Hassemer, G., Moroni, P., O'Leary, N., 2018. A nomenclatural revision of *Littorella* (Plantaginaceae, Plantagineae). Taxon 67, 1024–1028.

Hassemer, G., 2019. Novelties and notes on *Plantago* sect. *Virginica* (Plantaginaceae), including the description of a new species and a revised identification key. Webbia 74.

Heather, J.M., Chain, B., 2015. The sequence of sequencers: the history of sequencing DNA. Genomics 107, 1–8.

Hoggard, R.K., Kores, P.J., Molvray, M., Hoggard, G.D., Broughton, D.A., 2003. Molecular systematics and biogeography of the amphibious genus *Littorella* (Plantaginaceae). American Journal of Botany 90, 429–435.

Hou, C., Wikström, N., Strijk, J.S., Rydin, C., 2016. Resolving phylogenetic relationships and species delimitations in closely related gymnosperms using high-throughput NGS, Sanger sequencing and morphology. Plant Systematics and Evolution 302, 1345–1365.

Ishikawa, N., Yokoyama, J., Tsukaya, H., 2009. Molecular evidence of reticulate evolution in the subgenus *Plantago* (Plantaginaceae). American Journal of Botany 96, 1627–1635.

Iwanycki Ahlstrand, N.E., Verstraete, B., Hassemer, G., Dunbar-Co, S., Hoggard, R., Meudt, H.M., Rønsted, N., (in press). Ancestral range reconstruction of remote oceanic island species of *Plantago* (Plantaginaceae) reveals differing scales and modes of dispersal. Journal of Biogeography.

Jensen, S.R., Olsen, C.E., Rahn, K., Rasmussen, J.H., 1996. Iridoid glucosides in *Plantago alpina* and *P. altissima*. Phytochemistry 42, 1633–1636.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30, 772–780.

Klimko, M., Idzikowska, K., Truchan, M., Kreft, A., 2004. Seed sculpture of Polish species of the genus *Plantago* L. Acta Societatis Botanicorum Poloniae 73, 103–111.

Knope, M.L., Morden, C.W., Funk, V.A., Fukami, T., 2012. Area and the rapid radiation of Hawaiian *Bidens* (Asteraceae). Journal of Biogeography 39, 1206–1216.

Koboldt, D.C., Steinberg, K.M., Larson, D.E., Wilson, R.K., Mardis, E.R., 2013. The next-generation sequencing revolution and its impact on genomics. Cell 155, 27–38.

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

Li, Z., Wei, L., Hoggard, R.K., 2011. Plantaginaceae. In: Wu, Z., Raven, P.H., Hong, D. (Eds.), Flora of China, vol. 19, Science Press, Beijing, and Missouri Botanical Garden, St. Louis, pp. 495–503.

Liu, J.Z., Zhang, Q.M., Guo, S.H., Zhou, X.D., 1992. Seed morphology of *Plantago* in China and its taxonomic significance. Acta Phytotaxonomica Sinica 30, 118–125.

Meudt, H.M., 2011. Amplified fragment length polymorphism data reveal a history of auto-and allopolyploidy in New Zealand endemic species of *Plantago* (Plantaginaceae): new perspectives on a taxonomically challenging group. International Journal of Plant Sciences 172, 220–237.

Meudt, H.M., 2012. A taxonomic revision of native New Zealand *Plantago* (Plantaginaceae). New Zealand Journal of Botany 50, 101–178.

Meyers, S.C., Liston, A., 2008. The biogeography of *Plantago ovata* Forssk. (Plantaginaceae). International Journal of Plant Sciences 169, 954–962.

Moore, D.M., Chater, A.O., Cartier, D., 1976. Plantaginaceae. In: Tutin, T.G., Heywood, V.H., Burges, N.A., Moore, D.M., Valentine, D.H., Walters, S.M., Webb, D.A. (Eds.), Flora Europaea, vol. 4, Cambridge University Press, Cambridge, pp. 38–44.

Morey, M., Fernández-Marmiesse, A., Castiñeiras, D., Fraga, J.M., Couce, M.L., Cocho, J.A., 2013. A glimpse into past, present, and future DNA sequencing. Molecular Genetics and Metabolism 110, 3–24.

Murillo G., V.E., 2012. Revisión taxonómica del género *Plantago* L. (Plantaginaceae) en Chile. Scientia (Panamá) 22(2), 7–76.

Murray, B.G., Meudt, H.M., Tay, M.L., Garnock-Jones, P.J., 2010. New chromosome counts in New Zealand species of *Plantago* (Plantaginaceae). New Zealand Journal of Botany 48, 197–204.

Olmstead, R.G., dePamphilis, C.W., Wolfe, A.D., Young, N.D., Elisons, W.J., Reeves, P.A., 2001. Disintegration of the Scrophulariaceae. American Journal of Botany 88, 348–361.

Pilger, R.K.F., 1937. Plantaginaceae. In: Engler, H.G.A., Diels, F.L.E. (Eds.) Das Pflanzenreich, vol. 102. W. Engelmann, Leipzig, 466 pp.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Systematic Biology 53, 793–808.

Rahn, K., 1974. *Plantago* section *Virginica*: a taxonomic revision of a group of American plantains using experimental, taximetric and classical methods. Dansk Botanisk Arkiv 30(2), 1–180.

Rahn, K., 1978. Nomenclatural changes within the genus *Plantago* L., infraspecific taxa and subdivisions of the genus. Botanisk Tidsskrift 73, 106–111.

Rahn, K., 1984. *Plantago* sect. *Oliganthos* in southern South America, a taxonomic revision. Nordic Journal of Botany 4, 601–627.

Rahn, K., 1985. *Plantago* sect. *Carpophorae*, a taxonomic study. Nordic Journal of Botany 5, 143–151.

Rahn, K., 1992. Trichomes within the Plantaginaceae. Nordic Journal of Botany 12, 3–12.

Rahn, K., 1996. A phylogenetic study of the Plantaginaceae. Botanical Journal of the Linnean Society 120, 145–198.

Renoult, J.P., Kjellberg, F., Grout, C., Santoni, S., Khadari, B., 2009. Cyto-nuclear discordance in the phylogeny of *Ficus* section *Galoglychia* and host shifts in plant-pollinator associations. BMC Evolutionary Biology 9, 248.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61, 539–542.

Rosatti, T.J., 1984. The Plantaginaceae in the southeastern United States. Journal of the Arnold Arboretum 65, 533–562.

Rosenberg, M.S., Kumar, S., 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. Proceedings of the National Academy of Sciences 98, 10751–10756.

van Royen, P., 1983. The Alpine Flora of New Guinea, vol. 4. J. Cramer, Vaduz, pp. 2405–3516.

Rønsted, N., Göbel, E., Franzyk, H., Jensen, S.R., Olsen, C.E., 2000. Chemotaxonomy of *Plantago*. Iridoid glucosides and caffeoyl phenylethanoid glycosides. Phytochemistry 55, 337–348.

Rønsted, N., Chase, M.W., Albach, D.C., Bello, M.A., 2002. Phylogenetic relationships within *Plantago* (Plantaginaceae): evidence from nuclear ribosomal ITS and plastid *trn*L-F sequence data. Botanical Journal of the Linnean Society 139, 323–338.

Rønsted, N., Franzyk, H., Mølgaard, P., Jaroszewski, J.W., Jensen, S.R., 2003. Chemotaxonomy and evolution of *Plantago* L. Plant Systematics and Evolution 242, 63–82.

Samuelsen, A.B., 2000. The traditional uses, chemical constituents and biological activities of *Plantago major* L. A review. Journal of Ethnopharmacology 71, 1–21.

Savolainen, V., Chase, M.W., 2003. A decade of progress in plant molecular phylogenetics. Trends in Genetics 19, 717–724.

Schubert, M., Lindgreen, S., Orlando, L., 2016. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Research Notes 9, e88.

Shehata, A.A., Loutfy, M.H.A., 2006. On the taxonomy of Plantaginaceae Juss. *sensu lato*: evidence from SEM of the seed coat. Turkish Journal of Botany 30, 71–84.

Shipunov, A.B., 1998. The significance of the sculpture of the surface of seeds for systematics of genera *Plantago* L. and *Psyllium* Mill. (Plantaginaceae). Bulletin of Moscow Society of Naturalists, Biological Series 103, 41–51 [in Russian].

Sluys, R., 2013. The unappreciated, fundamentally analytical nature of taxonomy and the implications for the inventory of biodiversity. Biodiversity and Conservation 22, 1095–1105.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Soltis, D.E., Gitzendanner, M.A., Stull, G., Chester, M., Chanderbali, A., Chamala, S., Jordon-Thaden, I., Soltis, P.S., Schnable, P.S., Barbazuk, W.B., 2013. The potential of genomics in plant systematics. Taxon 62, 886–898.

Straub, S.C., 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, 349–364.

Sukumaran, J., Holder, M.T., 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics 26, 1569–1571.

Taskova, R., Evstatieva, L., Handjieva, N., Popov, S., 2002. Iridoid patterns of genus *Plantago* L. and their systematic significance. Zeitschrift für Naturforschung 57c, 42–50.

Tay, M.L., Meudt, H.M., Garnock-Jones, P.J., Ritchie, P.A., 2010a. DNA sequences from three genomes reveal multiple long-distance dispersals and non-monophyly of sections in Australasian *Plantago* (Plantaginaceae). Australian Systematic Botany 23, 47–68.

Tay, M.L., Meudt, H.M., Garnock-Jones, P.J., Ritchie, P.A., 2010b. Testing species limits of New Zealand *Plantago* (Plantaginaceae) using internal transcribed spacer (ITS) DNA sequences. New Zealand Journal of Botany 48, 205–224.

Tessene, M.F., 1969. Systematic and ecological studies on *Plantago cordata*. Michigan Botanist 8, 72–104.

Thiers, B., 2018. Index Herbariorum: a global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. Available from: http://sweetgum.nybg.org/science/ih (accessed 9 November 2018).

Turland, N.J., Wiersema, J.H., Barrie, F.R., Greuter, W., Hawksworth, D.L., Herendeen, P.S., Knapp, S., Kusber, W.-H., Li, D.-Z., Marhold, K., May, T.W., McNeill, J., Monro, A.M., Prado, J., Price, M.J., Smith, G.F., 2018. International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017. Regnum Vegetabile 159. Koeltz Botanical Books, Glashütten, 254 pp.

Tutel, B., 1982. Plantaginaceae. In: Davis, P.H., Edmondson, J.R., Mill, R.R., Tan, K. (Eds.), Flora of Turkey and the East Aegean Islands, vol. 7, University Press, Edinburgh, pp. 504–521.

Uribe-Convers, S., Carlsen, M.M., Lagomarsino, L.P., Muchhala, N., 2017. Phylogenetic relationships of *Burmeistera* (Campanulaceae: Lobelioideae): combining whole plastome with targeted loci data in a recent radiation. Molecular Phylogenetics and Evolution 107, 551–563.

Verdcourt, B., 1971. Plantaginaceae. In: Milne-Redhead, E., Polhill, R.M. (Eds.), Flora of Tropical East Africa, Crown Agents for Oversea Governments and Administrations, London, 8 pp.

Villarroel S., D., Wood, J.R.I., 2011. *Plantago pyrophila* (Plantaginaceae), a new species from the cerrados of eastern Bolivia. Kew Bulletin 66, 471–474.

Wagner, W.L., Herbst, D.R., Sohmer, S.H., 1990. Manual of the Flowering Plants of Hawai'i, vol. 2. Bishop Museum Special Publication 83. University of Hawaii Press and Bishop Museum Press, Honolulu, pp. 983–1854.

Wägele, H., Klussmann-Kolb, A., Kuhlmann, M., Haszprunar, G., Lindberg, D., Koch, A., Wägele, J.W., 2011. The taxonomist – an endangered race. A practical proposal for its survival. Frontiers in Zoology 8, 25–31.

Weryszko-Chmielewska, E., Matysik-Woźniak, A., Sulborska, A., Rejdak, R., 2012. Commercially important properties of plants of the genus *Plantago*. Acta Agrobotanica 65, 11–20.

Wong, C., Murray, B.G., 2012. Variable changes in genome size associated with different polyploid events in *Plantago* (Plantaginaceae). Journal of Heredity 103, 711–719.

Wong, C., Murray, B.G., 2014. In situ hybridization with genomic and rDNA probes reveals complex origins for polyploid New Zealand species of *Plantago* (Plantaginaceae). New Zealand Journal of Botany 52, 315–327.

Yamazaki, T., 1993. Plantaginaceae. In: Iwatsuki, K., Yamazaki, T., Boufford, D.E., Ohba, H. (Eds.), Flora of Japan, vol. 3a, Kodansha, Tokyo, pp. 384–386.

Zhu, A., Guo, W., Gupta, S., Fan, W., Mower, J.P., 2015. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. New Phytologist 209, 1747–1756.

Fig. 1. Photographs of species in *Plantago* subg. *Plantago*. **A**. *P. eriopoda*. **B**. *P. elongata*. **C**. *P. alpestris*. **D**. *P. euryphylla*. **E**. *P. glacialis*. **F**. *P.major*. **G**. *P. asiatica*.

Fig. 2. Photographs of species in *Plantago* subg. *Plantago*. **A**. *P. macrocarpa*. **B**. *P. tehuelcha*. **C**. *P. bradei*. **D**. *P. napiformis*. **E**. *P. rahniana*. **F**. *P. commersoniana*.

Fig. 3. Photographs of species in *Plantago* subg. *Plantago*. **A**. *P. cordata*. **B**. *P. triandra*. **C**. *P. spathulata*. **D**. *P. myosuros*. **E**. *P. udicola*. Photo credits: Mei Lin Tay (B, C, E) and Luís Adriano Funez (D).

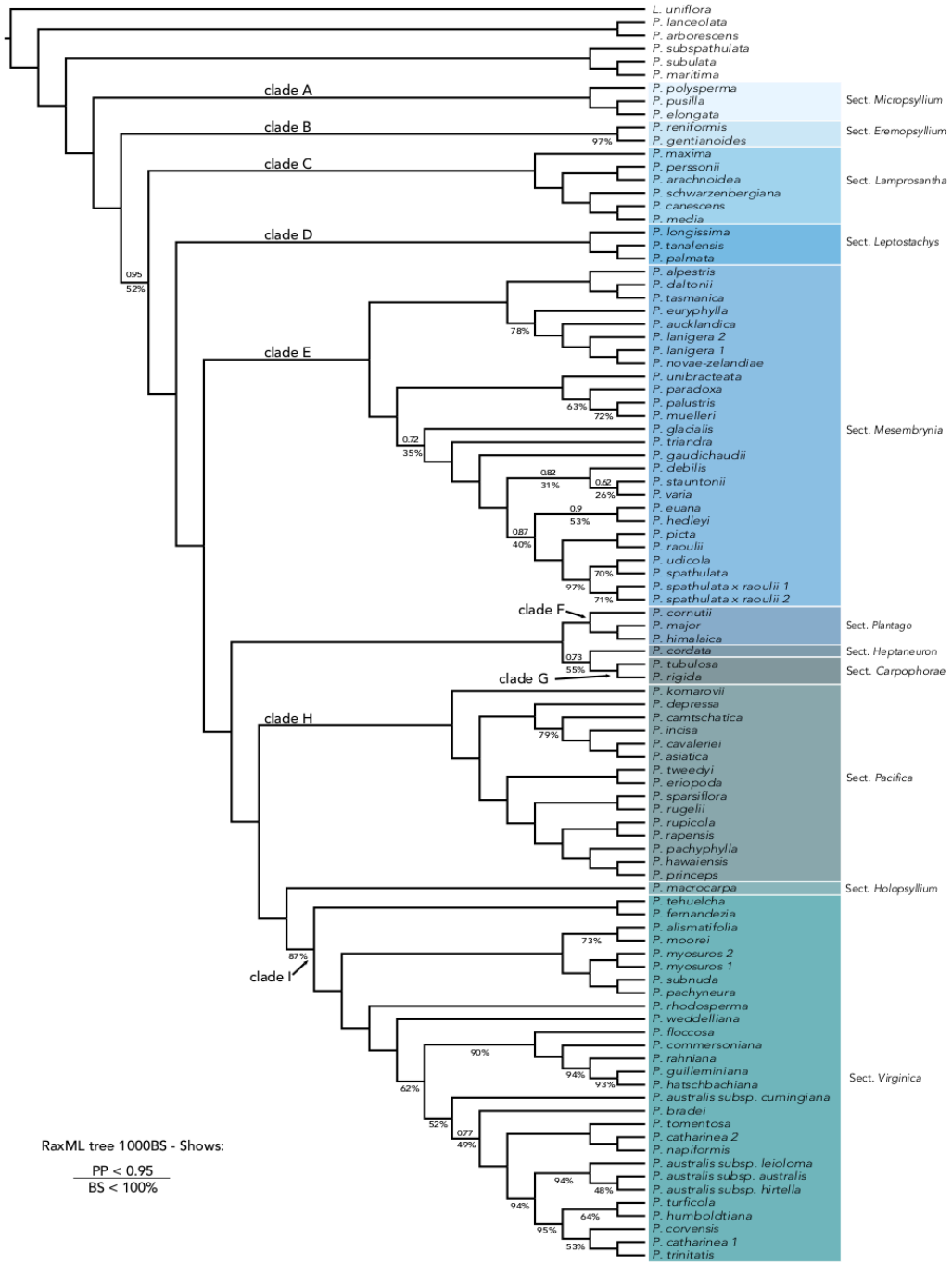

L. uniflora
P. lanceolata
P. arborescens
P. subspathulata
P. subulata
P. maritima
clade A
P. polysperma
P. pusilla
P. elongata
Sect. Micropsyllium
clade B
97%
P. reniformis
P. gentianoides
Sect. Eremopsyllium
clade C
P. maxima
P. perssonii
P. arachnoidea
P. schwarzenbergiana
P. canescens
P. media
Sect. Lamprosantha
clade D
P. longissima
P. tanalensis
P. palmata
Sect. Leptostachys
0.95
52%
P. alpestris
P. daltonii
P. tasmanica
78%
P. euryphylla
P. aucklandica
P. lanigera 2
P. lanigera 1
P. novae-zelandiae
clade E
P. unibracteata
P. paradoxa
63%
P. palustris
72%
P. muelleri
0.72
35%
P. glacialis
P. triandra
Sect. Mesembrynia
P. gaudichaudii
P. debilis
0.82
31%
0.62
26%
P. stauntonii
P. varia
0.9
53%
P. euana
P. hedleyi
0.87
40%
P. picta
P. raoulii
P. udicola
70%
97%
P. spathulata
71%
P. spathulata x raoulii 1
P. spathulata x raoulii 2
clade F
P. cornutii
P. major
Sect. Plantago
P. himalaica
P. cordata
Sect. Heptaneuron
0.73
55%
P. tubulosa
Sect. Carpophorae
clade G
P. rigida
P. komarovii
P. depressa
clade H
P. camtschatica
79%
P. incisa
P. cavaleriei
P. asiatica
P. tweedyi
P. eriopoda
Sect. Pacifica
P. sparsiflora
P. rugelii
P. rupicola
P. rapensis
P. pachyphylla
P. hawaiensis
P. princeps
P. macrocarpa
Sect. Holopsyllium
P. tehuelcha
P. fernandezia
P. alismatifolia
73%
P. moorei
87%
P. myosuros 2
clade I
P. myosuros 1
P. subnuda
P. pachyneura
P. rhodosperma
P. weddelliana
P. floccosa
90%
P. commersoniana
P. rahniana
94%
P. guilleminiana
93%
P. hatschbachiana
Sect. Virginica
62%
P. australis subsp. cumingiana
52%
P. bradei
0.77
49%
P. tomentosa
P. catharinea 2
P. napiformis
P. australis subsp. leioloma
94%
P. australis subsp. australis
48%
P. australis subsp. hirtella
94%
P. turficola
64%
P. humboldtiana
95%
P. corvensis
53%
P. catharinea 1
P. trinitatis
RaxML tree 1000BS - Shows:
PP < 0.95
BS < 100%

Fig. 4 (previous page). Phylogenetic hypothesis of *Plantago* subg. *Plantago*. Best tree obtained from the RaxML analyses based on plastome data. Clades with low support (PP<0.95 or BS<100%) are indicated. The updated classification of *Plantago* subg. *Plantago* is shown on the right.

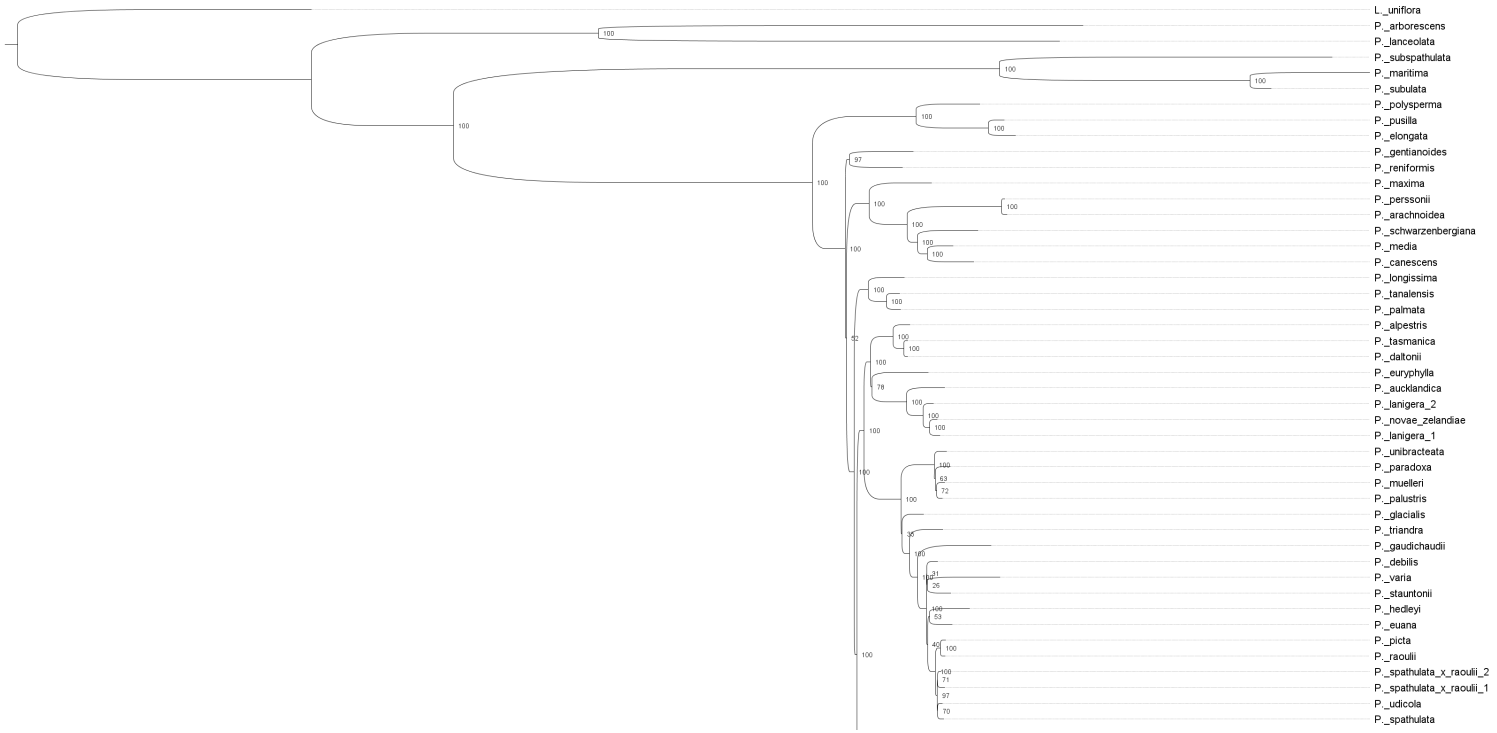
SUPPLEMENTARY MATERIAL FOR

The application of high-throughput sequencing for taxonomy: the case of *Plantago* subgenus *Plantago* (Plantaginaceae)

Gustavo Hassemer, Sam Bruun-Lund, Alexey B. Shipunov, Barbara G. Briggs, Heidi M. Meudt & Nina Rønsted

---

Figure S1. Specimen of *Plantago rhodosperma* in Texas (USA), from which DNA was extracted. Photo credits: A. B. Shipunov.


L._uniflora
P._arborescens
P._lanceolata
P._subspathulata
P._maritima
P._subulata
P._polysperma
P._pusilla
P._elongata
P._gentianoides
P._reniformis
P._maxima
P._perssonii
P._arachnoidea
P._schwarzenbergiana
P._media
P._canescens
P._longissima
P._tanalensis
P._palmata
P._alpestris
P._tasmanica
P._daltonii
P._euryphylla
P._aucklandica
P._lanigera_2
P._novae_zelandiae
P._lanigera_1
P._unibracteata
P._paradoxa
P._muelleri
P._palustris
P._glacialis
P._triandra
P._gaudichaudii
P._debilis
P._varia
P._stauntonii
P._hedleyi
P._euana
P._picta
P._raoulii
P._spathulata_x_raoulii_2
P._spathulata_x_raoulii_1
P._udicola
P._spathulata

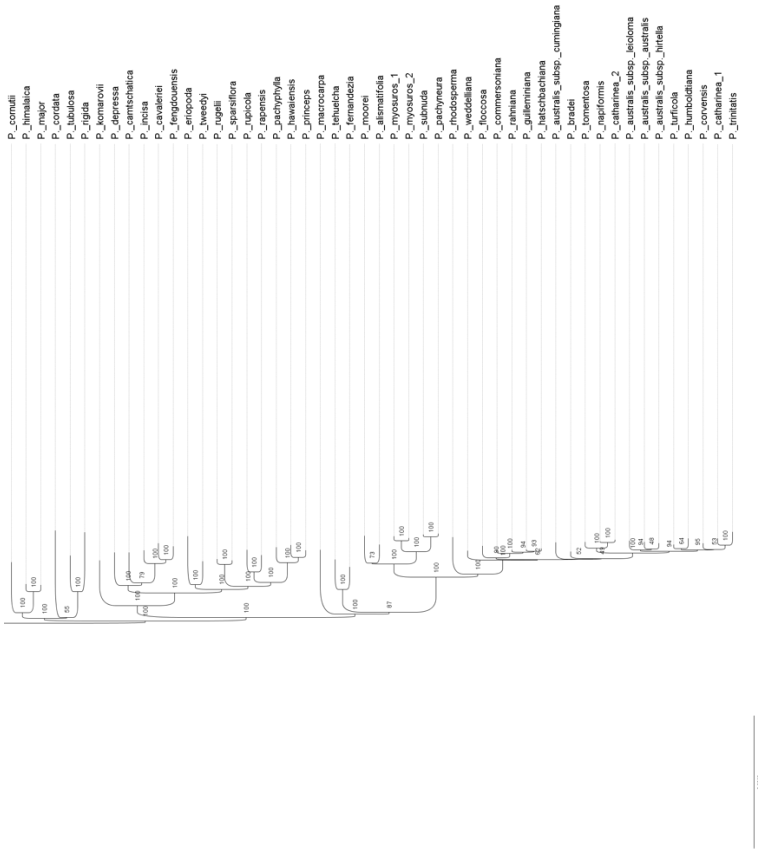
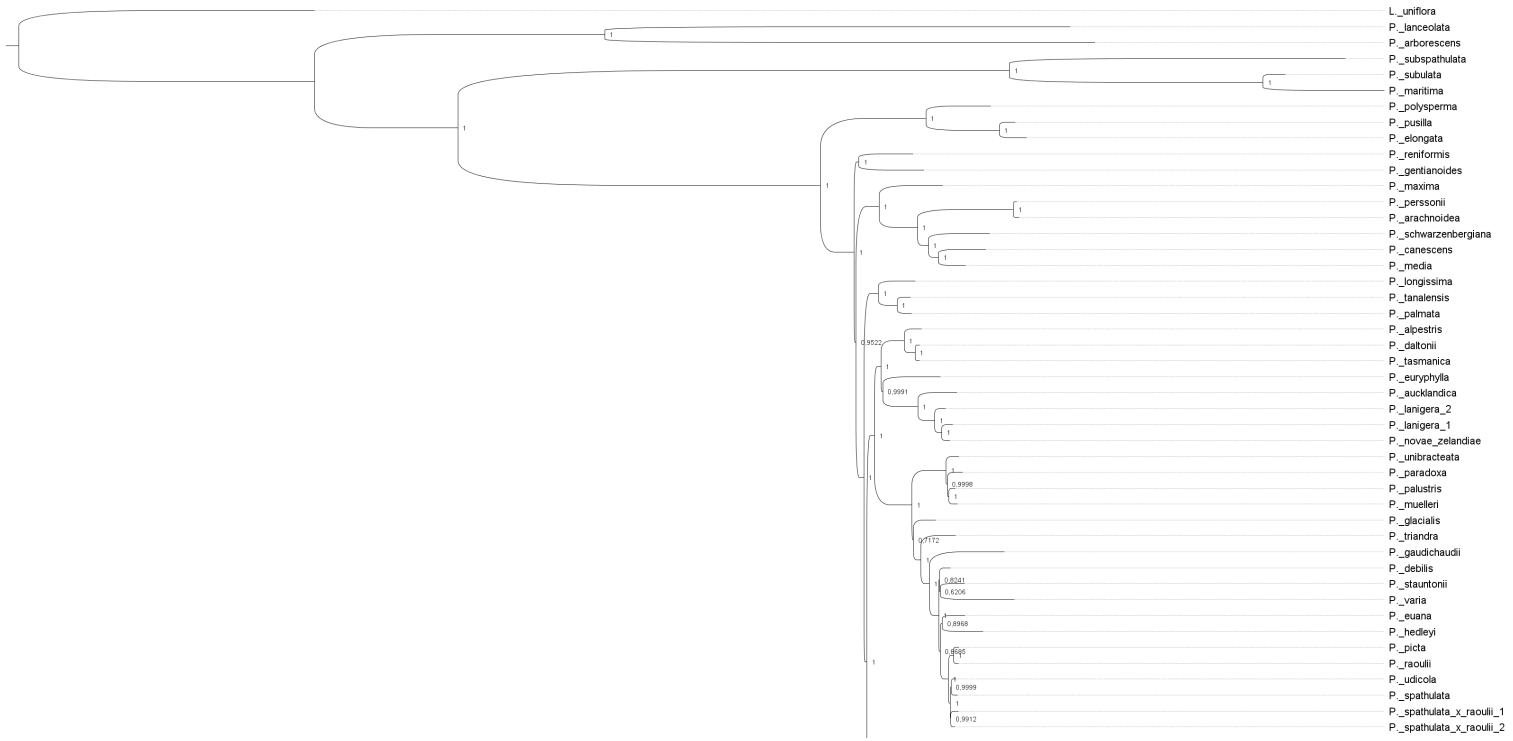


Figure S2. Best maximum likelihood tree based on plastome data.


L._uniflora
P._lanceolata
P._arborescens
P._subspathulata
P._subulata
P._maritima
P._polysperma
P._pusilla
P._elongata
P._reniformis
P._gentianoides
P._maxima
P._perssonii
P._arachnoidea
P._schwarzenbergiana
P._canescens
P._media
P._longissima
P._tanalensis
P._palmata
P._alpestris
P._daltonii
P._tasmanica
P._euryphylla
P._aucklandica
P._lanigera_2
P._lanigera_1
P._novae_zelandiae
P._unibracteata
P._paradoxa
P._palustris
P._muelleri
P._glacialis
P._triandra
P._gaudichaudii
P._debilis
P._stauntonii
P._varia
P._euana
P._hedleyi
P._picta
P._raoulii
P._udicola
P._spathulata
P._spathulata_x_raoulii_1
P._spathulata_x_raoulii_2
0.9522
0.9991
0.9998
0.7172
0.8241
0.6206
0.8968
0.9999
0.9912

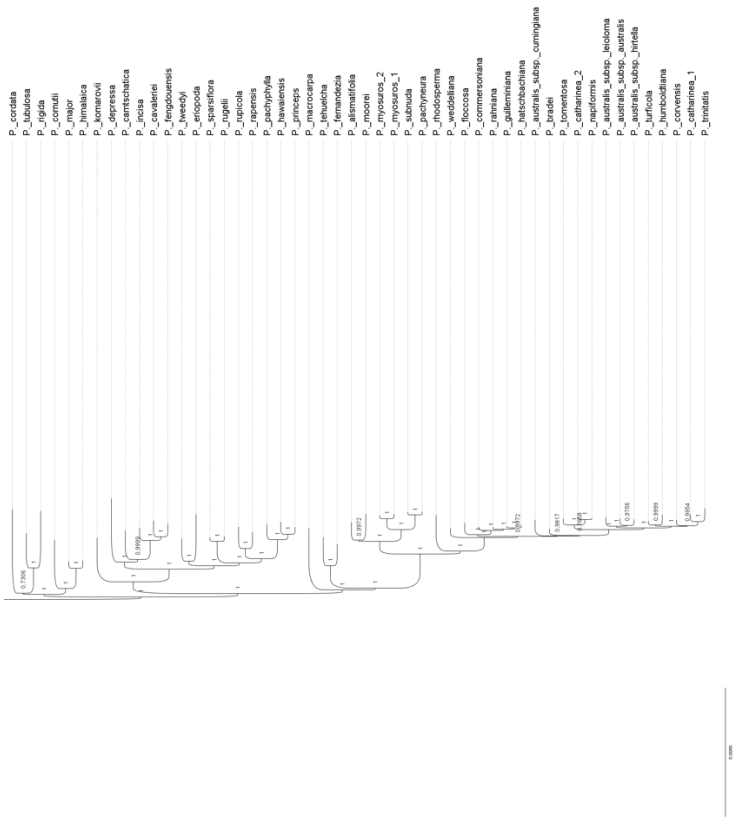


Figure S3. Best Bayesian inference tree based on plastome data.

Figure S4. Type of *Plantago cumingiana* (*H. Cuming s.n.*, LE-00016458). Copyright: V. L. Komarov Botanical Institute.

Appendix III

Hemmeligheden bag fignernes kærlighedsliv, evolution og religiøse betydning

Hemmeligheden om fignernes kærlighedsliv, evolution og religiøse betydning

Af Sam Bruun-Lund og Nina Rønsted

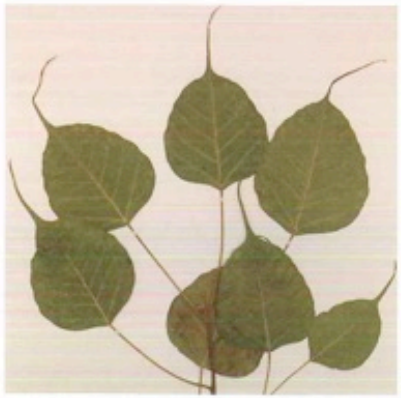
Figner *(Ficus)* er en af de største planteslægter med 750-800 kendte arter. *Ficus* tilhører morbærfamilien (Moraceae), der også indeholder en række andre kendte spiselige planter som morbær (*Morus*-arter), brødfrugt (*Atrocarpus altilis*) og jackfruit (*Atrocarpus heterophyllus*). I Danmark er den bedst kendte art fra *Ficus*-slægten *Ficus carica* (bornholmerfigen eller middelhavsfigen), og det er også fignerne fra denne art, man kan finde som tørrede i handlen – specielt omkring juletid. Foruden den gastronomiske interesse for figner er figentræer også planter med en lang kulturhistorie, som rækker flere tusind år tilbage og er nævnt både i Biblen og i andre religioners hellige skrifter.

Ficus hispidoides. *Fignen er i virkeligheden ikke en frugt, men en blomsterstand, der på indersiden er fyldt med små blomster. Foto: George Weiblen.*

Figner i kristendommen og buddhismen

Hvis man følger Biblens skabelsesberetning ordret, kom den første beklædningsgenstand fra *Ficus*-slægten. Det siges, at efter syndefaldet dækkede Eva og Adam sig med figenblade: "Derfor syede de figenblade sammen og bandt dem om livet" (1. Mos. 3, 7).

Man kan også finde beretninger om figner i buddhismen. For buddhister er *Ficus religiosa* meget betydningsfuld og vigtig. Det var netop under et træ af denne art, Buddha sad og gennem meditation kom frem til den fuldkomne viden – dvs. blev Buddha. *Ficus religiosa* er derfor et symbol på læren og vejen til Nirvana, og man kan finde *Ficus religiosa* i utallige helligdomsområder i Asien. Det helligste træ findes i Anuradhapura på Sri Lanka og siges at være en stikling fra netop det træ, Buddha sad under. Selve stiklingen kom til Sri Lanka 288 f.Kr. Den lille stikling blev sendt fra Nordindien som gave fra Kejser Asoka. I 1940'erne var træet ved at gå til af sygdom, men trods sine mere end 1700 år lever det stadigvæk i dag.



Ficus religiosa, *deponeret i herbariet ved Statens Naturhistoriske Museum (SNM). Det var under et træ af denne art, Buddha opnåede den fuldkomne viden. Senere er stiklinger fra træet blevet spredt ud over hele kloden og findes i mange helligdomsområder. Et træ af samme art kan ses i SNM's palmehus – dog uden direkte forbindelse til Buddhas træ. Foto: Sam Bruun-Lund.*

Ekstrem diversitet inden for slægten

Slægten Ficus repræsenterer mere end 80% af diversiteten i hele morbærfamilien. Gruppen af planter, der er tættest beslægtet med fignerne hedder Castilleae og indeholder kun ~60 arter, hvilket er bemærkelsesværdigt få sammenlignet med fignernes 750-800 arter, og arterne fra Castilleae ligner hinanden langt mere. *Ficus*-slægten udviser en bred vifte af livsformer – f.eks. stedsegrønne, løvfældende, buske, træer, epifytter, klatrende, urter og sågar kvælerfigner, der kvæler kæmpetræer i regnskoven i kampen om en plads i sollyset. Hvad skyldes denne ekstreme diversitet og forskellighed inden for én gruppe af planter? En af de faktorer, der med stor sandsynlighed har haft indflydelse på den utrolige diversitet, er fignernes unikke og raffinerede bestøvningsbiologi.

Ficus-slægten

Fælles for alle arter af *Ficus* er fignen. De fleste kender fignen som en frugt på linje med en dadel/abrikos. Men botanisk set er fignen ikke en frugt, men en indadvendt blomsterstand. Man kan forestille sig en sammenfoldet solsikke, hvor der på indersiden findes en masse små blomster tæt pakket sammen. I spidsen af fignen, modsat stilken, findes en pore, kaldet ostiolen, som er den eneste indgang til blomsterne. Ostiolen er beklædt med skælblade og giver kun særligt tilpassede bestøvere adgang til blomsterne på indersiden.

Slægten er udbredt i tropiske egne over hele verden. Enkelte arter vil også kunne findes i Middelhavsområdet, så som *Ficus carica* der også kan overleve på en lun vokseplads herhjemme. *Ficus carica* er dog ikke naturligt hjemmehørende i disse egne, men er indført fra det vestlige Asien og Mellemøsten.

Figner og deres bestøvere

Fælles for hele slægten er, at den er fuldkommen afhængig af en bestemt gruppe af bestøvere, nemlig hvepse af familien Agaonidae, der til gengæld benytter fignerne som æg-udklækningssted. Fignernes bestøvningsbiologi har udviklet sig til et utroligt samspil, som også involverer snyltehvepse, der ikke bidrager med bestøvning. Bestøvning foretages af en hvepsehun, der kravler ind i fignen gennem ostiolen. Hvepsen mister sine antenner og vinger i kampen om adgang til de ellers ufremkommelige blomster på indersiden. Hos nogle hvepsearter har hunnerne særlige pollensække med pollen fra en tidli-

Her ses bestøveren af Ficus septica, Ceratosolen juncundus, *som har fået adgang til blomsterne inden i fignen igennem ostiolen, der er dækket med skælblade. Foto: Finn Kjellberg (reproduceret med tilladelse fra Systematic Biology).*

gere figen, som de aktivt afsætter på blomsterne. Hos andre hvepsearter er pollen klistret til kroppen som hos humlebier, og bestøvningen foregår passivt, når hvepsen kravler på blomsterne. Hunhvepsen lægger æg i enkelte blomster, og larven ernærer sig fra blomsten. Senere parrer hanner og hunner sig. Hanhvepsenes eneste opgave er at gnave en udgang, som de unge hunner vil benytte til at forlade fignen. Hanhvepsen har ingen vinger og vil aldrig opleve ydersiden af fignen. Hunhvepsen finder en ny figen, og cyklussen gentager sig. Denne usædvanlige adfærd hos hvepsene og ofringen af mange blomster som mad til larverne er resultatet af mange millioner års evolutionært samspil mellem fignerne og deres bestøvere.

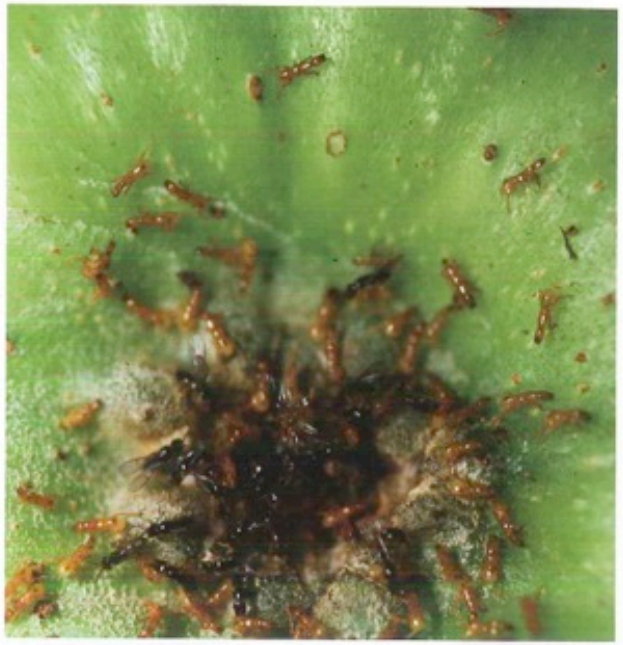
Samspillet er endnu mere imponerende, når man betænker, at figenhvepsene har en relativ kort levetid (få dage), som gør, at de ret hurtigt skal kunne lokalisere en egnet figenart til æglægning. Mange figenarter producerer derfor figner året rundt, og da figenproduktionen ikke er sæson- eller tidsbetinget udgør fignerne en vigtig fødekilde for mange fugle og andre dyr.

I mange år antog man, at hver art af figner kun blev bestøvet af én art af hvepse. Imidlertid har mindst 50 figenarter vist sig at have flere bestøvere, og biologer har derfor diskuteret, hvornår samspillet mellem fignerne og hvepsene mon opstod, og om det er dette samspil, der kan være årsagen til udviklingen af de mange arter af figner. I biologien bliver begrebet co-evolution brugt om den gensidige påvirkning, som forskellige arter kan udøve på hinandens udvikling/tilpasning over tid. På baggrund af mere end et årtis indsamling og studier af figner over det meste af verden, har det været muligt for os i samarbejde med flere internationale forskningsgrupper at teste for co-evolution ved bl.a. at sammenligne DNA-baserede stamtræer (og derved den hypotetiske udviklingshistorie) for både hvepsene og fignerne. Det viste sig, at associationen ikke var tilfældig, men at figen-hvepseforholdene var signifikant korrelerede. Det betyder, at hvepsene ikke tilfældigt har fundet egnede figner til æglægning, men at de mange figen- og hvepsearter er udviklet i gensidig afhængighed. Ved hjælp af forekomsten af fossile hvepse og figner har vi også kunnet datere fignerne og hvepsene til at være opstået for mere end 75 millioner år siden i den sene kridttid, hvilket betyder, at fignerne og hvepsene har levet sammen med dinosaurerne.

Fignernes oprindelse og evolution

De DNA-baserede metoder, man i mange år har brugt til at studere fignerne og hvepsene med, har ikke givet os information nok til entydigt at afgøre, hvor i verden samspillet mellem fignerne og hvepsene opstod. Nogle forskere mener, at det skete

De små Ceratosolen dentifer-*hvepse flokkes ved ostiole hos* Ficus hispidoides.
Foto: George Weiblen.



i det, der nu er Sydamerika, mens andre mener, at det må være sket i Asien. Vi har heller ikke med sikkerhed kunnet finde ud af, hvordan de mange hundrede arter er indbyrdes beslægtede, så der er fortsat en række uafklarede spørgsmål.

Inden for de seneste år er der udviklet banebrydende teknikker, såkaldt Next Generation Sequencing, der gør det muligt at afkode over 100 gange så meget information fra DNA som tidligere. Vi er nu gået i gang med det kæmpestore arbejde med disse metoder til at studere de mange hundrede figenarters DNA. Vores foreløbige resultater, der kun er baseret på godt 30 udvalgte arter, peger på, at de nye metoder med succes kan give os nok information til at afklare hele evolutionshistorien for fignerne. Foreløbig understøttes ideen om, at fignerne og hvepsene opstod i Sydamerika. De kommende år vil fortælle os meget mere om, hvordan og hvorfor fignerne og hvepsene har haft så stor succes med at kolonisere troperne verden over – en viden, som måske også kan hjælpe os med at forstå endnu mere om de processer, der har betydning den biologiske mangfoldighed, som er så vigtig for os, og hvordan vi kan passe bedre på den.

Sam Bruun-Lund er Ph.D.-studerende, og Nina Rønsted er lektor ved Statens Naturhistoriske Museum.

Appendix IV

Pressemeddelelse:
20 sjældne figentræer til Botanisk Have

11. december 2017

20 sjældne figentræer til Botanisk Have

JULEGAVE PÅ FORSKUD Figner er meget mere end sød julekonfekt med chokolade og pynt. Figentræerne er en slægt af træer, der har været utrolig succesfulde. De findes i dag bl.a. som slyng-figner, der vokser på sten eller i vand, som store savannetræer kaldet banyans og som kæmpe kvæler-figner, og for forskerne er det lidt af en gåde, hvorfor slægten har haft så stor succes. Forleden fik Botanisk Have i København en ekstraordinær stor julegave. Endda på forskud. Intet mindre end 20 nye sjældne figentræer, der snart vil kunne opleves i Palmehuset, og som på sigt skal gøre forskerne klogere på den tropiske plantes kringlede stamtræ.

Foto: Else Jorunn

- Figner er ikke kun den ene art, vi ofte spiser omkring juletid. Der findes et sted mellem 750-800 arter, som vokser over det meste af verden, hvor der er varmt. I Botanisk Have har vi i forvejen en fin samling af store og gamle figentræer, men den er nu blevet udvidet med unikke arter fra blandt andet Asien, Australien, Afrika og Sydamerika. Dermed får vi nu en af verdens vigtigste levende samlinger af figentræer, siger Sam Bruun-Lund, der er ph.d.-studerende ved Statens Naturhistoriske Museum, som Botanisk Have hører under. Han tilføjer:

- For en botaniker som mig er det her meget bedre end juleaften.

Udplantes i Palmehuset

Overtagelsen af fignerne er resultatet af et samarbejde mellem forskere og gartnere fra Statens Naturhistoriske Museum og Arboretet og Botanisk Hage, Universitetet i Bergen, Norge.

- Figentræer er en helt speciel og unik gruppe af planter, som har været fantastisk succesfulde med at tilpasse sig mange forskellige levesteder. De nye planter skal både bruges til DNA-arbejdet i forbindelse med kortlægningen af fignernes slægtskab og udbredelseshistorie, til undervisning og uddannelse af studerende på vores kurser ved Københavns Universitet og ikke mindst til forskellige formidlingsarrangementer i Botanisk Have, siger Sam Bruun-Lund og tilføjer:

- I øjeblikket er vi i fuld gang med at udplante figentræerne i Palmehuset, så publikum straks kan få glæde af dem.

Et stamtræ for alle arter

Det er et stort projekt at flytte så mange tropiske træer ind i de levende samlinger, og både gartnere og forskere fra begge botaniske haver har hjulpet til for at få den ekstraordinære store julegave på plads.

Figentræerne skal fremover indgår i et stort forskningsprojekt ledet af professor Nina Rønsted og ph.d.-studerende Sam Bruun-Lund. Projektet er finansieret af Det Fri Forskningsråd.

Projektets mål er at producere et stort stamtræ for alle 800 arter af figner for derefter at kunne forklare, hvorfor lige fignerne er blevet så forskellige, og at der er så mange arter i forhold til andre plante-grupper, der vokser de samme steder.

Relaterede nyheder

Gymnasieelever lærer at finde ny malariamedicin i regnskoven
21. nov. 2017

Naturhistorisk gave til alle
05. dec. 2017

Kontakt

Sam Bruun-Lund
Botaniker, Ph.d.-studerende
Statens Naturhistoriske Museum
Københavns Universitet
Mail: sam.bruunlund@snm.ku.dk
Mobil: 22 44 59 03

Foto

Fotos på denne side må bruges i forbindelse med presseomtale af historien. Klik på billederne for at downloade dem i høj opløsning. Krediter venligst fotografen.

Til DNA-arbejdet anvendes både de levende samlinger i Botanisk Have og hundredevis af tørrede indsamlinger fra museets herbarium. Med støtte fra Carlsbergfondet lykkedes det også at få indsamlet figner i Australien i sommers.

Der vil løbende blive formidlet om projektets fremskidt i Palmehuset, hvor figentræerne nu bliver plantet, så alle kan følge med og lære mere om de fantastiske figner.

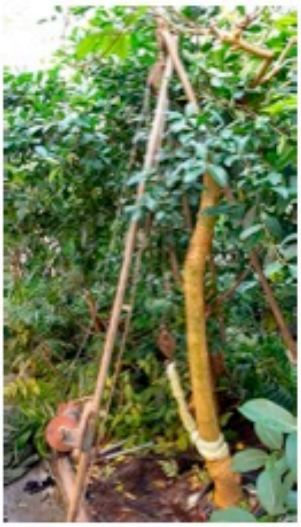


Foto: Else Jorunn

Statens Naturhistoriske Museum
Københavns Universitet
Gothersgade 130

Kontakt:
Statens Naturhistoriske Museums
snm@snm.ku.dk

FIG STREET
Yungaburra, QLD, Australia