Ilia J. Leitch
Editor-in-chief

Johann Greilhuber
Jaroslav Doležel
Jonathan F. Wendel
Editors

Plant Genome Diversity

Volume 2: Physical Structure, Behaviour and Evolution of Plant Genomes

Springer

Plant Genome Diversity Volume 2

Ilia J. Leitch
Editor-in-chief

Johann Greilhuber · Jaroslav Doležel · Jonathan F. Wendel
Editors

Plant Genome Diversity Volume 2

Physical Structure, Behaviour and Evolution of Plant Genomes

Springer

Editor-in-chief
Ilia J. Leitch
Jodrell Laboratory
Royal Botanic Gardens,
Kew Richmond, Surrey
United Kingdom

Editors
Johann Greilhuber
Department of Systematic and Evolutionary Botany
Faculty of Life Sciences
University of Vienna
Vienna
Austria

Jaroslav Doležel
Institute of Experimental Botany ASCR
Centre of the Region Hana for
Biotechnological and Agricultural
Research
Olomouc
Czech Republic

Jonathan F. Wendel
Department of Botany
Iowa State University
Ames, Iowa
USA

ISBN 978-3-7091-1159-8 ISBN 978-3-7091-1160-4 (eBook)
DOI 10.1007/978-3-7091-1160-4
Springer Wien Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012935228

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Ever since the origin of life, the evolution of living organisms and their hereditary information has been accompanied by the development of genetic machinery capable of storing, utilizing, and transmitting this information between generations. Importantly, this machinery has had to be flexible, able to respond to the environment and evolve. A characteristic feature of the genetic machinery in eukaryotes is the partitioning of the hereditary information into smaller portions—chromosomes. Indeed, the appearance of linear chromosomes was one of the great evolutionary inventions and paved the way for the formation of large and complex genomes, in plants as well as in animals. Any consideration of plant genome structure, evolution, and function is thus incomplete if it does not take into account its higher-order structure and the behaviour of its principal units—the chromosomes.

The chromosome theory of heredity, which linked the behaviour of Mendel's "factors" (units of inheritance) with that of chromosomes, was coined by Walter S. Sutton more than a century ago. This was followed, only a few decades later by Cyril D. Darlington's demonstration that the behaviour of chromosomes and meiotic crossing over in particular, was the main force behind evolution as opposed to single gene mutations and deletions. This set in motion the quest to understand the nature of inheritance, leading to the discovery of the structure of DNA and the advent of molecular biology and genomics. At this point the goal seemed clear—all that was needed was to establish the sequence of bases in the DNA. However, as increasing amounts of DNA sequences were generated, it became obvious that there was still a lot to discover about how DNA was organized within chromosomes and how the DNA sequence information was interpreted, processed, and utilized in the nuclear and cellular environments. The days when the DNA sequence itself was considered a holy grail are over and we now know that things are considerably more complicated.

Luckily, progress in genomics has been complemented by advances in understanding the dynamic structure of chromatin, the organization of interphase nuclei, and the behaviour of chromosomes during mitosis and meiosis. The latter includes novel insights into modified cell cycles, which may lead to chromosomes with more than two chromatids. Impressive progress has also been made in understanding the origin and function of specialized chromosomes (e.g., B chromosomes and sex chromosomes) and in appreciating the extent and significance of polyploidy in plant evolution. Although the frequent occurrence of polyploidy has been known for a long time based on chromosome counts and behaviour, the advent of DNA sequencing and comparative genomics has been instrumental in uncovering evidence of further rounds of polyploidy buried within the genome and now no longer visible at the chromosome level. Such studies have reinforced and extended our understanding of the significance of this mechanism as one of the main forces underlying the evolution and large diversity of many plant genomes. The effect has been multiplied by the extensive structural chromosome changes, which, together with alterations in chromosome number and genome size, can accompany plant speciation.

To fully understand and appreciate the diversity, functioning, and evolution of plant genomes, a holistic knowledge of the current status in each of these individual areas is vital, yet there is no single accessible source of information currently available. Thus, we felt it timely to fill this gap.

This book is the second volume of a two-volume set on Plant Genome Diversity. Our aim is to assist students and researchers by providing as complete an overview as possible of each respective area of research. We have succeeded in engaging leading experts in each field who describe the current state-of-the-art knowledge without overwhelming the reader with details that can be found elsewhere. What we offer in the present volume are 20 chapters whose topics have been chosen carefully to provide a complete picture. Each chapter can stand on its own and thus the reader does not need to read all chapters if he/she is only interested in a specific area. We sincerely hope that this model serves our readers well. It is up to them to decide if we have succeeded.

The 20 chapters deal with individual aspects of plant genome structure, function, and evolution and they are divided into five informal sections.

Evolutionary Framework for Studying the Diversity of Plant Genomes

Although we do not necessarily expect our readers to read all chapters, we do recommend that those interested in the evolution of plant genomes read the first chapter by Soltis and Soltis (Chap. 1) who provide an overview of plant phylogeny, with an emphasis on angiosperms. Among other things, they highlight research projects that have deposited phylogenetic trees in public databases and can be downloaded for analysis.

Architecture and Dynamics of the Plant Cell Nucleus

In nondiving cells, the chromosomes are organized within the nucleus, although the structural and functional complexity of this organization is still poorly understood. One can hardly imagine the intricacy of interactions of DNA with various molecules necessary to control tens of thousands of genes and process transcripts of genic and non-genic DNA. In addition, this is all taking place in a tightly packed nuclear environment, which also harbors structures needed for DNA synthesis and repair, chromosome reduplication, posttranscriptional modifications, and synthesis of ribosomal subunits, to name but a few. Jones and Langdon (Chap. 2) review nuclear organization and discuss the consequences of interspecific hybridization, which results in two different genomes being accommodated within a single nucleus. This cohabitation may not be peaceful and can result in dramatic structural and epigenetic reorganizations in subsequent generations.

The way DNA is organized and packaged into chromatin, particularly at the higher-order level has never been entirely clear although numerous models have been proposed. However, recent discoveries question even the existence of the 30-nm fibre, which traditionally has been considered to originate by folding the 11 nm nucleosome fibre. In Chap. 3, Takata et al. address this topic by describing the composition of chromatin in relation to chromosome condensation and DNA packing. Moreover, they present a novel model for chromosome structure, which suggests that the nucleosome fibres exist in a highly disordered state and do not form 30-nm chromatin fibres at all.

In addition to separating the nuclear environment from the cytoplasm, the nuclear envelope performs many important functions; one of which is the control of molecular traffic between both cellular compartments. Kiseleva et al. (Chap. 4) describe the composition of the nuclear envelope, the nuclear pore complexes, and their assembly and function and discuss possible interactions of the envelope with the cytoplasmic and nucleoplasmic components.

Nuclei are known to contain a variety of nuclear bodies, but only the nucleolus can be easily identified by optical light microcopy. It is where the cell produces ribosomes, which are required by the cell in large numbers. Shaw (Chap. 5) summarizes the current state of knowledge of the nucleolus, which is formed on nucleolar organizing regions of chromosomes.

For plants to grow and reproduce, the cells must divide either through mitosis or meiosis. The aim is that the hereditary material is faithfully transmitted to the daughter cells. Not only must the chromosomes be fully reduplicated but their chromatids must separate at the right moment and move in the right direction to form daughter nuclei. In addition, the nuclear envelope breaks down during cell division and this represents an additional major challenge for the genetic apparatus. Magyar et al. (Chap. 6) provide an insightful review on molecular events underlying the mitotic cell cycle and mitosis itself.

Following cell division, cell and tissue differentiation is often accompanied by modified cell cycles in which the mitosis step is omitted and the nuclear envelope does not break down. Maluszynska et al. (Chap. 7) outline these different types of endopolyploidy and describe the molecular pathways involved in switching from the mitotic to the endopolyploidization cycle and how the number of endocycles are regulated. They also review the occurrence of endopolyploidy, its biological significance, and the structure of endopolyploid nuclei.

The production of gametes provides an important means to generate genetic variation via recombination of parental chromatids and their random segregation. Given the complexity of the process, it is not surprising that it is unclear exactly how the mitotic machinery is modified for the purpose of meiosis. Nevertheless, Jenczewski et al. (Chap. 8) describe the current knowledge in this area, covering chromosome dynamics during meiosis, initiation of meiotic recombination, regulation of double strand break repair, crossover formation and interference, genetic control of crossing-over formation, and its distribution in polyploids.

Karyotype Diversity Across Plants and Trends in Evolution

One of the ways in which the diversity of plant genomes is manifested is through a wide range of chromosome numbers. Lysák and Schubert (Chap. 9) explain that in many cases this originates via chromosome rearrangements. The authors outline in detail the mechanisms of chromosome rearrangements detectable by microscopic techniques and highlight those that have had an impact on the alteration of chromosome number and structure during evolution and thus may have played a role in speciation.

The compartmentalization of genomes into chromosomes has provided opportunities for the development of specialized chromosomes. One such example is the B chromosome (often called supernumerary chromosome), and Houben et al. (Chap. 10) describe its structure, DNA composition, and evolution. The authors explain peculiarities in the behaviour of Bs during mitosis and meiosis and list various drive mechanisms responsible for retaining Bs in the population.

Sex chromosomes are another classic example of specialized chromosomes, and Janoušek et al. (Chap. 11) review sex determination systems in various plant groups and, based on taxonomic distribution, argue that dioecy has originated independently many times during evolution. The authors introduce the genus *Silene* as an excellent system to study the evolution of sex chromosomes and present the first ever evidence of sex dimorphism in dioecious plants.

Bureš et al. (Chap. 12) describe holocentric chromosomes which differ from the more common monocentric chromosomes by the way in which spindle microtubules attach along the whole chromosome length through kinetochores that cover a substantial part of their poleward surfaces during mitosis. They review the occurrence of holocentric chromosomes in plants and describe their chromatin structure and behaviour during mitosis and meiosis and the evolutionary processes that have contributed to the diversity of holocentric karyotypes.

The remaining three chapters in this section analyse karyotype diversity in three different groups of plants. Weiss-Schneeweiss and Schneeweiss (Chap. 13) provide a comprehensive account of karyotype diversity and evolutionary trends in angiosperms. They discuss in detail how changes in chromosome number, including dysploidy and aneuploidy, as well as changes in chromosome morphology contribute to the karyotype diversity observed. They also outline various cytogenetic methods which can be used to characterize chromosomes in a karyotype and study their changes during evolution and speciation. Murray (Chap. 14) presents a survey of chromosome numbers and size variation in gymnosperms, the sister group to angiosperms, and describes the methods used to analyse karyotype diversity in this group of seed plants. After reviewing available data, the author concludes that in contrast to angiosperms, gymnosperms are characterized by much greater uniformity in chromosome number and karyotype. The third chapter in this block is by Barker (Chap. 15) who focuses on karyotype and genome evolution in pteridophytes (monilophytes and lycophytes). He draws attention to the high chromosome numbers typical of many ferns, particularly the homosporous species, which on average contain over three-fold more chromosomes than the average flowering plant. Interestingly, there is currently no conclusive answer as to why this should be so although it is expected that the availability of complete genome sequences will contribute to solving this long-standing mystery.

Generative Polyploidy

The three chapters in this section evaluate various features of generative polyploidy, which is widespread in land plants. Husband et al. (Chap. 16) examine patterns of polyploid occurrence, such as the variation among taxonomic groups at or above the species level, intraspecific variation, variation in mechanisms of formation, geographic and ecological patterns of polyploid incidence, and associations between ploidy and reproduction.

Thanks to the advances in DNA sequencing and genomics there is now evidence to suggest that most seed plants have undergone at least one episode of poylploidization. Thus, one cannot consider the evolution of land plants without understanding polyploidy. Fawcett et al. (Chap. 17) explain how the episodes of ancient polyploidization can be identified and dated and describe the immediate consequences of polyploidization to genes and genomes. They also discuss changes within polyploid genomes during evolution and the contribution of polyploidization to the evolutionary success of descendant lineages.

While the majority of studies on polyploidy have focused on angiosperms, Rensing et al. (Chap. 18) consider the importance of polyploidy in haploid-dominant land plants, the bryophytes. Here, polyploidy may play an even more essential evolutionary role than in other evolutionary lineages, rendering a bryophyte more robust against somatic mutations, while changes in chromosome number through polyploidy can lead to changes in the sexual system. The need for more genomic data and model species is paramount and the sequencing of the genome of the moss *Physcomitrella patens* together with the eagerly anticipated genome sequences from other moss species and the liverwort *Marchantia polymorpha* in the near future should shed further light on genome evolution and the role of polyploidy in these haploid-dominant land plants.

Genome Size Diversity and Consequences

The book closes with chapters that consider the whole genome in bulk. Leitch and Leitch (Chap. 19) take advantage of the recent increase in the number of species with genome size data and provide a comprehensive review on diversity of genome sizes across all groups of land plants. Evaluation of individual groups suggests that most plant genomes are rather small, probably due to strong selection pressure to limit genome size. Importantly, the chapter also considers how the diversity in genome size might have evolved. The last chapter of this volume by Greilhuber and Leitch (Chap. 20) examines the phenotypic correlates of variation in genome size, which include cell size and cell division rate. It also discusses the theories to explain the causality behind this variation observed, considers alternative views, and puts important studies into focus.

There is no doubt that it was an ambitious goal to cover the broad range of biological phenomena related to the structure, function, and evolution of plant genomes. However, we were motivated by the lack of a single resource, which is so needed in this era of rapid DNA sequence data generation. The chapters included in this volume deliver exciting facts from the history and life of plant genomes and present unanswered questions and hypotheses. We hope that the readers will find that the time spent with the book is both enjoyable and stimulating.

This volume would not exist without the contribution of the authors of individual chapters. Busy leaders in their areas of research, they spared precious time to share with us their knowledge and visions. We cannot be grateful enough for this and we appreciate their efforts and patience when responding to our requests for revisions. The only reward for them may be a response from the readers. So why not contact them? Sincere thanks go to the publisher, Springer-Verlag, Vienna and New York, who initiated and accompanied this project and made the publication of this volume possible. We appreciate the careful and professional management of the project.

RBG Kew, United Kingdom — Ilia. J. Leitch
Vienna, Austria — Johann Greilhuber
Olomouc, Czech Republic — Jaroslav Doležel
Ames, Iowa, USA — Jonathan F. Wendel
March, 2012

Contents

1 Angiosperm Phylogeny: A Framework for Studies of Genome Evolution 1
Pamela S. Soltis and Douglas E. Soltis

2 The Plant Nucleus at War and Peace: Genome Organization in the Interphase Nucleus 13
R. Neil Jones and Tim Langdon

3 The Organization of Genomic DNA in Mitotic Chromosomes: A Novel View 33
Hideaki Takata, Sachihiro Matsunaga, and Kazuhiro Maeshima

4 Structural Organization of the Plant Nucleus: Nuclear Envelope, Pore Complexes and Nucleoskeleton 45
Elena Kiseleva, Jindriska Fiserova, and Martin W. Goldberg

5 The Plant Nucleolus 65
Peter Shaw

6 Cell Cycle Modules in Plants for Entry into Proliferation and for Mitosis 77
Zoltán Magyar, Masaki Ito, Pavla Binarová, Binish Mohamed, and Laszlo Bogre

7 Endopolyploidy in Plants 99
Jolanta Maluszynska, Bozena Kolano, and Hanna Sas-Nowosielska

8 Meiosis: Recombination and the Control of Cell Division 121
Eric Jenczewski, Raphael Mercier, Nicolas Macaisne, and Christine Mézard

9 Mechanisms of Chromosome Rearrangements 137
Martin A. Lysák and Ingo Schubert

10 Biology and Evolution of B Chromosomes 149
Andreas Houben, Ali Mohammad Banaei-Moghaddam, and Sonja Klemme

11 Chromosomes and Sex Differentiation 167
Bohuslav Janoušek, Roman Hobza, and Boris Vyskot

12 Holocentric Chromosomes 187
Petr Bureš, František Zedek, and Michaela Marková

13 Karyotype Diversity and Evolutionary Trends in Angiosperms 209
Hanna Weiss-Schneeweiss and Gerald M. Schneeweiss

14 Karyotype Variation and Evolution in Gymnosperms 231
Brian G. Murray

15 Karyotype and Genome Evolution in Pteridophytes 245
Michael S. Barker

16 The Incidence of Polyploidy in Natural Plant Populations: Major Patterns and Evolutionary Processes 255
Brian C. Husband, Sarah J. Baldwin, and Jan Suda

17 Significance and Biological Consequences of Polyploidization in Land Plant Evolution 277
Jeffrey A. Fawcett, Yves Van de Peer, and Steven Maere

18 Evolutionary Importance of Generative Polyploidy for Genome Evolution of Haploid-Dominant Land Plants 295
Stefan A. Rensing, Anna K. Beike, and Daniel Lang

19 Genome Size Diversity and Evolution in Land Plants 307
Ilia J. Leitch and Andrew R. Leitch

20 Genome Size and the Phenotype 323
Johann Greilhuber and Ilia J. Leitch

Index 345

Contributors

Dr. Sarah J. Baldwin Department of Integrative Biology, University of Guelph, Guelph, ON, Canada

Michael S. Barker Department of Ecology & Evolutionary Biology, University of Arizona, Tucson, USA, msbarker@email.arizona.edu

Dr. Anna K. Beike Plant Biotechnology, Faculty of Biology, University of Freiburg, Freiburg, Germany

Dr. Pavla Binarová Institute of Microbiology, ASCR, Prague 4, Czech Republic

Prof. Laszlo Bogre Royal Holloway, University of London, Centre for Systems and Synthetic Biology, Egham, UK, l.bogre@rhul.ac.uk

Prof. RNDr. Petr Bureš Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic, bures@sci.muni.cz

Dr. Jeffrey A. Fawcett Graduate University for Advanced Studies, Hayama, Kanagawa, Japan

Dr. Jindriska Fiserova Department of Biological and Biomedical Sciences, Durham University, Durham, UK

Dr. Martin W. Goldberg Department of Biological and Biomedical Sciences, Durham University, Durham, UK

Prof. Johann Greilhuber Department of Systematic and Evolutionary Botany, Faculty of Life Sciences, University of Vienna, Vienna, Austria

Dr. Roman Hobza Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno, Czech Republic

Dr. Andreas Houben Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany, houben@ipk-gatersleben.de

Prof. Dr. Brian C. Husband Department of Integrative Biology, Science Complex, University of Guelph, Guelph, ON, Canada, bhusband@uoguelph.ca

Dr. Masaki Ito Graduate School of Bioagricultural Sciences, Nagoya University, Nagoya, Japan

Dr. Bohuslav Janoušek Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno, Czech Republic

Dr. Eric Jenczewski Institut Jean-Pierre Bourgin, Institut National de Recherche Agronomique, cedex, France

Prof. R. Neil Jones Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Wales, UK, rnj@aber.ac.uk

Dr. Elena Kiseleva Laboratory of Morphology and Function of Cell Structure, Institute of Cytology and Genetics, Novosibirsk, Russia, elka@bionet.nsc.ru

Dr. Sonja Klemme Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Chromosome Structure and Function Laboratory, Gatersleben, Germany

Dr. Bozena Kolano Department of Plant Anatomy and Cytology, University of Silesia, Katowice, Poland

Dr. Daniel Lang Plant Biotechnology, Faculty of Biology, University of Freiburg, Freiburg, Germany

Dr. Tim Langdon Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Wales, UK

Prof. Andrew R. Leitch School of Biological and Chemical Sciences, Queen Mary, University of London, London, UK

Dr. Ilia J. Leitch Jodrell Laboratory Royal Botanic Gardens, Kew, Richmond, Surrey, UK, I.Leitch@kew.org

Dr. Martin Lysák Department of Experimental Biology, Faculty of Science, Masaryk University, Brno, Czech Republic, lysak@sci.muni.cz

Dr. Nicolas Macaisne Institut Jean-Pierre Bourgin, Institut National de Recherche Agronomique, Versailles, cedex, France

Prof. Steven Maere Department of Plant Systems Biology, VIB, Ghent, Belgium, stmae@psb.vib-ugent.be

Dr. Kazuhiro Maeshima Laboratory for Biological Macromolecules, Structural Biology Center, National Institute of Genetics, Mishima, Shizuoka, Japan, kmaeshim@lab.nig.ac.jp

Dr. Zoltán Magyar Institute of Plant Biology, Biological Research Centre, Szeged, Hungary

Prof. Jolanta Maluszynska Department of Plant Anatomy and Cytology, University of Silesia, Katowice, Poland, jolanta.maluszynska@us.edu.pl

Dr. Michaela Marková Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic

Dr. Sachihiro Matsunaga Department of Biotechnology, Graduate School of Engineering, Osaka University, Suita, Osaka, Japan

Dr. Raphael Mercier Institut Jean-Pierre Bourgin, Institut National de Recherche Agronomique, Versailles, cedex, France

Prof. Christine Mézard Station de Génétique et d'Amélioration des Plantes, Versailles, France, Christine.Mezard@versailles.inra.fr

Dr. Binish Mohamed Royal Holloway, University of London, Centre for Systems and Synthetic Biology, Egham, UK

Dr. Ali Mohammad Banaei Moghaddam Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

Prof. Brian G. Murray School of Biological Sciences, University of Auckland, Auckland, New Zealand, b.murray@auckland.ac.nz

Dr. Stefan A. Rensing FRISYS, Faculty of Biology, University of Freiburg, Freiburg, Germany, stefan.rensing@biologie.uni-freiburg.de

Prof. Ingo Schubert Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

Dr. Hanna Sas-Nowosielska Department of Plant Anatomy and Cytology, University of Silesia, Katowice, Poland

Prof. Gerald M. Schneeweiss Department of Systematic and Evolutionary Botany, Faculty Center Botany, University of Vienna, Vienna, Austria

Dr. Peter Shaw Cell and Developmental Biology Department, John Innes Centre, Norwich, UK, peter.shaw@bbsrc.ac.uk

Prof. Douglas E. Soltis Department of Biology and the Genetics Institute, University of Florida, Gainesville, FL, USA, dsoltis@botany.ufl.edu

Prof. Pamela S. Soltis Laboratory of Molecular Systematics and Evolutionary Genetics, Florida Museum of Natural History, University of Florida, Gainesville, FL, USA, psoltis@flmnh.ufl.edu

Prof. Jan Suda Department of Botany, Faculty of Science, Charles University in Prague, Prague, Czech Republic

Dr. Hideaki Takata Biological Macromolecules Laboratory, Structural Biology Center, National Institute of Genetics, Mishima, Shizuoka, Japan

Prof. Yves Van de Peer Department of Plant Systems Biology, VIB, Ghent, Belgium, yves.vandepeer@psb.vib-ugent.be

Prof. Boris Vyskot Institute of Biophysics, Czech Academy of Sciences, Laboratory of Plant Developmental Genetics, Brno, Czech Republic, vyskot@ibp.cz

Prof. Hanna Weiss-Schneeweiss Department of Systematic and Evolutionary Botany, Faculty Center Botany, University of Vienna, Vienna, Austria, Hanna.Weiss@univie.ac.at

Dr. František Zedek Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic

Angiosperm Phylogeny: A Framework for Studies of Genome Evolution

1

Pamela S. Soltis and Douglas E. Soltis

Contents

1.1 **Introduction** 1

1.2 **Methods of Phylogenetic Analysis: A Primer** 2

1.3 **The Phylogeny of Embryophytes: An Abbreviated Overview** 4

1.4 **The Phylogeny of Angiosperms: An Overview** 5
1.4.1 Major Clades 5
1.4.2 Repeated Radiations 6
1.4.3 Unresolved Relationships 7
1.4.4 "Big Trees" 7

1.5 **Studies of Genome Evolution in Angiosperms** 7

1.6 **Future Prospects** 8
1.6.1 New Scope, New Tools 8
1.6.2 Improved Access to Data, Trees, and Tools 8

References 9

1.1 Introduction

The angiosperms—or flowering plants—comprise an estimated 260,000 (Takhtajan 1997)–400,000 (Raven in Jarvis 2007) extant species and occupy nearly all habitats on Earth except the coldest arctic and polar regions and the deepest oceans. Their diversification has occurred over a relatively short timespan, with the fossil record placing the earliest angiosperms in the early Cretaceous, approximately 132 million years ago. Molecular clock estimates suggest that the angiosperms are perhaps older, dating to the Jurassic (e.g., Sanderson et al. 2004; Bell et al. 2005; Bell et al. 2010) or even the Triassic (Magallon 2010; Smith et al. 2010).

Our understanding of the phylogeny of angiosperms has improved dramatically in recent years through large-scale collaborative analyses (e.g., Chase et al. 1993; Soltis et al. 1999; Soltis et al. 2000; Hilu et al. 2003; Soltis et al. 2011) and the application of molecular data, from single genes to entire plastid genomes (e.g., Jansen et al. 2007; Moore et al. 2007; Moore et al. 2010). Likewise, many clade-specific analyses have clarified relationships within some of the largest groups of angiosperms: e.g., *Monocotyledoneae* (monocots *sensu* Cantino et al. 2007; subsequent italicized names refer to phylogenetically defined clades in Cantino et al. 2007), Chase et al. (2006); Caryophyllales, Brockington et al. (2009); *Eudicotyledonea*e (eudicots), Moore et al. (2010); *Campanulideae* (campanulids), Tank and Donoghue (2010). In less than 20 years time, our view of angiosperm phylogeny has been transformed from a nebulous series of possible transitions to a well-supported and well-resolved tree of explicit sister-group relationships (summarized in Fig. 1.1). The stability of this tree is reflected in the modest changes to the classification of the Angiosperm Phylogeny Group (APG) over the past decade (1998, 2003, 2009; summarized by Stevens 2001 onward) and in the development of a phylogenetic nomenclature for angiosperms (Cantino et al. 2007).

P.S. Soltis (✉)
Florida Museum of Natural History and the Genetics Institute, University of Florida, Gainesville, FL 32611, USA
e-mail: psoltis@flmnh.ufl.edu

I.J. Leitch et al. (eds.), *Plant Genome Diversity Volume 2*,
DOI 10.1007/978-3-7091-1160-4_1,

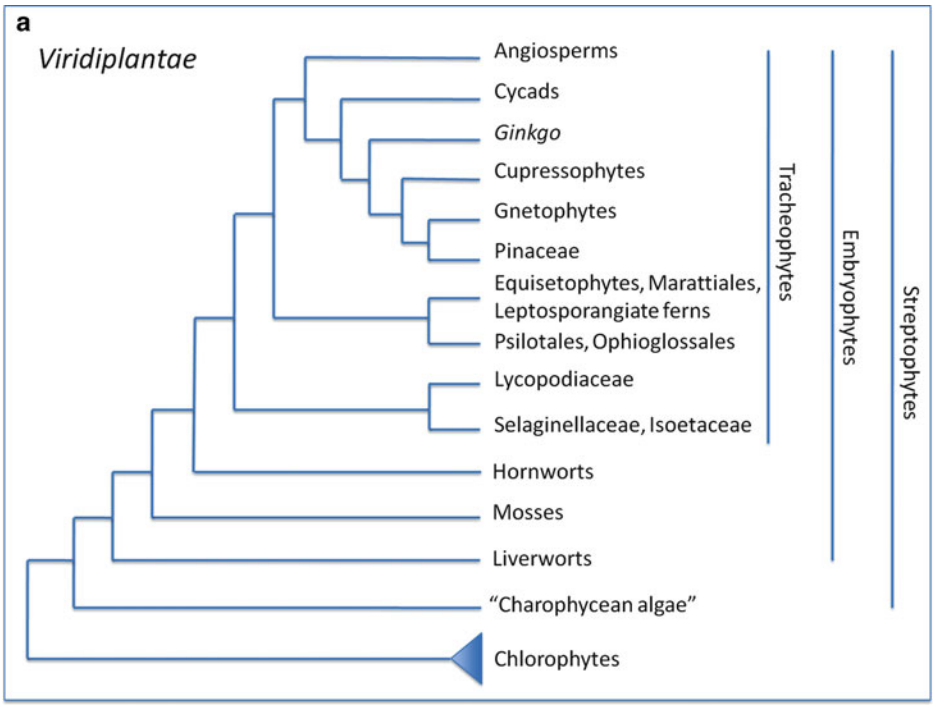

Fig. 1.1 (continued)

Phylogenetic trees of angiosperms have been used to address a range of evolutionary and ecological questions, such as the causes of diversification (Davies et al. 2004), the evolution of reproductive systems (e.g., Culley et al. 2002), the evolution of syncarpy and its role in pollination (Armbruster et al. 2002), and the relationship among phylogeny, biogeography, and biodiversity (Donoghue 2008). In addition, trees for which internal nodes have been dated (e.g., Wikstrom et al. 2001; Bell et al. 2005; Bell et al. 2010) have supplied a framework for many additional studies (Slingsby and Verboom 2006; Vamosi et al. 2006; Edwards et al. 2007; Webb et al. 2008).

Recent advances in angiosperm phylogenetics have also played a significant role in selecting taxa for genetic analysis and genome sequencing (e.g., Pryer et al. 2002; Soltis et al. 2008). For example, studies of gene family evolution have focused on representatives of basal angiosperm clades, basal eudicots, and selected monocots and eudicots (e.g., Kim et al. 2004; Kim et al. 2005; Zahn et al. 2005) to investigate patterns of gene duplication and loss. The results yield complex patterns of gene family dynamics—patterns that are not apparent through analysis of model systems alone. Likewise, genomic resources (e.g., BAC libraries) have been developed for a set of phylogenetically important plant species in anticipation of eventual genomic analysis and sequencing. Most recently, genome sequencing of *Amborella trichopoda*, the sister to all other extant angiosperms (e.g., Soltis et al. 1999; Soltis et al. 2000; Hilu et al. 2003; Leebens-Mack et al. 2005; Jansen et al. 2007), has been initiated, to provide an evolutionary reference for genome analysis within the angiosperms and across all green plants (Soltis et al. 2008; Chamala et al. 2011). The genome sequence of *Aquilegia* of Ranunculales, the sister to all other eudicots, will similarly provide an evolutionary reference for eudicots and a further point of comparison among the genomes of *Amborella*, monocots, and model eudicots.

Here we provide an overview of plant phylogeny, with an emphasis on angiosperms, based on the past two decades of research, to serve as the basis for investigating patterns of genome evolution. We give a summary, as well as many original citations, with an emphasis on those analyses that have deposited trees in public databases, such as TreeBASE, where they are available for download and analysis.

1.2 Methods of Phylogenetic Analysis: A Primer

The development of phylogenetic methods during the past decade has produced a perhaps baffling array of approaches, algorithms, and software. The state of the art a mere decade ago was maximum parsimony, with numerous options, e.g., TNT (Goloboff 1999), parsimony ratchet (Nixon 1999), to allow for analysis of perhaps several hundred taxa to a few thousand (Kallersjo et al. 1999) and one or a handful of genes. Concerns that sufficient tree space was searched were paramount, given the restrictions in memory and speed of most computers at the time. Maximum likelihood analyses were possible for only tens of taxa. In the early 2000s, major shifts occurred to model-based approaches as Bayesian methods (MrBayes, Huelsenbeck and Ronquist 2001; Huelsenbeck et al. 2001; and then BEAST, Drummond and Rambaut

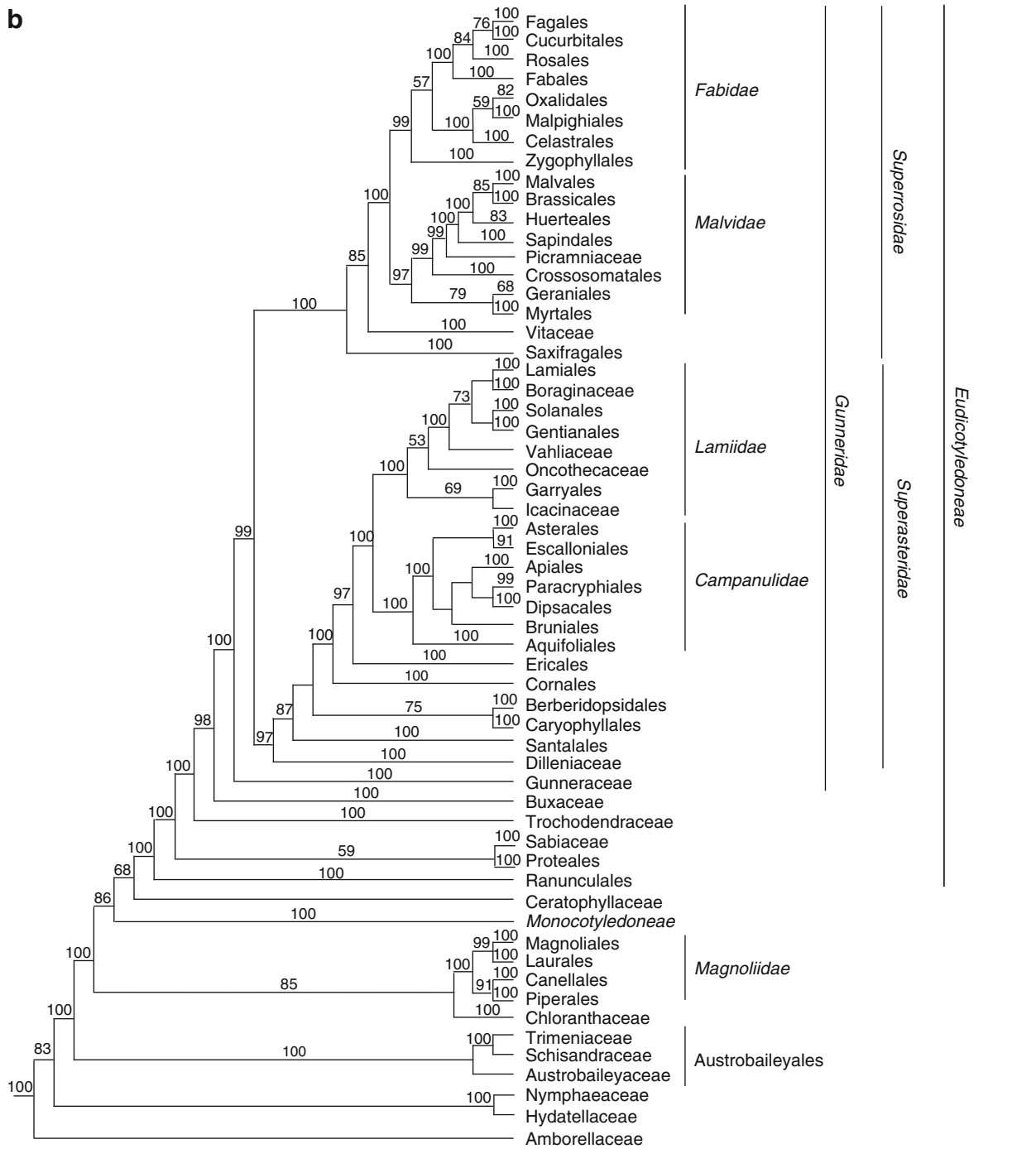

Fig. 1.1 Summary of phylogenetic relationships among major clades of green plants (*Viridiplantae*). (**a**) Overview, based on consensus of many studies. (**b**) Overview of angiosperm phylogeny, based on maximum likelihood analysis, with bootstrap values, redrawn from Soltis et al. (2011)

2007), along with maximum likelihood approaches using new algorithms (e.g., genetic algorithm, GARLI, Zwickl 2006; RAxML, Stamatakis 2006; Stamatakis et al. 2008), made it possible to reconstruct large trees (hundreds of taxa) with confidence scores (posterior probabilities or bootstrap values, respectively). Parallelization has helped to reduce run times dramatically for large problems, but has not been universally implemented to date (although dividing bootstrap analyses among an array of processors in a cluster is a form of parallel analysis that can considerably shorten run times).

Our assessment is that most projects today employ both parsimony and at least one model-based approach (typically RAxML or MrBayes).

Whereas many analyses of phylogeny reconstruction have embraced model-based methods, most analyses of character evolution continue to rely on parsimony, despite the implementation of both maximum likelihood and Bayesian methods for inferring ancestral states and mapping character variation. Although the reason for this bias is unclear, it may be that researchers are more comfortable applying statistical methods to tree selection than to character mapping, in which parsimony has an intuitive appeal. Although as in tree selection, likelihood, Bayesian, and parsimony methods typically produce similar patterns of character evolution, parsimony results may differ from likelihood and Bayesian reconstructions, particularly when branch lengths are short. We encourage expanded use of likelihood and/or Bayesian methods for character reconstructions, at least for comparison with parsimony results.

Most analyses of plant phylogeny to date have focused on plastid genes, with an emphasis at deep levels on *rbc*L, *atp*B, *ndh*F, and to some extent *mat*K. The former two have similar rates of evolution and are easily alignable, *ndh*F tends to evolve slightly more rapidly and is longer (although only part of the gene is sometimes used), and *mat*K has a higher rate of both nucleotide substitution and indels, leading to more difficult alignment. Of course, plastid genes provide only the evolutionary history of the plastid, and although this may not be a concern at the deepest levels of plant phylogeny, one must consider how well a plastid gene tree may reflect the organismal tree. To date, at deep levels, the only nuclear genes that have been used widely are the 18S and 26S ribosomal RNA genes. A number of MADS-box genes have been shown to track angiosperm phylogeny (Litt and Irish 2003; Kim et al. 2004; Kim et al. 2005; Zahn et al. 2005), but these genes have not yet been applied solely for the purpose of phylogeny reconstrution. However, a number of other nuclear genes (or their introns) have been used at more shallow levels: *LEAFY, APETALA3, PISTILLATA, ALCOHOL DEHYDROGENASE, GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE*, *CHALCONE SYNTHASE*, and *WAXY*, to name a few. All of this latter set of genes tend to have regions of 1,000 bp (plus or minus a few hundred) and are generally fairly easy to amplify with standard primers. However, the use of nuclear genes carries its own concerns, most notably issues of orthology, allelic diversity, and recombination, often requiring extensive cloning, sequencing of clones, and analyses of recombination prior to phylogenetic analysis. A set of mitochondrial genes (e.g., *mat*R, *atp*1, *nad*5, *rps*3) has also been applied to plant phylogeny. These genes tend to evolve more slowly than either plastid or nuclear genes used to date and can supply characters that are useful deep in plant phylogeny. However, mitochondrial-based trees have shown evidence of horizontal transfer of mitochondrial genes (e.g., Won and Renner 2003; Bergthorsson et al. 2004; Davis and Wurdack 2004) and should therefore be used in conjunction with other markers, particularly in groups that contain parasites.

1.3 The Phylogeny of Embryophytes: An Abbreviated Overview

A thorough summary of plant phylogeny is beyond the scope of this chapter; instead, we present a simple overview to set the stage for further discussion of genome evolution in plants and the phylogenetic placement of angiosperms, the focus of this chapter. Although important for understanding major patterns of plant evolution, especially in morphological and anatomical characters, the fossil record plays a less crucial role in understanding genome evolution, and we have therefore largely confined our discussion of plant phylogeny to extant taxa. However, the fossil record provides a requisite timeframe on our interpretation of phylogeny, and the phylogenetic placement of fossil groups may affect the topology of extant groups; we therefore introduce data from fossils as needed in this brief overview but we recognize that our treatment is incomplete.

The "green plants" (sometimes referred to as *Viridiplantae* or viridophytes) are a clade of at least half a million species with a fossil record that extends back nearly one billion years. They share a common cyanobacterial endosymbiotic event with red algae and glaucophytes and can be diagnosed by chlorophyll *b*, starch as the storage product for photosynthesis, and a stellate flagellar structure (e.g., Judd et al. 2008). A basal split in the green plants produced two clades, the chlorophytes (mostly marine "green algae") and streptophytes (which include freshwater "green algae" and embryophytes). *Mesostigma*, a freshwater "alga", has been identified as the sister to all other streptophytes. Subsequently branching lineages include the Klebsormidiales, the Zygnematales, and the Coleochaetales and Charales, the limits of which are not completely clear. *Chara* (and relatives) and Coleochaetales seem to be the sister group(s) of the embryophytes, although some recent studies place Zygnematales in this position (e.g., Timme et al. 2012). Embryophytes, or land plants, trace their history to at least the Ordovician and began to diversify extensively in the Silurian and Devonian. Morphological and anatomical synapomorphies of the embryophytes are a multicellular sporangium, thick-walled spores, multicellular gametangia, an embryo, and a cuticle.

Within the embryophytes, the phylogeny of the major clades is not fully resolved (Fig. 1.1a). For example, although traditionally recognized as a single taxonomic group, the bryophytes (consisting of mosses, liverworts,

and hornworts) are paraphyletic, and the branching order of these clades relative to the tracheophytes is not yet clear. All possible branching orders have been proposed. Most morphological and molecular data support liverworts as sister to all other embryophytes, and the major remaining disagreement is between the topology of (liverworts, (hornworts, (mosses + tracheophytes))), which is supported by the shared feature of a sporophyte apical meristem in mosses and tracheophytes, and (liverworts, (mosses, (hornworts + tracheophytes))), which is supported by the persistently green sporophyte in hornworts and tracheophytes (reviewed in Judd et al. 2008).

The tracheophytes (*Tracheophyta*) comprise two major clades, lycophytes (*Lycopodiophyta*) and euphyllophytes (*Euphyllophyta*). *Lycopodiophyta* comprises *Isoetes*, *Selaginella*, and Lycopodiaceae and forms a clade with some of the most prominent early vascular plants: *Cooksonia*, zosterophytes, and Lepidodendrales (Judd et al. 2008). Extant *Euphyllophyta* is united by a plastid genome inversion, multicellular sperm, overtopping, and terminal sporangia on lateral branches. The clade contains two major clades, *Monilophyta* and *Lignophyta*. The former is composed of Psilotales, Ophioglossales, Equisetales, Marattiales, and the leptosporangiate ferns (*Leptosporangiatae*). This assemblage of monilophytes was recognized by Kenrick and Crane (1997) on the basis of stem anatomy and was later supported by molecular data as well (Pryer et al. 2001; Pryer et al. 2004). The *Lignophyta* comprises several fossil lineages and the *Spermatophyta*, the seed plants.

Relationships among seed plants perhaps remain the most challenging in plant phylogeny. Extensive extinction within this clade has undoubtedly contributed to the difficulty of phylogeny reconstruction. The "gymnosperms" as typically recognized are paraphyletic and include several clades with extant members (cycads, *Ginkgo*, conifers, and gnetophytes) and several other groups that are only found in the fossil record (*Medullosa*, seed ferns, glossopterids, *Caytonia*, Bennettitales). However, the paraphyly of the "gymnosperms" is not apparent in molecular-based trees that typically (but not always) recover reciprocally monophyletic gymnosperms and angiosperms. When fossils are included in phylogenetic analyses of seed plants, disagreements exist with regard to both the placement of many of the non-flowering seed plants and in the sister group of the angiosperms, and there is little consensus on the overall phylogeny of all seed plants. It appears that glossopterids, *Caytonia*, and Bennettitales are more closely related to angiosperms than to other "gymnosperms" but beyond that, there is little resolution. When only extant seed plants are considered, disagreement still abounds. Most recent studies have found a topology in which extant gymnosperms and angiosperms are sister groups, with cycads either sister to all other extant gymnosperms or sister to *Ginkgo*, and with conifers and gnetophytes either sister to each other, or more often, gnetophytes nested within conifers or within Pinaceae. Despite extensive study using many sources of data and modes of analysis, the phylogenetic relationships among extant seed plants remain unresolved. For analyses that seek to examine patterns of evolution among angiosperms, this frustrating result precludes identification of the appropriate outgroup for comparative studies.

1.4 The Phylogeny of Angiosperms: An Overview

1.4.1 Major Clades

Angiosperm phylogeny has been studied extensively in recent decades, from the perspective of deep-level branching patterns to clades of closely related species. Ultimately and ideally, these results will be linked, either through supertree methods that combine published trees via shared taxa or through new supermatrix analyses that combine all data into a single matrix for analysis (see below). Here we will provide only an overview of the major clades and their interrelationships, followed by further discussion on some of the emergent patterns from analyses conducted to date.

Nearly all molecular-based analyses of the past decade have identified *Amborella* as the sister to all other extant angiosperms, most often alone or occasionally with Nymphaeales (Fig. 1.1b). All analyses are consistent in then placing Austrobaileyales (comprising *Austobaileya*, *Trimenia*, *Illicium*, and Schisandraceae) as the sister group to all other extant angiosperms. This large remainder, the *Mesangiospermae*, comprises *Magnoliidae* + Chloranthaceae as sister to *Monocotyledoneae* + *Eudicotyledoneae* + *Ceratophyllum* (Moore et al. 2007). Although long recognized as an ancient group, the placement of Chloranthaceae has been elusive, but recent analyses place them as sister to *Magnoliidae*. The relationship among magnoliids, monocots, and eudicots has been very difficult to disentangle, but plastid genome sequences support the sister-group relationship of monocots and eudicots (+ *Ceratophyllum*).

Relationships among major clades of monocots are now clear, but they do not follow traditional taxonomic circumscriptions. Acorales are sister to all other extant monocots, and a grade that includes Alismatales, followed by Petrosaviaceae, subtends a clade comprising the majority of monocot species diversity. Pandanales + Dioscoreales are sister to a clade of (Liliales, (Asparagales + *Commelinidae*)). One of the most substantial reorganizations of monocot classification is based on new understanding of relationships of the former Liliaceae. Although dismantling of this large family was proposed many years ago, the placements of its components have not always been clear. Progress has

been substantial, but questions remain. Likewise, relationships within Asparagales have been difficult to resolve, and although recent analyses have done much to resolve phylogeny, few morphological characters have been idenitifed to diagnose the component clades. The *Commelinidae* are diagnosed by starchy pollen, UV-fluorescent ferulic and coumaric acids in the cell walls, and *Strelitzia*-type epicuticular wax. The clade is large, comprising over 25,000 species, with diversity spanning grasses to palms. Component clades are Commelinales, Zingiberales, Arecales, and Poales (+ Dasypogonaceae).

Within eudicots, a basal grade consisting of Ranunculales, Proteales, Sabiales, Trochodendrales, and Buxales subtends the "core eudicots", or *Gunneridae*. Gunnerales are sister to the remaining *Gunneridae*, the *Pentapetalae*, which fall into two major clades (Moore et al. 2010; Soltis et al. 2011): *Superrosidae* and *Superasteridae* (*sensu* Soltis et al. 2011). *Superrosidae* comprises Saxifragales, Vitaceae, and *Rosidae*, whereas *Superasteridae* contains Santalales, Berberidopsidales, Caryophyllales, and *Asteridae*. Dilleniaceae, which has been associated with Caryophyllales in some previous analyses and unplaced in many others, remains unplaced, with alternative placements in *Superrosidae* and *Superasteridae* (see below). The positions of most major clades of core eudicots (*Gunneridae*) remain unchanged relative to earlier studies (e.g., Soltis et al. 1999; Soltis et al. 2000; Soltis et al. 2005), although additional resolution has been obtained in both the *Rosidae* and *Asteridae* (see Soltis et al. 2011, and references therein).

Phylogenetic analyses of the angiosperms at these deep levels have resulted in new classifications, such as the Angiosperm Phylogeny Group's (APG) (1998, 2003, 2009) classifications at the familial and ordinal levels, with rank-free names assigned to clades corresponding to groups larger than recognized orders. An alternative rank-free classification has also emerged (Cantino et al. 2007), with phylogenetic definitions provided for many of the clades that correspond to those recognized at the ordinal level and above in the APG system.

1.4.2 Repeated Radiations

A prominent pattern apparent in the angiosperm phylogenetic trees is a series of polytomies interspersed by regions of well-resolved relationships (see Soltis et al. 2005; Soltis et al. 2008; Wang et al. 2009; Soltis et al. 2010). Although polytomies may be due to insufficient data or taxon sampling, they may also represent real radiations. Within the angiosperms, several apparent radiations have persisted through the addition of new data and more taxa, suggesting that in fact these radiations are real.

The origin of the angiosperms themselves has often been considered a rapid radiation, based on the fossil record and Darwin's words themselves: *"The rapid rise and early diversification of angiosperms is an abominable mystery…"* (Darwin 1903). However, phylogenetic reconstructions suggest instead that the angiosperms radiated, not immediately upon their origin, but a few nodes subsequent to the common ancestor of all extant angiosperms (Mathews and Donoghue 1999; Soltis et al. 1999; Soltis et al. 2005). This radiation corresponds to the diversification of the *Mesangiospermae* (*sensu* Cantino et al. 2007), the clade comprising magnoliids, Chloranthaceae, monocots, Ceratophyllaceae, and eudicots—in other words, all angiosperms except *Amborella*, Nymphaeales, and Austrobaileyales (see Moore et al. 2007). Subsequent radiations appear to follow the origin and early diversification of many large clades of angiosperms: for example, within the eudicots (*Eudicotyledoneae*), within the core eudicots (*Gunneridae*), within *Rosidae* (and within the fabid and malvid clades, *Fabidae* and *Malvidae*, of *Rosidae*), within the *Asteridae*, and within clades of monocots, to name a few.

Attempts to find causes for these apparent radiations have met with mixed success. One of the most extensive analyses of possible factors associated with radiations addressed both the early radiation of the angiosperms themselves and subsequent radiations (Davies et al. 2004). Davies et al. (2004) tested a range of traits reflecting prominent hypotheses for the "success" of the angiosperms on rates of diversification and found no significant association at any level of the tree. Despite the lack of significance of specific features, radiations within the angiosperms may be explained by biotic factors and cospeciation with other clades. For example, modern ferns diversified alongside angiosperms, suggesting either that angiosperms provided new habitats for ferns or that the same causal factors allowed diversification of both clades (Schneider et al. 2004). Moreover, within the angiosperms, the radiation of the rosids is associated with radiations in several other clades, such as ants, amphibians, and even primates (see Wang et al. 2009, for review). Other possible biotic interactions, such as those with mycorrhizal fungi, may also have contributed to radiations of angiosperms.

Recent observations of gene duplications at or near nodes associated with radiations are suggestive of a causal role of genetic or genomic factors in these radiations themselves. For example, coincident gene duplications in multiple subfamilies of the MADS-box gene family prior to the origin of the angiosperms raised hypotheses about the role that these duplications in genes important in the specification of floral organ identity and other features of the flower may have played in the early evolution of angiosperms (Buzgo et al. 2005; De Bodt et al. 2005; Zahn et al. 2005). Likewise, similar patterns of duplication appear to be associated with

the early evolution of the eudicots, again suggesting a causal role in the floral changes that occurred at that point in angiosperm phylogeny and a further role in diversification (for reviews see Soltis et al. 2006; Soltis et al. 2009b). Finally, duplications in the *CYCLOIDEA* gene family suggest possible roles in both floral and species diversification in asterids (Howarth and Donoghue 2006).

Coincident gene duplications at specific nodes are suggestive of whole-genome duplications (WGD; see Buzgo et al. 2005; De Bodt et al. 2005; Zahn et al. 2005; reviewed in Soltis et al. 2009a; Soltis et al. 2009b), and it may be that WGD rather than duplications of specific floral genes triggered radiation. Whole-genome duplication (polyploidy) has, in fact, been suggested as the impetus for angiosperm success following the K-T boundary (Fawcett et al. 2009). Genomic data have revealed unsuspected episodes of WGD throughout green plants (see below; Blanc and Wolfe 2004; Cui et al. 2006; Soltis et al. 2009a; Jiao et al. 2011 for review). In several instances, the clade marked by WGD is more species-rich than the sister clade that lacks the duplication; however, genomic data are lacking for a sufficient number of species to allow for thorough statistical analyses of heterogeneity of diversification rates (Soltis et al. 2009a). Additional data for more species, so that WGD events can be plotted more accurately on a phylogenetic tree, are needed.

1.4.3 Unresolved Relationships

Given the recent progress in angiosperm phylogenetics, few major issues of deep-level relationships remain, although problems abound within clades recognized as "orders" and "families" *sensu* APG. Among deep-level problems, one of the most perplexing placements is that of Dilleniaceae, which occupies different positions depending on the data set and analysis, from sister to *Superrosidae* (*sensu* Soltis et al. 2011), to sister to *Superasteridae* (*sensu* Soltis et al. 2011), to sister to *Superasteridae* + *Superrosidae* (see Soltis et al. 2011, for discussion). Other prominent areas requiring further analysis include: (1) the branching order among basal eudicots; (2) relationships among major clades of rosids; (3) relationships within Malpighiales; (4) Lamiales. Most of these regions can most likely be resolved with additional taxa and DNA sequence data, although problem areas such as Malpighiales have recently received substantial attention, with at least some progress (Wurdack and Davis 2009).

1.4.4 "Big Trees"

Until recently, most tree reconstruction algorithms could not handle data sets of 1,000 or more terminals. However, recent modifications have yielded trees with many thousands of species (e.g., Goloboff et al. 2009; Smith et al. 2009; Smith et al. 2011). The most recent tree generated by Smith et al. (2011)—with 55,000 taxa!—used a newly adapted version of RAxML, demonstrating the ability to use model-based approaches for "big tree" reconstruction. The concern with such large trees, however, is their accuracy: with so many terminals, the thoroughness of the searches is reduced, raising the question of the accuracy of the results. The overall structure of the Smith et al. (2011) 55,000-taxon tree is quite similar to trees based on far fewer taxa (such as the 640-taxon tree of Soltis et al. 2011), suggesting that for many purposes, this very large tree will be very useful. However, without further diagnostics on the performance of tree reconstruction at this large scale, many close relationships may require cautious acceptance. And this may be an issue for studies aimed at reconstructing the evolution of genes or specific genomic traits at a fine scale. Nevertheless, breakthroughs in tree reconstruction will undoubtedly lead to new and exciting opportunities for learning about the evolution of plant genomes.

An alternative to the supermatrix approach described above is the construction of a supertree from smaller trees with overlapping taxa (see Sanderson et al. 1998; Davies et al. 2004). Supertree methods take advantage of vast amounts of data collected and analysed in the past to produce a summary of phylogenetic inferences contained in published trees. Although not without their own problems, supertrees offer a solution to generating large phylogenies. Recent methods that combine elements of supermatrix and supertree approaches show particular promise.

1.5 Studies of Genome Evolution in Angiosperms

Hypotheses on patterns of chromosomal evolution in the angiosperms abound. Classical perspectives were based on integrated inferences on ancient and recent polyploidy, processes of chromosomal fission and fusion, and relative "advancement" of a taxonomic group. More recently, it has been possible to test hypotheses of increases and decreases in chromosome number and genome size by mapping these characteristics across an explicit phylogenetic tree. Longstanding hypotheses of chromosomal evolution have suggested that the ancestral chromosome number for angiosperms ranged from $x = 6$–9, with $x = 7$ a commonly proposed base number (e.g., Stebbins 1950; Ehrendorfer et al. 1968; Stebbins 1971; Raven 1975; Grant 1981). Chromosome numbers in angiosperms vary dramatically, from $2n = 4$ (e.g., *Haplopappus gracilis*, Asteraceae) to $2n =$ c. 640 (*Sedum suaveolens*, Crassulaceae), a 160-fold difference (Uhl 1978). However, it is clear that genome size varies independently of chromosome number, with genome size

ranging from 1C = 0.065 pg to 1C = 152.23 pg, a c. 2,400-fold difference (Greilhuber et al. 2006; Pellicer et al. 2010; Leitch and Leitch, 2013 this volume). Previous reconstructions of genome size across angiosperms found that the ancestral genome was "very small", with 1C $\leq$ 1.4 pg (Leitch et al. 1998; Soltis et al. 2003), with multiple increases and decreases in genome size from this ancestral condition.

Another clear attribute of angiosperm genomes is polyploidy, and events of whole-genome duplication have also been mapped across an angiosperm tree (Soltis et al. 2009a). Whereas polyploidy has long been recognized as an important speciation mechanism in angiosperms, events of genome duplication were long considered to be confined to the tips of the tree, with a few putative cases of ancient polyploidy, e.g., Magnoliaceae, Lauraceae, Salicaceae (Stebbins 1950; Stebbins 1971). Remarkably, genome sequencing studies have revealed multiple rounds of genome duplication throughout the evolutionary history of angiosperms. The surprising finding that the very small genome of *Arabidopsis thaliana* has undergone multiple rounds of duplication set the stage for investigations of other unsuspected events of genome duplication, and all other angiosperms sequenced to date likewise exhibit signatures of ancient duplication (see Soltis et al. 2009a; Fawcett et al. 2013, this volume). As more genome sequences are generated, it will be possible to map these duplication events more clearly onto a phylogenetic tree. From there, hypotheses of possible causal effects of genome duplication may be tested.

1.6 Future Prospects

1.6.1 New Scope, New Tools

Despite the progress in reconstructing the major relationships within the angiosperms, much remains to be done to produce a comprehensive phylogeny. One approach to increasing taxon sampling is to include all species that are represented in GenBank, regardless of gene. Methods such as those explored by Driskell et al. (2004) have proven that even very sparse matrices can yield reasonable phylogenetic trees. A modification of this approach that includes only species for which any of a small, specified set of genes has been sequenced (Smith et al. 2009; Smith et al. 2011) has resulted in a 55,000-taxon tree, generated using maximum likelihood (via RAxML, Stamatakis 2006), that has been used to address the evolution of various traits (Smith et al. 2011). The ability to generate trees of this size is a recent breakthrough and sets the stage for future large-scale analyses. Whereas until recently, the limitation in phylogenetics was computational power, the bottleneck very soon will be a lack of data. Tools being developed by the iPlant Collaborative (iplantcollaborative.org) will soon be available for large-scale tree reconstruction and post-tree analyses.

An alternative approach to using data fortuitously available from GenBank is the deliberate generation of new data for taxa not represented in GenBank. For example, only a fraction of the estimated 15,000 genera of angiosperms are included in GenBank. Although most genera are likely not monophyletic (Judd et al. 2008), generic-level classification reflects a loose assessment of diversity and therefore a framework for ongoing phylogenetic analysis. However, substantial specimen collecting, including samples for DNA analysis, is needed to conduct such a study. Thus, given new developments in phylogenetic software, perhaps the major limitation to a comprehensive angiosperm phylogenetic tree is the lack of material for molecular analysis.

1.6.2 Improved Access to Data, Trees, and Tools

Until now, most phylogenetic reconstructions of morphological, physiological, ecological, or other characters have involved, initially, sharing of trees and data sets by systematists and, more recently, downloading published trees and data from public databases such as TreeBASE (treebase.org) and Dryad (datadryad.org). These analyses have been necessarily limited to those taxa included in prior phylogenetic trees, without representation of those taxa that might be of greater interest from the perspective of morphology or other traits. A solution is to reconstruct a new tree that includes such taxa, but large phylogenetic analyses may not be feasible for those interested in reconstructing patterns of character evolution. New approaches, such as an automatically generated tree with each new GenBank release (every 2 months) and the availability of data matrices and computational resources for tree estimation, are being developed by the iPlant Collaborative. Implementation of these tools will facilitate customized phylogenetic analyses and lead to "democratization" of angiosperm phylogenetics. Continued investment in phylogenetic cyberinfrastructure to address such issues as reticulation, horizontal gene transfer, and patterns of character evolution will lead to further advancements in our understanding of angiosperm evolution.

Acknowledgments This work was supported in part by the US National Science Foundation (grants EF-0431266 and PGR-0638595) and the NSF-funded iPlant Collaborative.

References

Angiosperm Phylogeny Group (1998) An ordinal classification for the families of flowering plants. Ann Mo Bot Gard 85:531–553

Angiosperm Phylogeny Group (2003) An update of the angiosperm phylogeny group classification for the orders and families of flowering plants. Bot J Linn Soc 141:399–436

APG, III [The Angiosperm Phylogeny Group] (2009) An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG III. Bot J Linn Soc 161:105–121

Armbruster WS, Debevec EM, Willson MF (2002) Evolution of syncarpy in angiosperms: theoretical and phylogenetic analyses of the effects of carpel fusion on offspring quantity and quality. J Evol Biol 15:657–672

Bell CD, Soltis DE, Soltis PS (2005) The age of the angiosperms: a molecular time-scale without a clock. Evolution 59:1245–1258

Bell CD, Soltis DE, Soltis PS (2010) The age and diversification of the angiosperms re-revisited. Am J Bot 97:1296–1313

Bergthorsson U, Richardson AO, Young GJ, Goertzen LR, Palmer JD (2004) Massive horizontal transfer of mitochondrial genes from diverse land plant donors to the basal angiosperm *Amborella*. Proc Natl Acad Sci USA 101:17747–17752

Blanc G, Wolfe KH (2004) Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell 16:1667–1678

Brockington SF, Alexandre R, Ramdial J, Moore MJ, Crawley S, Dhingra A, Hilu K, Soltis DE, Soltis PS (2009) Phylogeny of Caryophyllales and patterns of floral evolution. Int J Plant Sci 170:627–643

Buzgo M, Soltis PS, Kim S, Soltis DE (2005) The making of the flower. Biologist 52:149–154

Cantino P, Doyle J, Graham S, Judd W, Olmstead R, Soltis D, Soltis P, Donoghue M (2007) Towards a phylogenetic nomenclature of Tracheophyta. Taxon 56:822–846

Chamala S, Walts B, Albert V, dePamphilis C, Der J, Estill J, Leebens-Mack J, Lee S, Ma H, Rounsley S, Schuster S, Soltis D, Soltis P, Tomsho L, Wessler S, Wing R, Yu Y, Barbazuk B (2011) The *Amborella* genome project: generating a reference sequence for angiosperm evolutionary analysis. PAG abstract

Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, Duvall MR, Price RA, Hills HG, Qiu Y-L et al (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. Ann Mo Bot Gard 80:528–580

Chase MW, Fay MF, Devey D, Maurin O, Rønsted N, Davies J, Pillon Y, Petersen G, Seberg O, Tamura MN, Asmussen CB, Hilu K, Borsch T, Davis JI, Stevenson DW, Pires JC, Givnish TJ, Sytsma KJ, McPherson MA, Graham SW, Rai HS (2006) Multigene analyses of monocot relationships: a summary. In: Columbus JT, Friar EA, Porter JM, Prince LM, Simpson MG (eds) Monocots: comparative biology and evolution (excluding Poales), vol 22. Aliso, pp 63–75

Cui L, Wall PK, Leebens-Mack J, Lindsay BG, Soltis D, Doyle JJ, Soltis P, Carlson J, Arumuganathan A, Barakat A, Albert V, Ma H, dePamphilis CW (2006) Widespread genome duplications throughout the history of flowering plants. Genome Res 16:738–749

Culley TM, Weller SG, Sakai AK (2002) The evolution of wind pollination in angiosperms. Trends Ecol Evol 17:361–369

Darwin C (1903) Letter to J. D. Hooker. In: Darwin F, Seward AC (eds) More letters of Charles Darwin, vol 2. John Murray, London

Davies TJ, Barraclough TG, Chase MW, Soltis PS, Soltis DE, Savolainen V (2004) Darwin's abominable mystery: insights from a supertree of the angiosperms. Proc Natl Acad Sci USA 101:1904–1909

Davis CC, Wurdack KJ (2004) Host-to-parasite gene transfer in flowering plants: phylogenetic evidence from Malpighiales. Science 305:676–678

De Bodt S, Maere S, Van de Peer Y (2005) Genome duplication and the origin of angiosperms. Trends Ecol Evol 20:591–597

Donoghue MJ (2008) A phylogenetic perspective on the distribution of plant diversity. Proc Natl Acad Sci USA 105(suppl 1):11549–11555

Driskell AC, Ané C, Burleigh JG, McMahon MM, O'Meara BC, Sanderson MJ (2004) Phylogenetic utility of large sequence databases for building the tree of life. Science 306:1172–1174

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7:214

Edwards EJ, Stills CJ, Donoghue MJ (2007) The relevance of phylogeny to the studies of global change. Trends Ecol Evol 22:243–249

Ehrendorfer F, Krendl F, Habeler E, Sauer W (1968) Chromosome numbers and evolution in primitive angiosperms. Taxon 17:337–468

Fawcett JA, Maere S, Van de Peer Y (2009) Plants with double genomes might have had a better chance to survive the cretaceous-tertiary extinction event. Proc Natl Acad Sci USA 106:5737–5742

Fawcett JA, Van de Peer Y, Maere S (2013) Significance and biological consequences of polyploidization in land plants. In: Leitch IJ, Greilhuber J, Doležel J, Wendel JF (eds) Plant genome diversity, vol 2, Physical structure, behaviour and evolution of plant genomes. Springer-Verlag, Wien, pp 277–293

Goloboff PA (1999) Analyzing large data sets in reasonable times: solutions for composite optima. Cladistics 15:415–428

Goloboff PA, Catalano SA, Mirande JM, Szumik CA, Arias JS, Kallersjo M, Farris JS (2009) Phylogenetic analysis of 73060 taxa corroborates major eukaryotic groups. Cladistics 25:211–230

Grant V (1981) Plant speciation. Columbia University Press, New York

Greilhuber J, Borsch T, Müller K, Worberg A, Porembski S, Barthlott W (2006) Smallest angiosperm genomes found in Lentibulariaceae with chromosomes of bacterial size. Plant Biol 8:770–777

Hilu KW, Borsch T, Müller K, Soltis DE, Soltis PS, Savolainen V, Chase MW, Powell MP, Alice LA, Evens R, Sauquet H, Neinhuis C, Slotta TAB, Rohwer JG, Campbell CS, Chatrou LW (2003) Angiosperm phylogeny based on *mat*K sequence information. Am J Bot 90:1758–1776

Howarth DG, Donoghue MJ (2006) Phylogenetic analyses of the "ECE" (CYC/TB1) clade reveal duplications that predate the core eudicots. Proc Natl Acad Sci USA 103:9101–9106

Huelsenbeck JP, Ronquist F (2001) MrBayes: Bayesian inference of phylogeny. Bioinformatics 17:754–755

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback J (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294:2310–2314

Jansen RK, Cai Z, Raubeson LA, Daniell H, dePamphilis CW, Leebens-Mack J, Müller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee S-B, Peery R, McNeal J, Kuehl JV, Boore JL (2007) Analysis of 81 genes from 64 chloroplast genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci USA 104:19369–19374

Jarvis C (2007) Order out of chaos: Linnaean plant names and their types. Linnean Society of London and the Natural History Museum, London

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, dePamphilis CW (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473:97–100

Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ (2008) Plant systematics—a phylogenetic approach, 3rd edn. Sinauer, Sunderland

Kallersjo M, Albert V, Farris J (1999) Homoplasy increases phylogenetic structure. Cladistics 15:91–93
Kenrick P, Crane PR (1997) The origin and early diversification of land plants. Smithsonian Institution Press, Washington
Kim S, Soltis DE, Albert V, Yoo MJ, Farris JS, Soltis PS, Soltis DE (2004) Phylogeny and diversification of B-function MADS-box genes in angiosperms: evolutionary and functional implications of a 260-million-year-old duplication. Am J Bot 91:2102–2118
Kim S, Koh J, Yoo M-J, Kong H, Hu Y, Ma H, Soltis PS, Soltis DE (2005) Expression of floral MADS-box genes in basal angiosperms: implications for the evolution of floral regulators. Plant J 43:724–744
Leebens-Mack J, Raubeson LA, Cui L, Kuehl JV, Fourcade MH, Chumley TW, Boore JL, Jansen RK, dePamphilis CW (2005) Identifying the basal angiosperm node in chloroplast genome phylogenies: sampling one's way out of the Felsenstein zone. Mol Biol Evol 22:1948–1963
Leitch IJ, Chase MW, Bennett MD (1998) Phylogenetic analysis of DNA C-values provides evidence for a small ancestral genome size in flowering plants. Ann Bot 82:85–94
Leitch IJ, Leitch AR (2013) Genome size diversity and evolution in land plants. In: Leitch IJ, Greilhuber J, Doležel J, Wendel JF (eds) Plant genome diversity, vol 2, Physical structure, behaviour and evolution of plant genomes. Springer-Verlag, Wien, pp 307–322
Litt A, Irish VF (2003) Duplication and diversification in the *APETALA1/FRUITFULL* floral homeotic gene lineage: implications for the evolution of floral development. Genetics 165:821–833
Magallon S (2010) Using fossils to break long branches in molecular dating: a comparison of relaxed clocks applied to the origin of angiosperms. Syst Biol 59:384–389
Mathews S, Donoghue M (1999) The root of angiosperm phylogeny inferred from duplicate phytochrome genes. Science 286:947–950
Moore MJ, Bell CD, Soltis PS, Soltis DE (2007) Using plastid genomic-scale data to resolve enigmatic relationships among basal angiosperms. Proc Natl Acad Sci USA 104:19363–19368
Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE (2010) Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proc Natl Acad Sci USA 107:4623–4628
Nixon KC (1999) The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics 15:407–414
Pellicer J, Fay MF, Leitch IJ (2010) The largest eukaryotic genome of them all? Bot J Linn Soc 164:10–15
Pryer KM, Schneider H, Smith AR, Cranfill R, Wolf PG, Hunt JS, Sipes SD (2001) Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature 409:618–622
Pryer KM, Schneider H, Zimmer EA, Banks J (2002) Deciding among green plants for whole genome studies. Trends Plant Sci 7:550–554
Pryer KM, Schuettpelz ME, Wolf PG, Schneider H, Smith AR, Cranfill R (2004) Phylogeny and evolution of ferns (monilophytes) with a focus on the early leptosporangiate divergences. Am J Bot 91:1582–1598
Raven PH (1975) The bases of angiosperm phylogeny: cytology. Ann Mo Bot Gard 62:724–764
Sanderson MJ, Purvis A, Henze C (1998) Phylogenetic supertrees: assembling the trees of life. Trends Ecol Evol 13:105–109
Sanderson MJ, Thorne JL, Wikstrom N, Bremer K (2004) Molecular evidence on plant divergence times. Am J Bot 91:1656–1665
Schneider H, Schuettpelz E, Pryer KM, Cranfill R, Magallon S, Lupia R (2004) Ferns diversified in the shadow of angiosperms. Nature 428:553–557
Slingsby JA, Verboom GA (2006) Phylogenetic relatedness limits co-occurrence at fine spatial scales: evidence from the schoenoid sedges (Cyperaceae: Schoeneae) of the Cape Floristic Region, South Africa. Am Nat 168:14–27
Smith SA, Beaulieu JM, Donoghue MJ (2009) Mega-phylogeny approach for comparative biology: an alternative to supertree and supermatrix approaches. BMC Evol Biol 9:37
Smith SA, Beaulieu JM, Donoghue MJ (2010) An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. Proc Natl Acad Sci USA 107:5897–5902
Smith SA, Beaulieu JM, Stamatakis A, Donoghue MJ (2011) Understanding angiosperm diversification using small and large phylogenetic trees. Am J Bot 98:404–414
Soltis PS, Soltis DE, Chase MW (1999) Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402:402–404
Soltis DE, Soltis PS, Chase MW, Mort M, Albach D, Zanis M, Savolainen V, Hahn W, Hoot S, Fay M, Axtell M, Swensen S, Nixon K, Farris J (2000) Angiosperm phylogeny inferred from a combined data set of 18 S rDNA, rbcL and atpB sequences. Bot J Linn Soc 133:381–461
Soltis DE, Soltis PS, Bennett MD, Leitch IJ (2003) Evolution of genome size in the angiosperms. Am J Bot 90:1596–1603
Soltis DE, Soltis PS, Endress PK, Chase MW (2005) Phylogeny and evolution of the angiosperms. Sinauer, Sunderland
Soltis PS, Soltis DE, Kim S, Chanderbali A, Buzgo M (2006) Expression of floral regulators in basal angiosperms and the origin and evolution of the ABC model. Adv Bot Res 44:483–506
Soltis DE, Bell CD, Kim S, Soltis PS (2008) Origin and early evolution of angiosperms. Ann NY Acad Sci 1133:3–25
Soltis DE, Albert VA, Leebens-Mack J, Bell CD, Paterson A, Zheng C, Sankoff D, Wall PK, Soltis PS (2009a) Polyploidy and angiosperm diversification. Am J Bot 96:336–348
Soltis PS, Brockington SF, Yoo MJ, Piedrahita A, Latvis M, Moore MJ, Chanderbali AS, Soltis DE (2009b) Floral variation and floral genetics in basal angiosperms. Am J Bot 96:110–128
Soltis PS, Burleigh JG, Chanderbali AS, Yoo M-J, Soltis DE (2010) Gene and genome duplication in plants. In: Dittmar K, Liberles DA (eds) Evolution after genome duplication. Wiley-VCH, Weinheim, pp 369–398
Soltis DE, Smith SA, Cellinese N, Wurdack KJ, Tank DC, Brockington SF, Refulio-Rodriguez NF, Walker JB, Moore MJ, Carlsward BS, Bell CD, Latvis M, Crawley S, Black C, Diouf D, Xi Z, Rushworth CA, Gitzendanner MA, Sytsma KJ, Qiu Y-L, Hilu KW, Davis CC, Sanderson MJ, Beaman RS, Olmstead RG, Judd WS, Donoghue MJ, Soltis PS (2011) Angiosperm phylogeny: 17 genes, 640 taxa. Am J Bot 98:704–730
Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690
Stamatakis A, Hoover AP, Rougemont J (2008) A fast bootstrapping algorithm for the RAxML web-servers. Syst Biol 57:758–771
Stebbins GL (1950) Variation and evolution in plants. Columbia University Press, New York
Stebbins GL (1971) Chromosomal evolution in higher plants. Addison-Wesley, London
Stevens PF (2001 onward) Angiosperm phylogeny website. http://www.mobot.org/MOBOT/research/APweb/
Takhtajan A (1997) Diversity and classification of flowering plants. Columbia University Press, New York
Tank DC, Donoghue MJ (2010) Phylogeny and phylogenetic nomenclature of the Campanulidae based on an expanded sample of genes and taxa. Syst Bot 35:425–441
Timme RE, Bachvaroff TR, Delwiche CF (2012) Broad phylogenomic sampling and the sister lineage of land plants. PLoS One 7:e29696. doi:10.1371/journal.pone.0029696
Uhl C (1978) Chromosomes of Mexican *Sedum* II. Section *Pachysedum*. Rhodora 80:491–512

Vamosi JC, Knight TM, Steets JA, Mazer SJ, Burd M, Ashman TL (2006) Pollination decays in biodiversity hotspots. Proc Natl Acad Sci USA 103:956–961

Wang H, Moore MJ, Soltis PS, Bell CD, Brockington SF, Alexandre R, Davis CC, Latvis M, Manchester SR, Soltis DE (2009) Rosid diversification and the rapid rise of angiosperm-dominated forests. Proc Natl Acad Sci USA 106:3853–3858

Webb CO, Cannon CH, Davies SJ (2008) Ecological organization, biogeography, and the phylogenetic structure of tropical forest tree communities. In: Carson WP, Schnitzer SA (eds) Tropical forest community ecology. Wiley-Blackwell, Oxford, pp 79–97

Wikström N, Savolainen V, Chase MW (2001) Evolution of the angiosperms: calibrating the family tree. Proc Roy Soc Lond B Bio 268:2211–2220

Won H, Renner SS (2003) Horizontal gene transfer from flowering plants to *Gnetum*. Proc Natl Acad Sci USA 100: 10824–10829

Wurdack KJ, Davis CC (2009) Malpighiales phylogenetics: gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. Am J Bot 96:1551–1570

Zahn LM, Kong H, Leebens-Mack JH, Kim S, Soltis PS, Landherr LL, Soltis DE, dePamphilis CW, Ma H (2005) The evolution of the *SEPALLATA* subfamily of MADS-box genes: a pre-angiosperm origin with multiple duplications throughout angiosperm history. Genetics 169:2209–2223

Zwickl DJ (2006) GARLI. Genetic algorithm for rapid likelihood inference. Software available at http://www.bio.utexas.edu/grad/zwickl/web/garli.html

The Plant Nucleus at War and Peace: Genome Organization in the Interphase Nucleus

2

R. Neil Jones and Tim Langdon

Contents

2.1 **Introduction** 13

2.2 **General Aspects of Organization in the Peaceful Nucleus** 14
2.2.1 Chromosome Territories (CT) 15
2.2.2 Modifications of Chromatin and DNA 15

2.3 **The Nucleus Under Strain** 17
2.3.1 Plasticity of Gene Family Organization 17
2.3.2 Limits to Expression Neighbourhoods 18
2.3.3 Hybrid Vigour 19
2.3.4 Mobile Elements 19

2.4 **The Nucleus at War** 19
2.4.1 Ploidy Levels 20
2.4.2 Changes at the Gene Level 21
2.4.3 Adjustments at the Chromosome Level 23

2.5 **Concluding Remarks** 26

References 27

2.1 Introduction

The organization of the plant nucleus has been seen as a key to understanding the workings of plants themselves for the past two centuries. This aspiration seems to be being finally realised as new sequencing technologies are helping to draw together old strands of research as well as supporting novel approaches to bridge the gap between cytological and molecular scales of description; and recent advances in understanding epigenetic processes and mechanisms are providing a fresh perspective on studies of interphase organization. The inevitable caveat is that new questions may need to be answered, as the comparative detail now available throws up additional complexities which were previously hidden. Chief among these is the extent to which the genome is not constant within a species. Limited surveys of genes and genomic regions in maize revealed some years ago that there was more genetic diversity within this one species than between humans and chimpanzees; and a more recent genome wide survey indicates that if anything this was an underestimate (Gore et al. 2009). More generally in both animals and plants, intragenic copy number variation in non-repetitive sequences is being found to rival that seen in heterochromatic repeats; 10% or more of coding sequences show copy number variation in maize and rice (Ding et al. 2007; Springer et al. 2009). Nuclear organization in many plants must therefore be capable of accommodating a wide range of structural and nucleotide polymorphisms without suffering detrimental effects; indeed, maize itself demonstrates that the hybridization of divergent parents may be strongly advantageous (Shull 1948). This plasticity may explain the ability of many wide crosses to eventually generate stable derivatives, which in turn is likely to underpin many of the numerous examples of reticulate evolution being found. Nevertheless, interspecific hybridizations have often been shown to trigger genome wide reactions, frequently followed by structural and epigenetic reorganization in subsequent generations, which can be seen as a war between parental genomes. Here we review some of the aspects of nuclear organization that may underlie

R.N. Jones (✉)
Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Edward Llwyd Building, Penglais Campus, Aberystwyth, Wales SY23 3DA, UK
e-mail: rnj@aber.ac.uk

I.J. Leitch et al. (eds.), *Plant Genome Diversity Volume 2*,
DOI 10.1007/978-3-7091-1160-4_2, © Springer-Verlag Wien 2013

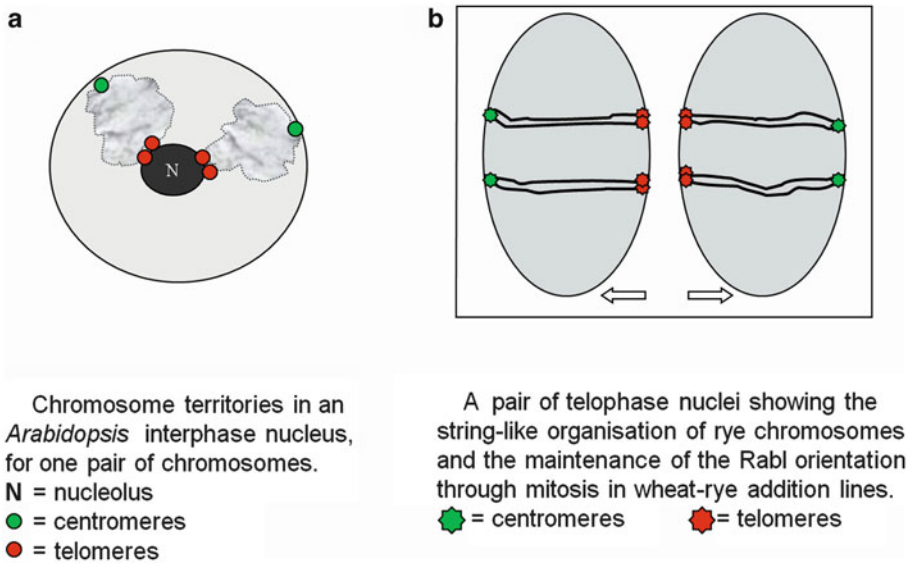

Fig. 2.1 Highly schematic representation of the disposition of chromosomes in species with a small (**a**) *(Arabidopsis)* and a large (**b**) (hexaploid wheat) genome. Probes identify the centromeres (*red*) and telomeres (*green*). The chromosome territories in *Arabidopsis* are visualized with chromosome-specific probes from pools of BAC contigs; whereas in wheat individual chromosomes, present as a wheat/rye addition line 1R, can be seen by probing with whole genome DNA which discriminates between the dispersed repeats of wheat and rye chromatin

incompatibilities, and indicate some of the changes that may be required to restore if not peace, then at least a truce.

2.2 General Aspects of Organization in the Peaceful Nucleus

A number of significant discoveries on the structure of the nucleus were made early on, at the limits of what could then be observed with the light microscope. In 1885 Rabl first described the arrangement of the chromosomes of salamanders in terms of the orientation of their centromeres and telomeres. He explained how the anaphase configuration, with the telomeres at the nuclear periphery of adjacent daughter cells and the centromeres at the relic poles, was maintained through to the following interphase (Rabl 1885). Laibach (1907) later made the remarkable finding that the number of chromocentres in *Arabidopsis* corresponded to the number of chromosomes, and Heitz (1928) demonstrated the continuity of the structure of chromosomes throughout the cell cycle in *Pellia endiviifolia*, using blocks of constitutive heterochromatin as markers for the identity of individual chromosomes. An early inference was that the interphase nucleus would be made up of discrete chromosomal territories (CTs) (Boveri 1909), and each subsequent development of microscopic techniques has improved our understanding of the nuclear architecture of CTs and its impact on gene expression (reviewed in Rouquette et al. 2010). Most recently, direct 3D analysis of living cells using green fluorescent protein (GFP) tags and BACs, and of accurately fixed material by confocal methods, has been extended by indirect analysis via sequencing of DNA retrieved from common pools of cross linked proteins (3C, 4C, 5C and HiC methods), which is providing resolution beyond the limits of microscopy (Rajapakse and Groudine 2011). Nevertheless, the role and even the extent of CTs remains unclear, and it appears that this is as much because of inherent variability as because of technical limitations.

This variability is well illustrated by the inconsistent occurrence even of Rabl arrangements. They do appear to be a fixture of plant species with relatively large genomes (such as wheat, oat, barley, *Vicia faba* and *Allium cepa*, all with genomes in excess of 1C = 5,000 Mb), where probes identifying alien segments can be used to show chromosome arms assuming a string-like form running between the centromeric and telomeric nuclear poles (Abranches et al. 1998) (Fig. 2.1). On the other hand, plant species with small genomes (such as sorghum, rice, *Arabidopsis*, and brassicas, with genomes of less than 1C = 1,000 Mb) generally lack the Rabl arrangement. This can be observed, for example, by 3D reconstruction of living cells of *A. thaliana* after tagging with a GFP probe, where the centromeres are predominantly dispersed around the nuclear periphery in different cell types (Fang and Spector 2005) although the telomeres are associated within the nucleus core around the nucleolus (Armstrong et al. 2001). As a result individual *Arabidopsis* chromosomes appear as radial euchromatic loops emanating from the heterochromatic centromeric chromocentres when painted with specific BAC markers (Fransz et al. 2002). Maize, with 1C = 3,000 Mb of