

**GENETICS, GENOMICS  
AND BREEDING OF  
CONIFERS**

## **Genetics, Genomics and Breeding of Crop Plants**

### *Series Editor*

Chittaranjan Kole  
Department of Genetics and Biochemistry  
Clemson University  
Clemson, SC  
USA

### **Books in this Series:**

#### *Published or in Press:*

- Jinguo Hu, Gerald Seiler & Chittaranjan Kole: *Sunflower*
- Kristin D. Bilyeu, Milind B. Ratnaparkhe & Chittaranjan Kole: *Soybean*
- Robert Henry & Chittaranjan Kole: *Sugarcane*
- Kevin Folta & Chittaranjan Kole: *Berries*
- Jan Sadowsky & Chittaranjan Kole: *Vegetable Brassicas*
- James M. Bradeen & Chittaranjan Kole: *Potato*
- C.P. Joshi, Stephen DiFazio & Chittaranjan Kole: *Poplar*
- Anne-Françoise Adam-Blondon, José M. Martínez-Zapater & Chittaranjan Kole: *Grapes*
- Christophe Plomion, Jean Bousquet & Chittaranjan Kole: *Conifers*

#### *Books under preparation:*

- Dave Edwards, Jacqueline Batley, Isobel Parkin & Chittaranjan Kole: *Oilseed Brassicas*

# GENETICS, GENOMICS AND BREEDING OF CONIFERS

*Editors*

**Christophe Plomion**

INRA

UMR BIOGECO

Cestas

France

**Jean Bousquet**

Centre d'étude de la forêt

Université Laval

Québec

Canada

**Chittaranjan Kole**

Department of Genetics and Biochemistry

Clemson University

Clemson, SC

USA



**CRC Press**  
Taylor & Francis Group  
an **informa** business  
[www.crcpress.com](http://www.crcpress.com)

6000 Broken Sound Parkway, NW  
Suite 300, Boca Raton, FL 33487  
270 Madison Avenue  
New York, NY 10016  
2 Park Square, Milton Park  
Abingdon, Oxon OX14 4RN, UK



**Science Publishers**  
Enfield, New Hampshire

iv *Genetics, Genomics and Breeding of Conifers*

Published by Science Publishers, P.O. Box 699, Enfield, NH 03748, USA  
An imprint of Edenbridge Ltd., British Channel Islands

E-mail: [info@scipub.net](mailto:info@scipub.net)

Website: [www.scipub.net](http://www.scipub.net)

**Marketed and distributed by:**



Copyright reserved © 2011

ISBN 978-1-57808-000-0

*The views expressed in this book are those of the author(s) and the publisher does not assume responsibility for the authenticity of the findings/conclusions drawn by the author(s). Also no responsibility is assumed by the publishers for any damage to the property or persons as a result of operation or use of this publication and/or the information contained herein.*

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying or otherwise, without the prior permission of the publisher, in writing. The exception to this is when a reasonable part of the text is quoted for purpose of book review, abstracting etc.

This book is sold subject to the condition that it shall not, by way of trade or otherwise be lent, re-sold, hired out, or otherwise circulated without the publisher's prior consent in any form of binding or cover other than that in which it is published and without a similar condition including this condition being imposed on the subsequent purchaser.

Printed in the United States of America

## **Preface to the Series**

Genetics, genomics and breeding has emerged as three overlapping and complimentary disciplines for comprehensive and fine-scale analysis of plant genomes and their precise and rapid improvement. While genetics and plant breeding have contributed enormously towards several new concepts and strategies for elucidation of plant genes and genomes as well as development of a huge number of crop varieties with desirable traits, genomics has depicted the chemical nature of genes, gene products and genomes and also provided additional resources for crop improvement.

In today's world, teaching, research, funding, regulation and utilization of plant genetics, genomics and breeding essentially require thorough understanding of their components including classical, biochemical, cytological and molecular genetics; and traditional, molecular, transgenic and genomics-assisted breeding. There are several book volumes and reviews available that cover individually or in combination of a few of these components for the major plants or plant groups; and also on the concepts and strategies for these individual components with examples drawn mainly from the major plants. Therefore, we planned to fill an existing gap with individual book volumes dedicated to the leading crop and model plants with comprehensive deliberations on all the classical, advanced and modern concepts of depiction and improvement of genomes. The success stories and limitations in the different plant species, crop or model, must vary; however, we have tried to include a more or less general outline of the contents of the chapters of the volumes to maintain uniformity as far as possible.

Often genetics, genomics and plant breeding and particularly their complimentary and supplementary disciplines are studied and practiced by people who do not have, and reasonably so, the basic understanding of biology of the plants for which they are contributing. A general description of the plants and their botany would surely instill more interest among them on the plant species they are working for and therefore we presented lucid details on the economic and/or academic importance of the plant(s); historical information on geographical origin and distribution; botanical origin and evolution; available germplasm and gene pools, and genetic and cytogenetic stocks as genetic, genomic and breeding resources; and

basic information on taxonomy, habit, habitat, morphology, karyotype, ploidy level and genome size, etc.

Classical genetics and traditional breeding have contributed enormously even by employing the phenotype-to-genotype approach. We included detailed descriptions on these classical efforts such as genetic mapping using morphological, cytological and isozyme markers; and achievements of conventional breeding for desirable and against undesirable traits. Employment of the *in vitro* culture techniques such as micro- and megaspore culture, and somatic mutation and hybridization, has also been enumerated. In addition, an assessment of the achievements and limitations of the basic genetics and conventional breeding efforts has been presented.

It is a hard truth that in many instances we depend too much on a few advanced technologies, we are trained in, for creating and using novel or alien genes but forget the infinite wealth of desirable genes in the indigenous cultivars and wild allied species besides the available germplasms in national and international institutes or centers. Exploring as broad as possible natural genetic diversity not only provides information on availability of target donor genes but also on genetically divergent genotypes, botanical varieties, subspecies, species and even genera to be used as potential parents in crosses to realize optimum genetic polymorphism required for mapping and breeding. Genetic divergence has been evaluated using the available tools at a particular point of time. We included discussions on phenotype-based strategies employing morphological markers, genotype-based strategies employing molecular markers; the statistical procedures utilized; their utilities for evaluation of genetic divergence among genotypes, local landraces, species and genera; and also on the effects of breeding pedigrees and geographical locations on the degree of genetic diversity.

Association mapping using molecular markers is a recent strategy to utilize the natural genetic variability to detect marker-trait association and to validate the genomic locations of genes, particularly those controlling the quantitative traits. Association mapping has been employed effectively in genetic studies in human and other animal models and those have inspired the plant scientists to take advantage of this tool. We included examples of its use and implication in some of the volumes that devote to the plants for which this technique has been successfully employed for assessment of the degree of linkage disequilibrium related to a particular gene or genome, and for germplasm enhancement.

Genetic linkage mapping using molecular markers have been discussed in many books, reviews and book series. However, in this series, genetic mapping has been discussed at length with more elaborations and examples on diverse markers including the anonymous type 2 markers such as RFLPs, RAPDs, AFLPs, etc. and the gene-specific type 1 markers such as EST-SSRs, SNPs, etc.; various mapping populations including  $F_2$ , backcross,

recombinant inbred, doubled haploid, near-isogenic and pseudotestcross; computer software including MapMaker, JoinMap, etc. used; and different types of genetic maps including preliminary, high-resolution, high-density, saturated, reference, consensus and integrated developed so far.

Mapping of simply inherited traits and quantitative traits controlled by oligogenes and polygenes, respectively has been deliberated in the earlier literature crop-wise or crop group-wise. However, more detailed information on mapping or tagging oligogenes by linkage mapping or bulked segregant analysis, mapping polygenes by QTL analysis, and different computer software employed such as MapMaker, JoinMap, QTL Cartographer, Map Manager, etc. for these purposes have been discussed at more depth in the present volumes.

The strategies and achievements of marker-assisted or molecular breeding have been discussed in a few books and reviews earlier. However, those mostly deliberated on the general aspects with examples drawn mainly from major plants. In this series, we included comprehensive descriptions on the use of molecular markers for germplasm characterization, detection and maintenance of distinctiveness, uniformity and stability of genotypes, introgression and pyramiding of genes. We have also included elucidations on the strategies and achievements of transgenic breeding for developing genotypes particularly with resistance to herbicide, biotic and abiotic stresses; for biofuel production, biopharming, phytoremediation; and also for producing resources for functional genomics.

A number of desirable genes and QTLs have been cloned in plants since 1992 and 2000, respectively using different strategies, mainly positional cloning and transposon tagging. We included enumeration of these and other strategies for isolation of genes and QTLs, testing of their expression and their effective utilization in the relevant volumes.

Physical maps and integrated physical-genetic maps are now available in most of the leading crop and model plants owing mainly to the BAC, YAC, EST and cDNA libraries. Similar libraries and other required genomic resources have also been developed for the remaining crops. We have devoted a section on the library development and sequencing of these resources; detection, validation and utilization of gene-based molecular markers; and impact of new generation sequencing technologies on structural genomics.

As mentioned earlier, whole genome sequencing has been completed in one model plant (*Arabidopsis*) and seven economic plants (rice, poplar, peach, papaya, grapes, soybean and sorghum) and is progressing in an array of model and economic plants. Advent of massively parallel DNA sequencing using 454-pyrosequencing, Solexa Genome Analyzer, SOLiD system, Heliscope and SMRT have facilitated whole genome sequencing in many other plants more rapidly, cheaply and precisely. We have included

extensive coverage on the level (national or international) of collaboration and the strategies and status of whole genome sequencing in plants for which sequencing efforts have been completed or are progressing currently. We have also included critical assessment of the impact of these genome initiatives in the respective volumes.

Comparative genome mapping based on molecular markers and map positions of genes and QTLs practiced during the last two decades of the last century provided answers to many basic questions related to evolution, origin and phylogenetic relationship of close plant taxa. Enrichment of genomic resources has reinforced the study of genome homology and synteny of genes among plants not only in the same family but also of taxonomically distant families. Comparative genomics is not only delivering answers to the questions of academic interest but also providing many candidate genes for plant genetic improvement.

The 'central dogma' enunciated in 1958 provided a simple picture of gene function—gene to mRNA to transcripts to proteins (enzymes) to metabolites. The enormous amount of information generated on characterization of transcripts, proteins and metabolites now have led to the emergence of individual disciplines including functional genomics, transcriptomics, proteomics and metabolomics. Although all of them ultimately strengthen the analysis and improvement of a genome, they deserve individual deliberations for each plant species. For example, microarrays, SAGE, MPSS for transcriptome analysis; and 2D gel electrophoresis, MALDI, NMR, MS for proteomics and metabolomics studies require elaboration. Besides transcriptome, proteome or metabolome QTL mapping and application of transcriptomics, proteomics and metabolomics in genomics-assisted breeding are frontier fields now. We included discussions on them in the relevant volumes.

The databases for storage, search and utilization on the genomes, genes, gene products and their sequences are growing enormously in each second and they require robust bioinformatics tools plant-wise and purpose-wise. We included a section on databases on the gene and genomes, gene expression, comparative genomes, molecular marker and genetic maps, protein and metabolomes, and their integration.

Notwithstanding the progress made so far, each crop or model plant species requires more pragmatic retrospect. For the model plants we need to answer how much they have been utilized to answer the basic questions of genetics and genomics as compared to other wild and domesticated species. For the economic plants we need to answer as to whether they have been genetically tailored perfectly for expanded geographical regions and current requirements for green fuel, plant-based bioproducts and for improvements of ecology and environment. These futuristic explanations have been addressed finally in the volumes.



We are aware of exclusions of some plants for which we have comprehensive compilations on genetics, genomics and breeding in hard copy or digital format and also some other plants which will have enough achievements to claim for individual book volume only in distant future. However, we feel satisfied that we could present comprehensive deliberations on genetics, genomics and breeding of 30 model and economic plants, and their groups in a few cases, in this series. I personally feel also happy that I could work with many internationally celebrated scientists who edited the book volumes on the leading plants and plant groups and included chapters authored by many scientists reputed globally for their contributions on the concerned plant or plant group.

We paid serious attention to reviewing, revising and updating of the manuscripts of all the chapters of this book series, but some technical and formatting mistakes will remain for sure. As the series editor, I take complete responsibility for all these mistakes and will look forward to the readers for corrections of these mistakes and also for their suggestions for further improvement of the volumes and the series so that future editions can serve better the purposes of the students, scientists, industries, and the society of this and future generations.

Science publishers, Inc. has been serving the requirements of science and society for a long time with publications of books devoted to advanced concepts, strategies, tools, methodologies and achievements of various science disciplines. Myself as the editor and also on behalf of the volume editors, chapter authors and the ultimate beneficiaries of the volumes take this opportunity to acknowledge the publisher for presenting these books that could be useful for teaching, research and extension of genetics, genomics and breeding.

Chittaranjan Kole



## **Preface to the Volume**

Conifers are woody plants, the great majority being trees. They represent 650 species, some ranking as the largest, tallest, and longest living non-clonal terrestrial organisms on Earth. They are of immense ecological importance, dominating many terrestrial landscapes and representing the largest terrestrial carbon sink. They are evolutionary distinct from angiosperm trees on many accounts and with their extraordinary large genomes, they provide a different view of plant genome biology and evolution. They are also of great economic importance, as they are primarily used for timber and paper production worldwide. Domestication of some of these species was started about 60 years ago through traditional genetic improvement programs. It has resulted in advances in overall growth, wood quality, pest resistance and adaptation, but breeding still remains a slow process because of long generation intervals typical of most conifers and because most traits cannot be correctly evaluated at an early stage.

During the past 20 years, more and more sophisticated genomics tools have been developed to describe the extreme plasticity and variability of these species at different levels of integration (from genes up to phenotypes) and are now being integrated into breeding to accelerate the domestication process by a more precise exploitation of genetic diversity. Application of genomic-based science is also playing an important role in understanding the evolution, patterns of nucleotide variation and the molecular basis of quantitative traits and adaptation. Altogether, this new knowledge is also expected to help delineate more efficient gene conservation strategies.

This book will give the reader an in-depth review of the current state-of-the-art of genetic and genomic research conducted in conifers. Each chapter is the product of specialists in their field. Their goal was to report on the latest trends and findings and at the same time, promote awareness and make this knowledge accessible to the vast majority. Accordingly, the chapters are well documented and illustrated. Their contribution is greatly appreciated.

The book begins with an exhaustive description of the conifers in terms of classification, geographical distribution, life history and ecology, morphology and fossil history as well as phylogenetics (Chapter 1). It is followed by a chapter devoted to their economic importance and the

development of conifer breeding programs worldwide, which lead to significant improvement of productivity and quality (Chapter 2). Chapter 3 deliberates on various classical and molecular cytogenetical tools useful to elucidate evolution, integrate physical and genetic maps, conserve species and assist in marker-based breeding. Chapter 4 describes the applications of neutral genetic markers from the perspectives of conservation genetics, phylogeography and gene flow studies. In Chapter 5, research efforts on linkage mapping, emerging gene maps as well as QTL detection and architecture are reviewed. An exhaustive review of investigations on candidate genes is provided in Chapter 6, from estimates of nucleotide diversity and recombination to new-generation selection signatures studies and the development of association mapping and outlier detection approaches. The ever-increasing applications of molecular markers into breeding from the management operations to selection strategies are considered in Chapter 7. Switching to more functional aspects, Chapters 8 and 9 review the current status of our understanding of transcriptome, proteome and metabolome modifications in responses to developmental changes and environmental constraints. The rapid advances in sequencing and cataloging the conifer gene space are also reported (Chapter 8). As a prerequisite for the sequencing of a conifer genome, insights into the characteristics of the large conifer genomes, especially with respect to the composition and evolution of transposable elements, are provided in Chapter 10. The book ends with refreshing views on the challenges faced by the conifer genomics community and how the pace of rapid advancement of the “omic” sciences might affect our understanding of conifer biology and the future use of conifer genetic resources (Chapter 11).

This book is a testimony to the substantial progress made in the field of conifer genetics and genomics and the definite value of conifers as a model system. Although the tools and concepts that are presented will continue to evolve rapidly, we hope this volume will provide a solid foundation for further development in conifer and more generally in forest tree genetics, genomics and breeding.

Christophe Plomion  
Jean Bousquet  
Chittaranjan Kole

# Contents

<i>Preface to the Series</i>	v
<i>Preface to the Volume</i>	xi
<i>List of Contributors</i>	xv
<i>Abbreviations</i>	xxiii
<b>1. The Conifers (Pinophyta)</b>	<b>1</b>
<i>David S. Gernandt, Ann Willyard, John V. Syring and Aaron Liston</i>	
<b>2. Economic Importance, Breeding Objectives and Achievements</b>	<b>40</b>
<i>T.J. Mullin, B. Andersson, J.-C. Bastien, J. Beaulieu, R.D. Burdon, W.S. Dvorak, J.N. King, T. Kondo, J. Krakowski, S.J. Lee, S.E. McKeand, L. Pâques, A. Raffin, J.H. Russell, T. Skrøppa, M. Stoehr and A. Yanchuk</i>	
<b>3. Cytogenetics</b>	<b>128</b>
<i>M. Nurul Islam-Faridi and C. Dana Nelson</i>	
<b>4. Neutral Patterns of Genetic Variation and Applications to Conservation in Conifer Species</b>	<b>141</b>
<i>Francesca Bagnoli, Bruno Fady, Silvia Fineschi, Sylvie Oddou-Muratorio, Andrea Piotti, Federico Sebastiani and Giovanni G. Vendramin</i>	
<b>5. Genetic Mapping in Conifers</b>	<b>196</b>
<i>Kermit Ritland, Konstantin V. Krutovskiy, Yoshihiko Tsumura, Betty Pelgas, Nathalie Isabel and Jean Bousquet</i>	
<b>6. Patterns of Nucleotide Diversity and Association Mapping</b>	<b>239</b>
<i>González-Martínez SC, Dillon S, Garnier-Géré PH, Krutovskiy KV, Alía R, Burgarella C, Eckert AJ, García-Gil MR, Grivet D, Heuertz M, Jaramillo-Correa JP, Lascoux M, Neale DB, Savolainen O, Tsumura Y and Vendramin GG</i>	
<b>7. Integration of Molecular Markers in Breeding</b>	<b>276</b>
<i>Rowland D. Burdon and Phillip L. Wilcox</i>	

<b>8. Transcriptomics</b>	<b>323</b>
<i>John J. Mackay and Jeffrey F. D. Dean</i>	
<b>9. Recent Advances in Proteomics and Metabolomics in Gymnosperms</b>	<b>358</b>
<i>Rebecca Dauwe, Andrew Robinson and Shawn D. Mansfield</i>	
<b>10. Toward the Conifer Genome Sequence</b>	<b>389</b>
<i>Michele Morgante and Emanuele De Paoli</i>	
<b>11. Future Prospects</b>	<b>404</b>
<i>Jeffrey F.D. Dean</i>	
<i>Index</i>	<b>439</b>
<i>Color Plate Section</i>	<b>000</b>