

Andrew Groover
Quentin Cronk *Editors*

Comparative and Evolutionary Genomics of Angiosperm Trees

Plant Genetics and Genomics: Crops and Models

Volume 21

Series Editor

Richard A. Jorgensen

More information about this series at <http://www.springer.com/series/7397>

Andrew Groover • Quentin Cronk

Editors

Comparative and Evolutionary Genomics of Angiosperm Trees

 Springer

Editors

Andrew Groover
Pacific Southwest Research Station
United States Forest Service
Davis, CA
USA

Quentin Cronk
Department of Botany
University of British Columbia
Vancouver, BC
Canada

ISSN 2363-9601 ISSN 2363-961X (electronic)
Plant Genetics and Genomics: Crops and Models
ISBN 978-3-319-49327-5 ISBN 978-3-319-49329-9 (eBook)
DOI 10.1007/978-3-319-49329-9

Library of Congress Control Number: 2017955083

© Springer International Publishing AG 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Angiosperm trees display a dizzying array of diversity in morphology, anatomy, physiology and biochemistry. This diversity has been described and catalogued by various scientific disciplines, but the molecular, genetic, and evolutionary mechanisms underlying this diversity have only recently been explored.

This book, *Comparative and Evolutionary Genomics of Angiosperm Trees* marks a change in focus of tree genomics, from single species to comparative approaches. Excitingly, advances in genomic and sequencing technologies are ushering a new era of research broadly termed comparative genomics, which simultaneously exploits and describes the evolutionary origins and genetic regulation of traits of interest. Effective comparative genomic approaches for trees are enabled by an explosion in genomic data including an increasing number of complete genome sequences available for angiosperm trees, and extensive gene expression data available for a wider array of species. We believe that there is a great potential role for comparative approaches for the study of angiosperm trees, both with regards to understanding the fundamental evolution and development, as well as addressing problems of economic or ecological importance.

This book is intended as resource to provide background on the diverse biological subject areas pertaining to comparative and evolutionary genomic approaches of angiosperm trees. We elected not to make genomic technologies (e.g. the latest sequencing technologies) or computational approaches a main focus of the book, as they are already covered by other literature, and also are rapidly changing. Instead, the chapters focus on biological, genomic, and evolutionary aspects of angiosperm trees that provide information and perspectives that will support researchers broadening the focus of their research. We hope this will provide a valuable resource, and have longevity of relevance that will outlive the particulars of current-day technical approaches.

The first section of the book provides background on the evolution and diversification of angiosperm trees, as well as description of the salient features and diversity of the unique physiology and wood anatomy of angiosperm trees. The second section describes developments in the most model advanced angiosperm tree species (poplars) as well as species that are emerging models. The third section

describes the structural features and evolutionary histories of angiosperm tree genomes, followed by a fourth section focusing on the genomics of traits of biological, ecological and economic interest.

We would like to acknowledge the significant efforts of the authors of each chapter, and the overall high quality of the writing and information contained within their chapters.

Davis, CA, USA
Vancouver, BC, Canada

Andrew Groover
Quentin Cronk

Introduction: Comparative Genomics of Angiosperm Trees: A New Era of Tree Biology

Quentin C.B. Cronk and Andrew T. Groover

Abstract Forest tree genomics has made enormous strides in recent years, by describing the expression and function of genes influencing tree growth and development, and even sequencing the entire genomes of select “model” tree species. We believe that the next chapter of forest tree genomics will focus on cross-species comparative approaches, which will have the ability to provide fundamental new insights into the unique biology and evolutionary history of tree species. Angiosperm trees in particular are fascinating in light of evolution. Angiosperm trees represent the extensive genome evolution, including whole genome duplications, exhibited by different angiosperm lineages. Angiosperm trees also present amazing morphological, physiological and biochemical diversity, providing the opportunity to use comparative genomic approaches to understand the evolutionary origin and diversification of traits associated with trees. This book provides background on biological, genomic, and evolutionary aspects of angiosperm trees, in support of researchers exploring the use of comparative and evolutionary genomic approaches. This introduction briefly reviews the diversity of angiosperm trees and sets out the conceptual framework for comparative and evolutionary study of angiosperm tree biology using genomic tools, and highlights individual chapters within this book.

Keywords Evolution • Wood Developmental Biology • Population Genomics • Angiosperm Trees • Comparative Genomics

Q.C.B. Cronk (✉)
University of British Columbia, Vancouver, BC V6T 1Z4, Canada
e-mail: quentin.cronk@ubc.ca

A.T. Groover
Pacific Southwest Research Station, US Forest Service, Davis, CA, USA
Department of Plant Biology, University of California Davis, Davis, CA, USA

A World of Trees

The carbon dioxide monitoring station at Mauna Loa, Hawaii, has revealed a steady rise of atmospheric carbon dioxide, year-on-year. However, within this rise there is strong seasonal variation. Slightly more carbon dioxide is detectable in the northern hemisphere winter and spring and slightly less in summer and fall. Allowing for a time lag of atmospheric mixing, this is attributable to the vast areas of northern hemisphere boreal and temperate forests losing biomass and releasing carbon dioxide in the fall and taking up carbon dioxide (1.5×10^{10} tons net) with new growth in the spring (Bolin and Keeling 1963). Nothing could more clearly indicate the impact of forests on the biosphere, with trees exerting an influence on atmospheric chemistry detectable half a world away in the central Pacific.

Forests also have direct impacts on human lives. As producers of raw material for industrial processes they provide employment, directly or indirectly, for millions. Forests heat homes and cook food: 6.1 %, or 772 million tons of oil equivalent (MTOE), of primary global energy comes from wood (FAO 2014). This is mainly in the rural and developing world, where fuelwood is directly gathered from forests, but increasingly woody biomass is being grown as bioenergy crops.

As reservoirs of biodiversity, forests are important biomes for biodiversity conservation. More intangible but nevertheless important are the ecosystem services forests provide: climatic moderation, erosion and landslip prevention, and watershed protection, to name a few. Forests are also human habitat, providing food and shelter for many. Even in the urbanised West they are increasingly valued for recreation, and in terms of providing a “connection to nature” which is increasingly recognised as a necessary part of human development (FAO 2014). Additionally, forests and related disturbances (e.g. wildfire) directly affect an increasing number of people living at the wildland-urban interface.

Globally, natural (minimally exploited) and semi-natural (exploited but minimally managed) forests are under threat. This is mainly through conversion, by burning and logging, into agricultural landscapes (or “agrosapes”). However many

Table 1 Status of forestry plantations in 2000

Plantation type	Area (000 ha)
<i>Acacia</i> (Leguminosae)	8317
<i>Eucalyptus</i> (Myrtaceae)	17,860
<i>Hevea</i> (Euphorbiaceae)	9885
<i>Tectona</i> (Lamiaceae)	5716
Other angiosperm	33,556
Total angiosperm	75,334
Conifer(all)	58,134
Unspecified	53,618
TOTAL area (000 ha)	187,086
Rate of planting (000 ha/year)	4493

Source: The Global Forest Resources Assessment 2000 (FAO 2000)

of these agrosapes are subsequently used for the establishment of tree plantations in which intensively-managed, fast-growing, short-rotation trees are grown efficiently for industrial purposes (Table 1). Such plantations can, and should, take some of the pressure off natural forests and provide wide ranging socio-economic benefits, although their benefits are often disputed.

Angiosperm Trees as Genomic Subjects

This book deliberately focuses on angiosperm trees. Conifers, although important, are only dealt with here when appropriate to highlight the evolution of traits found in angiosperm trees. Angiosperm trees are much more diverse than conifers, as well as being the major component of most short-rotation plantation systems (FAO 2001; Table 1). While angiosperm trees present some challenges for biologists due to their large size and long generation times, modern genomic technology has done much to make them scientifically tractable and has allowed the exploitation of characteristics specific to trees that make them particularly attractive as biological subjects. First, wood itself is a globally important trait, not only economically valuable but also pivotal to the functioning of the biosphere. The enormous variety of angiosperm woody plants, and wood types, makes possible comparative and evolutionary approaches. Secondly, many trees have large populations and wide ranges, and show adaptation to diverse climates, and therefore make good subjects for the study of local adaptation. While putting trees in the laboratory is challenging, it has proved possible to take the laboratory to the tree and use wild populations as “natural breeding experiments”. Genomic characterization of natural variation grown in common gardens has allowed the molecular basis of traits to be uncovered, for instance using genome-wide association studies (GWAS).

Angiosperm Tree Diversity

There are about 369,000 species of angiosperms (RBG Kew 2016), of which between one fifth and one quarter can be considered trees (i.e. 70,000–90,000). The angiosperms, including those that are trees, display a huge amount of morphological variation, variation that makes comparative approaches particularly rewarding. Economically, there are ca. 1575 angiosperm trees with widely traded timber (Mark et al. 2014). The majority of these are tropical, reflecting the high diversity of tropical forest biomes. The most important family of tropical timber trees by far is the Leguminosae (Fabaceae), followed by the Meliaceae. A summary of some of the most important tropical tree families for forestry is given by Cronk and Forest (chapter 1: Table 4, this volume).

In temperate regions the situation is somewhat different. There is less overall tree diversity in temperate biomes and there Fagaceae is arguably the most important forestry family (chapter 1; Table 4). When temperate forestry trees are mapped on a

summary angiosperm phylogenetic tree it can be seen that the species are highly clustered in one clade (rosid 1). When the same is done for tropical trees the spread is more even, reflecting greater tropical diversity. Genomic resources for trees are rapidly developing. Every year new genome projects are announced. Happily, any discussion in this book will quickly be out of date, which indicates the health and dynamism of the subject. Nevertheless it may be useful to make a few observations on the situation as it now exists (2016). Emerging genomes of forest trees are discussed by Sollars and Buggs (chapter 4, this volume). The best resource by far is that provided by poplar (*Populus trichocarpa*). This was the third plant genome to be completed (after *Arabidopsis* and rice) and is of high quality (Douglas, chapter 3, this volume). Many of the emerging genomes are from the same major clade (rosid 1), such as the Fagaceae genomes.

However, given the importance of the legumes in tropical forestry, the absence of a high quality tree legume genome is notable. An *Acacia* genome is nearing completion, and *Acacia* is the most important plantation legume. However a tree legume in the same clade as crop legumes like soya (*Glycine max*) would allow tremendous synergy between legume tree genomics and legume crop genomics.

The ease with which whole shot-gun sequencing can now be carried out may herald an end of the “model tree” or “model plant” paradigm. Genomic resources can now be generated for any tree, however rare or obscure. This is to be welcomed. However a note of caution should be sounded. While sequencing is easy, assembly and annotation are not. A multiplicity of poorly assembled and annotated genomes may turn out to be more trouble than they are worth.

Model Species versus Comparative Trait-Based Approaches

Woodiness is a labile trait, and it has been modified to different extents, from shrubs to forest giants, in different angiosperm lineages. It has been lost, and sometimes regained, in many lineages. The diversity of woods, from balsa to ebony and teak, differing anatomically and chemically, provide a resource for the study of the molecular basis of this trait. Trait-based approaches, i.e. choosing a trait and following it wherever it iterates in organisms, stand in opposition to the model organism-based approach in which a single organism is chosen as one in which to study many traits (or at least as many as that organism possesses). The model organism has many advantages, not least the ability to build on a growing body of organism specific protocols and resources. However the weakness of the model organism approach is the limited view of any particular trait that it offers – without examining the trait and associated regulatory mechanisms in additional species it is not possible to know if findings are unique to the model species under study, or what the evolutionary history of the trait might be. Consequently, the immense power of comparative and evolutionary approaches cannot be brought to bear when working with a single model species.

Happily the tension between the model organism approach and a comparative approach is rapidly becoming obsolete. Any organism can now be a “model” at least

in the sense that genome resources can be generated readily. Sollars and Buggs (chapter 4, this volume) detail some recent genome projects of forest trees, and, judging by the speed at which new projects are coming on stream, this is likely to be the tip of a very large iceberg. Furthermore there is the possibility of using comparisons between classic plant models such as *Arabidopsis* and emerging tree models (such as poplar). The arabidopsis-poplar comparative model approach has already proved its power in many examples (e.g. Rottmann et al. 2000; Johnson and Douglas 2007).

The huge variety of angiosperm trees provides a scientific opportunity when unleashed by genomics. Although every taxon is a potential genomics model, working with organisms that reach vast sizes and do not domesticate well in the lab has been daunting. Rather than domesticate trees to the lab, the lab has been taken to the tree. Tree biologists have learnt to use the forest as the “growth chamber” and natural populations as “breeding experiments”. Common gardens, while expensive to set up, are long lasting (if land tenure issues can be solved) and have the potential to supply large quantities of data to multiple studies over many years (Fetter, Gugger & Keller, chapter 13, this volume).

Evolutionary and Comparative Genomics for Angiosperm Trees

Comparative approaches to tree biology can work at many different scales (Table 2), from the comparison of different genetic individuals in the same species, to the comparison of very different tree species in different plant families. The immense

Table 2 Comparative approaches to angiosperm tree genomics

Comparison	Questions	Examples
1. Between plant families, e.g. <i>Eucalyptus</i> (Myrtaceae) vs <i>Populus</i> (Salicaceae)	Conserved pathways in major traits, e.g. wood formation	Hefer et al. (2015), He and Groover, chapter 10, this volume
2. Between related genera, e.g. <i>Populus</i> vs <i>Salix</i>	Differences in genomic architecture, e.g. differences in sex locus architecture between willows and poplars	Olson, Hamrick and Moore, chapter 7, this volume
3. Interspecific, e.g. <i>Populus deltoides</i> vs <i>Populus trichocarpa</i>	Speciation genomics, adaptive introgression	Bawa and Holliday, chapter 8, this volume
4. Interpopulation, e.g. northern vs southern or montane vs lowland	Environmental genomics, origin of adaptations	Fetter, Gugger and Keller, chapter 13, this volume
5. Intrapopulation	Allelic variation, balancing selection, gene flow/selection balance	Fetter, Gugger and Keller, chapter 13, this volume

morphological variety of angiosperm trees lends itself in particular to comparative approaches. Top level comparisons give answers to the central question of what aspects of biology are fundamental and which are special responses of a limited clade (evolutionary lineage). This approach has been used in an attempt to define core wood genes by comparing eucalyptus and poplar (Hefer et al. 2015) for instance.

Comparative approaches can also be used in a phylogenetic context to understand how key regulators were co-opted in the origin of novel traits. The moss versions of the key wood developmental regulators, NAC transcription factors, have been shown to have a conserved function, able to act as wood regulators in angiosperms, despite mosses having no wood. This is strong evidence that it is not the primary function of this gene that has changed to facilitate the origin of wood, but the downstream modules that it regulates (Xu et al. 2014).

Environmental and Population Genomics: Exploiting Genetic Diversity

Trees are generally highly outbreeding and every wild tree is generally a genetically unique individual. Every comparison between two or more trees can therefore be very revealing of the effect of their different genetic constitutions and environmental histories on phenotypes. Trees have a number of features that makes them particularly amenable for comparative biology at the population level. Some of these features are highlighted briefly here, and detailed in chapter 7 by Olson, Hamrick and Moore, and chapter 8 by Bawa and Holliday.

First, they tend to exist as large populations with a large ranges covering more than one climatic zone. They are therefore highly suitable for landscape genomics and studies of climatic adaptation. Secondly, they are generally highly outbreeding, with effective gene flow, minimising population structure. This contrasts with inbreeding herbs like *Arabidopsis thaliana* which exist as strongly structured populations (a mosaic of inbred lines). As population structure is a confounding factor in all studies of genotype:phenotype association and local adaptation, anything that reduces it is important. Additionally, hybridization among species is common in some tree genera (e.g. *Populus*, *Salix*, *Eucalyptus*), presenting intriguing questions about speciation and unique ecological attributes of hybrids in forested landscapes.

Genome-wide association studies (GWAS) using natural variation in populations have begun to disentangle the molecular basis of important tree traits such as large growth, perenniality and architecture (McKown et al. 2014) as well as wood structure and the chemistry (Porth et al. 2013). Adaptation to environment is one of the most important biological phenomena, on which much of our agricultural and forestry productivity is based. With an environment changing rapidly due to CO₂ climate forcing, it becomes more important than ever to understand.

Big Questions for Angiosperm Tree Genomics

It is difficult to predict the future but a number of big questions seem to be emerging at the intersection of genomics and tree biology. These include epigenetics, structural genome variation, the genomic basis of the origin of woodiness and the genomic basis of tree traits, such as architecture and sexuality.

1. **Epigenetics.** Trees are long-lived organisms that cannot move to avoid environmental stress, so they have to endure it (Bräutigam et al. 2013). To what extent does environmentally-induced epigenetic regulation contribute to survival? As many trees can be cloned, it is possible to conduct what in human biology would be called “identical twin studies”, but on a vast scale.
2. **Structural genome variation.** Angiosperm tree species display the surprising history of genome duplication and rearrangements that have occurred at various points in angiosperm evolution and lineages. Structural variation in angiosperm genomes is presented in chapter 5 by Street. As discussed in chapter 6 by Hussey, Wegrzyn and Vasquez-Gross, a closely related topic is the complex evolution of gene families in angiosperm tree lineages, including how gene families undergo selection and fractionation. Additionally, angiosperm trees exist in large populations and are known to produce, at low level, structural genome variants such as triploids (Mock et al. 2012), aneuploids (with an extra copy of a particular chromosome) and segmental variants (with translocations, inversions or duplications of parts of chromosomes). Using new genomic approaches these can now be detected within populations. What is their effect on biology, speciation and genome evolution? Discussion of these and related topics is given by Bawa and Holiday (chapter 8).
3. **Woodiness** is a globally significant trait and the molecular control of its evolution is a pressing question. It is well known that various herbaceous lineages have lost woodiness only to gain it again in some clades. These clades of recently evolved woodiness provide promising experimental systems for the investigation of the evolution of woodiness itself (Moyers and Rieseberg 2013; Davin et al. 2016). Variation in wood structure and anatomy is summarized in chapter 2 by Spicer, and a discussion of the molecular regulation of wood development is presented by He and Groover in chapter 10.
4. **Genomic basis of important traits of trees,** including comparison of these traits to similar if not homologous traits in non-tree species. For example, tree architecture is tremendously complex. At its simplest, it is evident that some trees have narrow crowns whereas others are broad and spreading. Such traits are of great importance in commercial forestry and pomology (Segura et al. 2008) but are also important in the fundamental understanding of plant development. The regulation of tree architecture is presented in chapter 9 (this volume) by Hearn. Trees are also well known for the variety of sex expression from hermaphroditism to monoecy, dioecy or a mixture (polygamy). When studied in comparative framework, tree genomics may be expected to shed much light on the molecular pathways underlying such variation (Geraldes et al. 2015; Olson, Hamrick & Moore, chapter 7, this volume; Fetter, Guggen and

Keller, chapter 13, this volume). The regulation of perennial growth is fundamental to trees, and the regulation of phase change and phenology in trees is presented in chapter 11 (this volume) by Brunner, Varkonyi-Gasic, and Jones. Trees have also evolved additional strategies to deal with abiotic stress including drought, which is discussed in chapter 12 by Bastiaanse. Trees also have a variety of interactions with other organisms, including both pathogens and insect pests, as well as beneficial symbionts. Some of these interactions are described in chapter 14 by Plett and Plett.

These examples merely touch on some of the exciting basic science that will emerge from tree genomics over the next few years, and a bright future can be anticipated. However, as already alluded to, angiosperm dominated forests and plantations are of great ecological, economic and social importance. And many of these forests are in peril. Effective comparative genomic approaches can also provide new tools for applied forest management.

Conservation of Forest Biodiversity and Forest Genetic Resources

Forests, particularly tropical wet forests, are enormously rich in biodiversity. It is difficult to draw the line between trees and shrubs, but of the 369,000 species of flowering plants (RBG Kew 2016), between 70,000 and 90,000 can be regarded as trees. These are the largest organisms in terms of biomass on the planet, reaching, in the case of forest giants like *Eucalyptus regnans*, over 100 m in height. Many ordinary trees commonly reach 30 m or more. These are huge organisms by any standards. In turn trees support a pyramid of dependent organismal diversity. Erwin found 1143 beetle species on the tropical tree *Luhea seemannii*, of which an estimated 162 were only found on that tree species (Erwin 1982).

As well as the organismal diversity there is the genetic diversity of the trees themselves, which is essential to maintaining vigorous, well-adapted tree populations. Genomic tools are hugely powerful in characterising this genetic diversity (e.g. Geraldès et al. 2014; Fetter, Gugger and Keller, chapter 13, this volume) and in relating it to adaptation and the environment, whether through phenology (Brunner, Varkonyi-Gasic and Jones, chapter 11, this volume), abiotic stress (Bastiaanse, Theroux-Rancourt and Tixier, chapter 12, this volume) or species interactions (Plett and Platt, chapter 14, this volume).

Protection of Forests against Emerging Pests, Pathogens and Environmental Stressors Including Climate Change

Trees are large, long-lived resources that cannot move, and as such they are “sitting ducks” for pathologies of various sorts. Past epidemics such as Dutch elm disease and chestnut blight have been devastating. Plantation trees of a single genotype are

particularly vulnerable during their multi-decadal lifespan. The arrival of a new pest or pathogen, which may have multiple generations per year, is a real possibility during the relatively long rotation age. Resistance breeding, as well as planting of mixed genotypes, can mitigate risk on landscape scales. Genomics can help by permitting a gene-based understanding of resistance (Plett and Plett, chapter 14), and by providing breeding tools, as in “genomic selection” (Denis and Bouvet 2011). On the other hand, we are increasingly realising the beneficial role that fungal and bacterial endophytes in roots and leaves can play in pest and pathogen resistance. Such endophytes are very difficult to characterize by traditional means but lend themselves to metagenomic approaches.

Climate change is a major threat to forests, particularly as it now appears that the pace of change may exceed the ability of adapted genotypes to migrate to new climatic optima. Assisted migration (including planting trees that are adapted not to the present climate, but to future climate) is likely to emerge as a very important, and hotly debated, issue. Again, genomics has a role to play. Chapter 12 (Bastiaanse, Theroux-Rancourt and Aude Tixier) highlights our current understanding of how trees respond to drought stress and how genomics may aid in identifying genes and genotypes conferring resistance to drought and other abiotic stress.

The Future of Angiosperm Tree Genomics

The scientific issues highlighted in this introduction, and the other chapters in this book, indicate the health and excitement of the subject. Trees have traditionally been difficult to study and genomic tools now permit a catch-up. Much work needs to be done however. Despite the enormous importance of forests and forestry, this sector has fewer resources than the health and agriculture sectors that have already been deeply impacted by genomic technology.

The biology of trees provides many important problems that are unique to, or characteristic of, trees as opposed to other plants. For instance, most tree species are highly outbred and suffer from inbreeding depression. Hybrids, on the other hand, can show significant hybrid vigour. Indeed the highest yielding plantation crop in the world is a hybrid eucalyptus (*E. grandis* × *E. urophylla*) in Brazil. One explanation for hybrid vigour may be that trees carry a large genetic load of sublethal alleles, as suggested by genetic studies (Bradshaw and Stettler 1994). Surprisingly however, recent functional genomic work in poplar showed that most of the genome can be reduced to a haploid state without lethality (Henry et al. 2015). The interplay of inbreeding depression and hybrid vigour is just one example of many opportunities for new genomic studies in angiosperm trees. The size and genetic basis of load in trees has been a big unknown but genomic techniques offer potential solutions to this important problem. This is just one example of many opportunities for new genomic studies in angiosperm trees. As the chapters in this book will show, with powerful genomic tools at hand the solution to these and other problems in tree biology are within grasp.

Acknowledgements The editors wish to acknowledge the helpful assistance of Eric Stannard and other Springer staff in bringing this book to fruition. Work in the laboratory of QC is funded by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants Program (grant no. RGPIN-2014-05820). Work in the laboratory of AG is funded by USDA AFRI (grant no. 2015-67013-22891) and DOE Office of Science, Office of Biological and Environmental Research (BER grant no. DE-SC0007183).

Postscript While this book was in preparation, but after the manuscript of his chapter was submitted, we received the news of the tragic death of Carl Douglas in a climbing accident in the mountains of British Columbia. We have lost a great colleague and trusted friend. Carl was a true leader in the field of the genomics of angiosperm trees and will be greatly missed. We dedicate this volume as a small tribute to his memory.

References

- Bolin B, Keeling CD. Large-scale atmospheric mixing as deduced from the seasonal and meridional variations of carbon dioxide. *J Geophys Res.* 1963;68(13):3899–920.
- Bradshaw Jr HD, Stettler RF. Molecular genetics of growth and development in *Populus*. II Segregation distortion due to genetic load. *Theor Appl Genet.* 1994;89(5):551–8.
- Bräutigam K, Vining KJ, Lafon-Placette C, Fossdal CG, Mirouze M, Marcos JG, Fluch S, Fraga MF, Guevara M, Abarca D, Johnsen Ø. Epigenetic regulation of adaptive responses of forest tree species to the environment. *Ecol Evol.* 2013;3(2):399–415.
- Davin N, Edger PP, Hefer CA, Mizrahi E, Schuetz M, Smets E, Myburg AA, Douglas CJ, Schranz ME, Lens F. Functional network analysis of genes differentially expressed during xylogenesis in social woody Arabidopsis plants. *Plant J.* 2016;86(5):376–90.
- Denis M, Bouvet JM. Genomic selection in tree breeding: testing accuracy of prediction models including dominance effect. *BMC Proc.* 2011;5(Suppl 7, IUFRO Tree Biotechnology Conference 2011: From Genomes to Integration and Delivery):O13.
- Erwin TL. Tropical forests: their richness in Coleoptera and other arthropod species. *Coleopt Bull.* 1982;36(1):74–5.
- FAO. Global forest resources assessment 2000 – main report (FRA 2000). Forestry Paper No. 140, Food and Agriculture Organization of the United Nations, Rome. 2001.
- FAO. State of the world's forests – 2014. Rome: Food and Agriculture Organization of the United Nations; 2014.
- Geraldes A, Farzaneh N, Grassa CJ, McKown AD, Guy RD, Mansfield SD, Douglas CJ, Cronk QCB. Landscape genomics of *Populus trichocarpa*: the role of hybridization, limited gene flow, and natural selection in shaping patterns of population structure. *Evolution.* 2014;68:3260–80.
- Geraldes A, Hefer CA, Capron A, Kolosova N, Martinez-Nuñez F, Soolanayakanahally RY, Stanton B, Guy RD, Mansfield SD, Douglas CJ, Cronk QC. Recent Y chromosome divergence despite ancient origin of dioecy in poplars (*Populus*). *Mol Ecol.* 2015;24(13):3243–56.
- Hefer CA, Mizrahi E, Myburg AA, Douglas CJ, Mansfield SD. Comparative interrogation of the developing xylem transcriptomes of two wood-forming species: *Populus trichocarpa* and *Eucalyptus grandis*. *New Phytol.* 2015;206(4):1391–405.
- Henry IM, Zinkgraf MS, Groover AT, Comai L. A system for dosage-based functional genomics in poplar. *Plant Cell.* 2015;27(9):2370–83.

- Johnson LA, Douglas CJ. *Populus trichocarpa* MONOPTEROS/AUXIN RESPONSE FACTOR5 (ARF5) genes: comparative structure, sub-functionalization, and *Populus-Arabidopsis* micro-synten. *Botany*. 2007;85(11):1058–70.
- Mark J, Newton AC, Oldfield S, Rivers M. The international timber trade: a working list of commercial timber tree species. Richmond: Botanic Gardens Conservation International; 2014.
- McKown AD, Klápště J, Guy RD, Gerales A, Porth I, Hannemann J, Friedmann M, Muchero W, Tuskan GA, Ehling J, Cronk QC, El-Kassaby YA, Mansfield SD, Douglas CJ. Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *New Phytol*. 2014;203(2):535–53.
- Mock KE, Callahan CM, Islam-Faridi MN, Shaw JD, Rai HS, Sanderson SC, Rowe CA, Ryel RJ, Madritch MD, Gardner RS, Wolf PG. Widespread triploidy in western North American aspen (*Populus tremuloides*). *PLoS One*. 2012;7(10):e48406.
- Moyers BT, Rieseberg LH. Divergence in gene expression is uncoupled from divergence in coding sequence in a secondarily woody sunflower. *Int J Plant Sci*. 2013;174:1079–89.
- Porth I, Klápště J, Skyba O, Hannemann J, McKown AD, Guy RD, DiFazio SP, Muchero W, Ranjan P, Tuskan GA, Friedmann MC, Ehling J, Cronk QCB, El-Kassaby YA, Douglas CD, Mansfield SD. Genome-wide association mapping for wood characteristics in *Populus* identifies an array of candidate single nucleotide polymorphisms. *New Phytol*. 2013;200:710–26.
- RBG Kew. The state of the world's plants report – 2016. Kew: Royal Botanic Gardens; 2016.
- Rottmann WH, Meilan R, Sheppard LA, Brunner AM, Skinner JS, Ma C, Cheng S, Jouanin L, Pilate G, Strauss SH. Diverse effects of overexpression of LEAFY and PTLF, a poplar (*Populus*) homolog of LEAFY/FLORICAULA, in transgenic poplar and *Arabidopsis*. *Plant J*. 2000;22(3):235–45.
- Segura V, Cilas C, Costes E. Dissecting apple tree architecture into genetic, ontogenetic and environmental effects: mixed linear modelling of repeated spatial and temporal measures. *New Phytol*. 2008;178:302–14.
- Xu B, Ohtani M, Yamaguchi M, Toyooka K, Wakazaki M, Sato M, Kubo M, Nakano Y, Sano R, Hiwatashi Y, Murata T. Contribution of NAC transcription factors to plant adaptation to land. *Science*. 2014;343:1505–8.

Contents

The Evolution of Angiosperm Trees: From Palaeobotany to Genomics	1
Quentin C.B. Cronk and Félix Forest	
Variation in Angiosperm Wood Structure and its Physiological and Evolutionary Significance	19
Rachel Spicer	
<i>Populus</i> as a Model Tree	61
Carl J. Douglas	
Emerging Genomics of Angiosperm Trees	85
Elizabeth Sollars and Richard Buggs	
Structural Genomics of Angiosperm Trees: Genome Duplications, Ploidy, and Repeat Sequences	101
Nathaniel R. Street	
Evolutionary Histories of Gene Families in Angiosperm Trees	121
S.G. Hussey, Jill L. Wegrzyn, and H.A. Vasquez-Gross	
Breeding Systems, Mating Systems, and Genomics of Gender Determination in Angiosperm Trees	139
Matthew S. Olson, J.L. Hamrick, and Richard Moore	
Genomics of Speciation in Temperate and Boreal Angiosperm Trees	159
R. Bawa and J.A. Holliday	
Perennial Growth, Form and Architecture of Angiosperm Trees	179
David J. Hearn	
The Genomics of Wood Formation in Angiosperm Trees	205
Xinqiang He and Andrew T. Groover	
Phase Change and Phenology in Trees	227
Amy M. Brunner, Erika Varkonyi-Gasic, and Rebecca C. Jones	

Abiotic Stress 275
Héloïse Bastiaanse, Guillaume Thérroux-Rancourt, and Aude Tixier

**Landscape Genomics of Angiosperm Trees:
From Historic Roots to Discovering New Branches
of Adaptive Evolution** 303
Karl C. Fetter, Paul F. Gugger, and Stephen R. Keller

**Pathogenic and Mutualistic Symbiotic Interactions
in Angiosperm Trees** 335
K.L. Plett and J.M. Plett

Erratum 355

Index 357

Contributors

Héloïse Bastiaanse Pacific Southwest Research Station, US Forest Service, Davis, CA, USA

R. Bawa Department of Forest Resources and Environmental Conservation, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Amy M. Brunner Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, USA

Richard Buggs School of Biological and Chemical Sciences, Queen Mary University of London, London, UK

Royal Botanic Gardens Kew, Richmond, Surrey, UK

Quentin C.B. Cronk University of British Columbia, Vancouver, BC, Canada

Carl J. Douglas Department of Botany, University of British Columbia, Vancouver, BC, Canada

Karl C. Fetter Department of Plant Biology, University of Vermont, Burlington, VT, USA

Félix Forest Royal Botanic Gardens, Kew, Richmond, Surrey, UK

Andrew T. Groover Pacific Southwest Research Station, US Forest Service, Davis, CA, USA

Department of Plant Biology, University of California Davis, Davis, CA, USA

Paul F. Gugger Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, MD, USA

J.L. Hamrick Department of Plant Biology, University of Georgia, Athens, GA, USA

Xinqiang He School of Life Sciences, Peking University, Beijing, China

David J. Hearn Towson University, Department of Biological Sciences, Towson, MD, USA

J.A. Holliday Department of Forest Resources and Environmental Conservation, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

S.G. Hussey Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), Genomics Research Institute (GRI), University of Pretoria, Pretoria, South Africa

Rebecca C. Jones School of Biological Sciences, University of Tasmania, Hobart, TAS, Australia

Stephen R. Keller Department of Plant Biology, University of Vermont, Burlington, VT, USA

Richard Moore Department of Biology, Miami University, Oxford, OH, USA

Matthew S. Olson Department of Biological Sciences, Texas Tech University, Lubbock, TX, USA

K.L. Plett Hawkesbury Institute for the Environment, Western Sydney University, Richmond, NSW, Australia

J.M. Plett Hawkesbury Institute for the Environment, Western Sydney University, Richmond, NSW, Australia

Elizabeth Sollars School of Biological and Chemical Sciences, Queen Mary University of London, London, UK

Royal Botanic Gardens Kew, Richmond, Surrey, UK

Rachel Spicer Connecticut College, Department of Botany, New London, CT, USA

Nathaniel R. Street Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, Umeå, Sweden

Guillaume Thérout-Rancourt Department of Plant Sciences, University of California, Davis, CA, USA

Aude Tixier Department of Plant Sciences, University of California, Davis, CA, USA

Erika Varkonyi-Gasic The New Zealand Institute for Plant & Food Research Limited (Plant & Food Research), Mt Albert, Private Bag 92169, Auckland Mail Centre, Auckland, New Zealand

H.A. Vasquez-Gross Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, USA

Jill L. Wegrzyn Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, USA

The Evolution of Angiosperm Trees: From Palaeobotany to Genomics

Quentin C.B. Cronk, Félix Forest

Abstract Angiosperm trees now rival the largest conifers in height and many species reach over 80 m high. The large tree life form, with extensive secondary xylem, originated with the progymnosperms and gymnosperms in the Devonian and Carboniferous. However evidence suggests that the ancestor of extant angiosperms was not a tree but either a herb or understory shrub. Angiosperm fossil woods are rare in the early Cretaceous but become common in the mid-Cretaceous. The “reinvention” of wood in the Cretaceous produced a novel xylary morphospace that has since been extensively explored by subsequent evolution. Today, large timber trees are absent in the early diverging lineages of the angiosperms, and conventional wood has been lost in the monocots. There are a few timber trees in the magnoliid clade. Most timber trees are in the rosid clade, particularly the fabids (e.g. Fabaceae) but also in the Malvids (e.g. Meliaceae). Timber trees are less common in the strongly herbaceous asterid clade but some important timbers are also found in this lineage such as teak, *Tectona grandis* (Lamiaceae). Genomic resources for angiosperm trees are developing rapidly and this, coupled with the huge variation in woody habit, make angiosperm trees a highly promising comparative system for understanding wood evolution at the molecular level.

Keywords Wood • Fossils • Evolution • Xylogenesis

Introduction

The tallest known angiosperm tree is “Centurion”, a large *Eucalyptus regnans* from Tasmania measuring 99.6 m in height, 12 m around at the base, with an above ground biomass of 215 tonnes and an annual increment approaching one tonne

Q.C.B. Cronk (✉)
University of British Columbia, Vancouver, BC V6T 1Z4, Canada
e-mail: quentin.cronk@ubc.ca

F. Forest
Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK

© Springer International Publishing AG 2017

A.T. Groover and Q.C.B. Cronk (eds.), *Comparative and Evolutionary Genomics of Angiosperm Trees*, Plant Genetics and Genomics: Crops and Models,

DOI 10.1007/7397_2016_31

Published Online: 03 Feb 2017

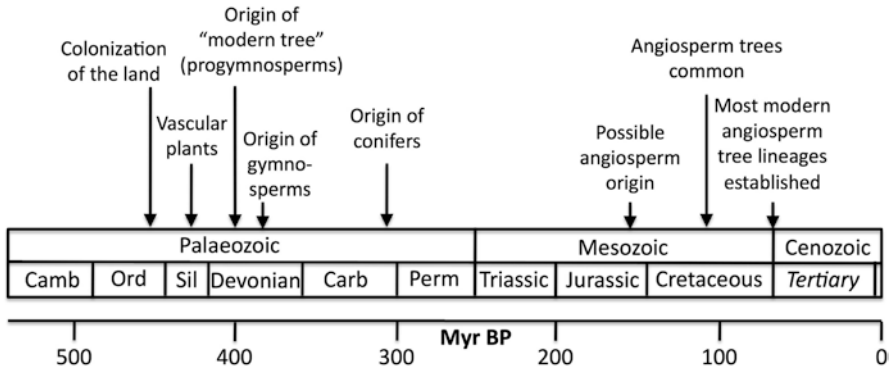


Fig. 1 Timeline of the evolution of the angiosperm tree habit (see also Table 1)

(Sillett et al. 2015). This is a big tree by any measure. *E. regnans* is only the third largest tree species after two conifers, the coast redwood and coastal Douglas⁺ fir of western north America. However, a disputed nineteenth century record, at 132.6 m, would make *E. regnans* the world's tallest tree, if correct.

Although conifers hold most of the height records, angiosperms are notable in certain categories. Their capacity for clonal growth means that “Pando”, a large aspen (*Populus tremuloides*) clone is the largest organism by biomass. It is estimated that the Pando clone covers 46.3 hectares (DeWoody et al. 2008). *Ficus benghalensis*, the Indian banyan tree, has the largest spread of a single tree. Descending prop roots stabilize branches which can grow horizontally almost indefinitely. One individual, “Thimmamma Marrimanu”, is said to occupy an area of 1.9 hectares.

It was not always so. It is likely that angiosperms were once small and evolved in the shadow of gymnosperms. They took millions of years to attain ecological dominance over gymnosperms (at least in most forest biomes) and millions more to fully rival the tallest gymnosperms in size. In terms of geological history however, the rise of the angiosperms was spectacularly rapid. This chapter will follow the road to ecological dominance of the angiosperm tree. Tree genomics holds the promise of being able to understand this rise and to understand the genomic toolbox (Schrader et al. 2004; Groover 2005; Carocha et al. 2015) used to build such forest giants. The origins of this toolkit lie in the distant evolutionary past. It is therefore useful to start the journey at the earliest land plants and what genes they brought with them, from their aquatic ancestors, onto the land (Floyd and Bowman 2007), around 470 million years ago (Mya). A timeline of land plant evolution is given in Fig. 1. By comparing genomes of trees with those of the early diverging extant land plants, such as the moss *Physcomitrella patens*, we can determine the conserved developmental modules which have been reshaped, co-opted or re-used, in order to build massive organisms (Xu et al. 2014).

Early Land Plants and the Building Blocks of Complexity

Increased knowledge of the biology of *Physcomitrella patens* has allowed us a glimpse into what aspects of the molecular machinery of woodiness is

shared with early divergent clades of land plants and may be considered to have involved soon after the colonisation of the land. Three examples can be cited: (1) The KNOX proteins. These are essential for meristems function and are involved in cambial activity (Champagne and Ashton 2001; Singer and Ashton 2007; Sakakibara et al. 2008). (2) The NAC transcription factors. These are essential as regulators of xylogenesis (Xu et al. 2014). (3) The lignin pathway, including genes like 4CL (Silber et al. 2008; Souza et al. 2008). There are many examples of evolution re-using and co-opting existing genes rather than evolving novel genes de novo. It is therefore unsurprising that we can find the molecular origins tree-building in distantly related simple plants and infer their presence in the common ancestor of extant land plants.

Vascularization: The Lignin Revolution

The earliest land plants had no vascular tissue and were unable to transport water for long distances. This constrained their size and ensured dependence on external water: they were ectohydric, with water transport by capillarity along the external surface of the plant.

The endohydric condition (internal water transport) is characteristic of the more robust mosses and therefore predates vascularisation and its precursor. Endohydric mosses can absorb substrate water through basal rhizoids, and are resistant to water loss because of a cuticle-like external layer. In their stems they have well-developed conducting strands or hydroids (Zamski and Trachtenberg 1976). Examples of endohydric mosses include *Polytrichum* spp., and the largest mosses, such as *Dawsonia superba* and *Dendroligotricum dendroides* (Atala and Alfaro 2012). The hydroids together make up a conducting tissue called the hydrome (there is also a phloem-like tissue called leptome). It is debatable whether the hydrome and leptome of mosses are directly homologous to xylem and phloem or whether they represent independent evolution. Nevertheless the early stages of evolution of xylem and phloem must have been equivalent to hydrome and leptome, and therefore we can use those tissues to help us understand vascularisation (Edwards et al. 2003). Hydrome is often particularly well-developed in the sporophytes of mosses, which also have stomata. Stomata therefore predate the origin of xylem but their association with hydrome in sporophytes of mosses indicates the likely co-evolution between conducting tissues and stomata (Ligrone et al. 2012).

Plants with simple patterns of xylem and phloem (primary vasculature) occur in the fossil record (Table 1) with the appearance of the rhyniophytes (Kenrick and Crane 1991) (leafless herbs, now extinct), lycophytes (microphyllous herbs) and ferns (megaphyllous herbs). The lycophytes and ferns also evolved strategies to form tree-like organisms, which will be discussed later, but these are more akin to giant herbs than true trees.

The tracheid, eventually with its lignified cell wall (Boyce et al. 2003), represented a great improvement over the hydroid in terms of water transport and the

existence of xylem (tracheid tissue) permitted water transport over greater distances, thus allowing for greater organismal stature.

The development of primary xylem is a first step in the ontogeny of shoots, even in large extant trees, and there is no reason to suppose that primary vascularization is not fundamentally homologous throughout vascular plants. Unfortunately there are rather few simple vascular plants with completely sequenced genomes. *Selaginella* (a lycophyte) is one (Banks et al. 2011), and there are plans to sequence the model fern, *Ceratopteris richardii* (Veronica Di Stilio, pers. comm.). Further resources of this sort would be extremely valuable in studying the evolution of vascularisation.

Currently our knowledge of primary vascularisation at the molecular level comes, of course, from *Arabidopsis*. The developmental course involves, first, the differentiation and division of procambial strands (Yang and Wang 2016). Second comes differentiation into protoxylem and protophloem and thirdly the formation of metaxylem and metaphloem. Genes involved in these processes are numerous but include HD-ZIP III genes that direct xylem development. Polar auxin transport (PAT) is very important: auxin mediated transcription of MONOPTEROS (MP) leads to the expression of the HD-ZIP gene AtHB8, as well as the PIN1 auxin transporter which maintains MP transcription in a positive feedback loop (Ohashi-Ito and Fukuda 2010). A key question is how these gene modules of primary vascularization evolved in the early history of life on land (Xu et al. 2014).

Vascular Elaboration: Trees without Woody Trunks

Woody trunks, as discussed in the next section, are a feature of the progymnosperms, gymnosperms and angiosperms, together forming a single clade the lignophytes. However, tree-like organisms evolved in the lycophytes and ferns, despite less extensive development of secondary vasculature. There are no extant tree-like lycophytes, but in the fossil record of the Carboniferous, *Lepidodendron* and *Sigillaria* grew to tree-like proportions in the coal swamps (Phillips and Dimichele 1992; DiMichele and Bateman 1996). These tree lycophytes had some secondary xylem, produced from a unifacial, and more or less continuous, cambium. However, this secondary xylem did not occupy the bulk of the stem, but only a small central core. The majority of the stem was parenchymatous with structural rigidity provided by a well-developed outer bark. They could almost be described as “giant herbs”, particularly as the arborescent stems were determinate and short lived. The closest living relative to these giant herbs is now the quillwort group, comprising the small aquatic *Isoetes* and its small shrubby relative *Stylites* (Karrfalt and Hunter 1980; Larsen and Rydin 2016). These extant plants would be an exciting genomic resource for the study of the origin of the tree habit in fossil lycophytes. However, no genomic resources are currently available for quillworts.

The tree habit has also evolved in ferns (monilophytes) and indeed we still have extant examples in tree ferns (Pteridophyta), such as *Dicksonia* and *Cyathea*. These

can grow several metres high but have no secondary xylem. Instead they have a complex primary vasculature (dictyostele) in a massive, mainly parenchymatous stem. Much of the structural rigidity of the stem comes from a tough outer layer of persistent leaf-bases or adventitious rootlets. These too have the aspect of a “giant herb”. An extinct group of free-sporing plants, the Cladoxylales, attained considerable stature (Soria and Meyer-Berthaud 2004). The cladoxylalean *Eospermatopteris* (*Wattieza*) has left Devonian fossils in New York State that are massive trunks from organisms formerly 8 m or more high (Stein et al. 2007). Finally, a rather divergent group of ferns (in the broad sense), the horsetails or sphenophytes, are also represented by large organisms in the fossil record. The arborescent sphenophytes produced a small amount of secondary xylem in their long and narrow stems, which grew many metres high (Rossler and Noll 2006; Roessler et al. 2012). The stems are cylinders of wood around a central pith. This is a different strategy but they are still hard to fit into the modern concept of “tree” and again have something of a resemblance to “giant herbs”, particularly as the aerial stems were likely mostly determinate and short-lived.

The Evolution of Woody Trunks: The Progymnosperm Legacy

The woody trunk or “hyperstele” (massive secondary development of the primary vasculature or stele) evolved first in the progymnosperms. The formation of woody trunks required the extensive production of secondary vascular tissue from a persistent and highly active continuous cambium. This permits massive stems (trunks) that are very largely composed of xylem with little or no pith (although some seed plants do have trunks with extensive pith - see below). The structural strength of such secondarily thickened stems allows for indeterminate growth into very large organisms. This contrasts with the determinate tree-like stems of arborescent lycophytes and horsetails which spring up to considerable heights to reproduce, but might not persist.

The position of the first “modern tree” is generally given to the progymnosperm *Archaeopteris* (Meyer-Berthaud et al. 1999). Progymnosperms do not have seeds but are instead free-sporing, indicating that the woody trunk evolved well before the seed. Progymnosperms appear to have had tracheids with bordered pits (Dannenhoffer and Bonamo 2003). Fossil wood of *Archaeopteris* is commonly referred to in paleontological literature as *Callixylon* (Beck 1960) and studies of well preserved *Callixylon* wood shows not only the presence of bordered pits but also of possible torus structures (Beck et al. 1982). If this interpretation is correct, then *Archaeopteris* shows advanced features of conifer wood. Wood features characteristic of conifers therefore predates conifers and even the seed habit.

When gymnosperms appear in the fossil record, they carry forward the massive wood construction of the progymnosperms (Savidge 2008). Their generally large size, together with the advantage of the seed, gave rise to ecological success. Millions of years of increasing gymnosperm dominance of the earth eventually left

only epiphytic, aquatic, marginal or forest understory niches for lycophytes and ferns. However, not all gymnosperms have the massive trunks of conifers. The solid wood and pith-free trunks of conifers is a form of construction termed “pynoxylic”, whereas the trunk of cycads is “manoxylic”, with a large pith. *Ginkgo* has a mixed stem anatomy with the short shoots being manoxylic and the long shoots pynoxylic. The final group of extant gymnosperms, Gnetales, is remarkable for its diverse habits, from lianas to small shrubs, and for the presence of vessels, independently derived from those of angiosperms. This diversity indicates the ability of this lineage to utilize different modes of woodiness in addition to the massive woodiness of conifers.

It is worth noting that these new innovations leading to massive wood partly involve the distribution and number of tracheids and partly changes to functional efficiency the tracheid itself. The basic tracheid building block is little changed from the earliest vascular plants to appear on the land, but some innovations have arisen, such as the bordered pit complete with torus (characteristic of some conifers). Many ferns and lycophytes have undifferentiated pit margins (although bordered pits have been noted in some).

Far more conspicuous, however, is the huge increase in distribution and volume of xylem. The key innovation for these lies in the persistent and continuous cambium. The challenge for genomics and development is therefore in understanding the specification and maintenance of the cambium (Groover and Robischon 2006).

Wood Reinvention: The Evolution of Angiospermous Wood

The first unequivocal angiosperm fossils are pollen grains that first appear in the early Cretaceous (from 135 Mya). At first rare, angiosperm pollen quickly increases in abundance, first in low latitudes, later in higher latitudes. By the end of the Cretaceous the angiosperms were clearly the dominant organisms of the biosphere. The first macrofossil evidence is *Archaeofructus* (Sun et al. 2002), an aquatic herb from the early Cretaceous (125 Mya). However, molecular dating studies consistently suggest angiosperm origins well before this, usually some time in the Jurassic. The fossil flower *Euanthus* (Liu and Wang 2016), from the late Jurassic (160 Mya) is not universally accepted as an angiosperm. Enigmatic fossils from the Triassic such as the “monocot-like” leaf *Sanmiguelia* and some angiosperm-like triassic fossil pollen (Hochuli and Feist-Burkhardt 2013), hint at an even earlier origin. If angiosperms did originate before the Cretaceous the problem becomes how they remained so rare for so long. Darwin (1903) recognized this problem when he wrote in a letter in 1875: “the presence of even one true angiosperm in the Lower Chalk [early Cretaceous] makes [one] inclined to conjecture that plant[s] of this great division must have been largely developed in some isolated area, whence owing to geographical changes, they at last succeeded in escaping, and spread quickly over the world” (Darwin and Seward 1903).

Features of the first angiosperm may be looked for by examining the early diverging lineages *Amborella*, Nymphaeales and Austrobaileyales (Table 2). There are two growth forms here: *Amborella* and the Austrobaileyales are generally shrubs or lianas adapted to low light and high soil disturbance in humid tropical forest understory, such as stream banks in tall forest. This model of the early angiosperm is the “dark and disturbed” hypothesis (Feild et al. 2004). The Nymphaeales are very different. They are aquatic and adapted to sunny open water. This has led to the suggestion that the early angiosperms might have been aquatic: the “aquatic palaeoherb” hypothesis (Sun et al. 2002; Feild and Arens 2005).

Either way, there is no evidence that the ancestral angiosperm was a tall forest tree. Interestingly, both *Amborella* and the Austrobaileyales have a seedling phase in which they form multiple scandent shoots from a basal lignotuber (Feild and Arens 2005). Sometimes the scandent habit persists as in lianous species of *Schisandra*, *Austrobaileya* and *Trimenia*. This is interesting as the lianous habit is potentially a driver for the evolution of vessels, as may have been the case in *Gnetum* (of which many species are lianous). High hydraulic conductivity per unit area is important in the thin stems of lianas. It should be noted however that lianous species without vessels are known (Feild et al. 2012).

Waterlilies (Nymphaeaceae) have large creeping rhizomes that are often perennial. Despite the large size they have no secondary xylem and do not form a vascular cambium (although there may be a cork cambium). Instead the primary vasculature is scattered and the bulk of the rhizome is of aerenchymatous ground tissue.

Whatever specialized niche, whether understory shrub, liana or aquatic herb that early angiosperms occupied, the ecological conditions were apparently permissive to a distinctive “reinvention” of wood: now with vessels and small bordered pits with a homogeneous pit membrane (lacking a torus). Notably, if the ancestral angiosperm was an aquatic herb then woodiness, and even the vascular cambium (if this was ancestrally lost), might have had to be re-evolved.

Whatever the ecological drivers, when angiosperms increased in numbers and stature in the mid-Cretaceous to compete with gymnosperms in the forest canopy, they possessed a remarkable new vesseliferous wood. There are only a few angiosperms with only tracheids (i.e. vessels completely absent) this feature may be ancestral in *Amborella* but it is an evolutionary reversal elsewhere (Winteraceae in the magnoliids, and *Trochodendron* in the eudicots). As Feild and Arens state: “vessel origin appears to allow for the exploitation over new morphospace of xylem hydraulic design” (Feild and Arens 2005). This new morphospace has been fully exploited in subsequent angiosperm evolution.

A likely further reinvention of wood occurred in the monocot clade, which appears to have diversified from an herbaceous ancestor. The palms are monocots with an anomalous “wood”, formed from extended production of fibre-capped vascular bundles distributed throughout the ground tissue. Compared to dicotyledon xylogenesis this seems bizarre and it produces “wood” unlike any other. This is not wood if that is defined as secondary xylem, but if wood is defined more generally as usable lumber then the word applies. Palm wood is functionally very effective, sup-

porting tall trees (up to 60 m in the case of the wax palm *Ceroxylon quindiuense*) and producing internationally traded and locally important hard tropical lumbers such as “red palm” lumber from *Cocos nucifera* and “black palm” lumber from *Borassus flabellifer*. The trunk is filled with functioning vascular bundles and there is no heartwood of non-functioning vessel elements. Furthermore the absence of a peripheral cambium reduces vulnerability to fire (Tomlinson 2006). The recent sequencing of the genomes of oil palm (*Elaeis guinaeensis*) (Singh et al. 2013) and date palm (*Phoenix dactylifera*) (Al-Dous et al. 2011; Al-Mssallem et al. 2013) has created opportunities for understanding the distinctive growth of palm trunks at the molecular developmental level.

There are now genomes available or soon-to-be available for several early divergent clades of the angiosperms, notably *Amborella* (Albert et al. 2013). These genomes will be of great significance for comparative work that seeks to elucidate the evolutionary developmental origin of angiosperm wood. Finally, mention should be made of a bizarre rootless aquatic dicot angiosperm, *Ceratophyllum* (Iwamoto et al. 2015), which lacks xylem, even primary xylem. As the xylogenesis pathway has been deleted in this plant, it represents a “natural knockout” experiment, which might one day be attractive to researchers.

Forest Giants: The Origin of Large Angiospermous Trees

The rapid rise and diversification of the angiosperms during the Cretaceous is well documented from fossil evidence. However much of the early differentiation appears to have been in the form of herbs (Jud 2015) and shrubs (Feild and Arens 2005). Fossil angiosperm wood does not appear until the Aptian and Albian (126–99 Mya) and does not become common until the late Cretaceous (84–65.5 Mya). At the same time, findings of fossil gymnosperm wood fall (Peralta-Medina and Falcon-Lang 2012). Recent fossil flower finds allow the identification of magnoliids as well as early diverging clades of eudicots (Proteales and Buxales) (Doyle 2015). By the end of the Cretaceous (65 Mya) the majority of eudicot lineages were well established and the abundance of fossil woods indicate that large eudicots were dominant in forests globally (Table 3).

The first diverging extant eudicot lineages include many herbaceous and shrubby clades (Bremer et al. 2009; Angiosperm Phylogeny Group 2016). However, the order Proteales includes the plane trees (*Platanus*), which are of large stature. Extinct platanoids (Maslova 2010) of various kinds may have been among the first eudicot forest dominants.

The delimitation of the eudicot clades used here (Table 3) follows the recent APG classifications (Bremer et al. 2009; Angiosperm Phylogeny Group 2016). The asterid clade of eudicots includes rather few large trees (although many herbs, as in the predominantly herbaceous family Asteraceae). *Gmelina* and *Tectona* (teak) in the mint family Lamiaceae are notable exceptions. By contrast, the rosid clade (containing about a quarter of flowering plants) contains the majority of

Table 1 Some major plant lineages mentioned in text

Name	Origin (approx.)	Notes
Liverworts (marchantiophytes)	Mid-Ordovician (470 Mya)	Earliest land plants (mid-Ordovician) are probably referable here; or at least were similar in form to modern marchantiophytes
Mosses (bryophytes)	? Silurian (c. 440 Mya)	Fossil record poor
Hornworts (anthocerotophytes)	? Silurian (c. 435 Mya)	Fossil record poor
Rhyniophytes (extinct)	Mid-Silurian (430 Mya)	The early vascular plant, the rhyniophyte <i>Cooksonia</i> is first known from fossils in Ireland (Edwards and Feehan 1980).
Lycophytes	Late Silurian (420 Mya)	The first fossil evidence is the relatively small lycophyte <i>Baragwanathia</i> from Australia
Ferns (monilophytes)	Likely late Silurian (420 Mya)	The earliest fern (in the broad sense) fossil is generally considered to be the mid-Devonian (c. 390 Mya) <i>Ibyka</i> . (possibly representing a lineage ancestral to sphenopsids) However the fern and lycophyte lineages are likely to have split before this
Progymnosperms (extinct)	Mid Devonian (c. 400 Mya)	The mid-Devonian Aneurophytales are the first exemplars. The first “modern tree”, <i>Archaeopteris</i> , first appears in the upper Devonian (380 Mya)
Gymnosperms	Late Devonian (385 Mya)	The first gymnosperms are not referable to any extant groups. The first members of extant groups, such as the first putative conifers, arose in the late Carboniferous (310 Mya)
Angiosperms	Possibly late Jurassic (160 Mya)	Unequivocal angiosperm pollen first appears in the early Cretaceous (from 135 Mya). The fossil flower <i>Euanthus</i> (Liu and Wang 2016), if accepted as angiospermous, pushes the origin back at least to the late Jurassic (160 Mya)

Table 2 Characters of the major clades of angiosperms

Clade	Life form	Xylem characters
<i>Amborella</i>	Shrub	Extensive bifacial vascular cambium, tracheids only
Nymphaeales	Aquatic herbs	Primary xylem only, vessels
Austrobaileyales	Shrubs, lianas	Extensive bifacial vascular cambium, vessels
Magnoliids	Shrubs, trees, lianas, herbs	Extensive bifacial vascular cambium, vessels (but tracheids in Winteraceae)
Monocots	Herbs (rarely trees, e.g. palms)	Primary xylem only (but sometimes with anomalous secondary xylogenesis, e.g. palms), vessels
Eudicots	Herbs, shrubs, lianas, trees	Extensive bifacial vascular cambium, vessels (but tracheids in <i>Trochodendron</i> and <i>Tetracentron</i>)

In addition the Chloranthaceae (a small clade of tropical shrubs of uncertain placement but near magnoliids) has characters of magnoliids. *Ceratophyllum*, an anomalous aquatic genus, has no vasculature