

Compendium of Plant Genomes
Series Editor: Chittaranjan Kole

Malali Gowda
Ambardar Sheetal
Chittaranjan Kole *Editors*

The Neem Genome

Compendium of Plant Genomes

Series Editor

Chittaranjan Kole, Raja Ramanna Fellow, Government of India,
ICAR-National Research Center on Plant Biotechnology, Pusa,
New Delhi, India

Whole-genome sequencing is at the cutting edge of life sciences in the new millennium. Since the first genome sequencing of the model plant *Arabidopsis thaliana* in 2000, whole genomes of about 100 plant species have been sequenced and genome sequences of several other plants are in the pipeline. Research publications on these genome initiatives are scattered on dedicated web sites and in journals with all too brief descriptions. The individual volumes elucidate the background history of the national and international genome initiatives; public and private partners involved; strategies and genomic resources and tools utilized; enumeration on the sequences and their assembly; repetitive sequences; gene annotation and genome duplication. In addition, synteny with other sequences, comparison of gene families and most importantly potential of the genome sequence information for gene pool characterization and genetic improvement of crop plants are described.

Interested in editing a volume on a crop or model plant?

Please contact Prof. C. Kole, Series Editor, at ckoleorg@gmail.com

More information about this series at <http://www.springer.com/series/11805>

Malali Gowda · Ambardar Sheetal ·
Chittaranjan Kole
Editors

The Neem Genome

 Springer

Editors

Malali Gowda
TransDisciplinary University Yelahanka
Bengaluru, Karnataka, India

Ambardar Sheetal
TransDisciplinary University Yelahanka
Bengaluru, Karnataka, India

Chittaranjan Kole
Raja Ramanna Fellow,
Government of India
ICAR-National Research Center
on Plant Biotechnology, Pusa
New Delhi, India

ISSN 2199-4781

ISSN 2199-479X (electronic)

Compendium of Plant Genomes

ISBN 978-3-030-16121-7

ISBN 978-3-030-16122-4 (eBook)

<https://doi.org/10.1007/978-3-030-16122-4>

© Springer Nature Switzerland AG 2019

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

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

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

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

*This book series is dedicated to my wife Phullara,
and our children Sourav, and Devleena*
Chittaranjan Kole

Preface to the Series

Genome sequencing has emerged as the leading discipline in the plant sciences coinciding with the start of the new century. For much of the twentieth century, plant geneticists were only successful in delineating putative chromosomal location, function, and changes in genes indirectly through the use of a number of “markers” physically linked to them. These included visible or morphological, cytological, protein, and molecular or DNA markers. Among them, the first DNA marker, the RFLPs, introduced a revolutionary change in plant genetics and breeding in the mid-1980s, mainly because of their infinite number and thus potential to cover maximum chromosomal regions, phenotypic neutrality, absence of epistasis, and codominant nature. An array of other hybridization-based markers, PCR-based markers, and markers based on both facilitated construction of genetic linkage maps, mapping of genes controlling simply inherited traits, and even gene clusters (QTLs) controlling polygenic traits in a large number of model and crop plants. During this period, a number of new mapping populations beyond F₂ were utilized and a number of computer programs were developed for map construction, mapping of genes, and mapping of polygenic clusters or QTLs. Molecular markers were also used in the studies of evolution and phylogenetic relationship, genetic diversity, DNA fingerprinting, and map-based cloning. Markers tightly linked to the genes were used in crop improvement employing the so-called marker-assisted selection. These strategies of molecular genetic mapping and molecular breeding made a spectacular impact during the last one and a half decades of the twentieth century. But still, they remained “indirect” approaches for elucidation and utilization of plant genomes since much of the chromosomes remained unknown and the complete chemical depiction of them was yet to be unraveled.

Physical mapping of genomes was the obvious consequence that facilitated the development of the “genomic resources” including BAC and YAC libraries to develop physical maps in some plant genomes. Subsequently, integrated genetic–physical maps were also developed in many plants. This led to the concept of structural genomics. Later on, emphasis was laid on EST and transcriptome analysis to decipher the function of the active gene sequences leading to another concept defined as functional genomics. The advent of techniques of bacteriophage gene and DNA sequencing in the 1970s was extended to facilitate sequencing of these genomic resources in the last decade of the twentieth century.

As expected, sequencing of chromosomal regions would have led to too much data to store, characterize, and utilize with the-then available computer software could handle. But the development of information technology made the life of biologists easier by leading to a swift and sweet marriage of biology and informatics, and a new subject was born—bioinformatics.

Thus, the evolution of the concepts, strategies, and tools of sequencing and bioinformatics reinforced the subject of genomics—structural and functional. Today, genome sequencing has traveled much beyond biology and involves biophysics, biochemistry, and bioinformatics!

Thanks to the efforts of both public and private agencies, genome sequencing strategies are evolving very fast, leading to cheaper, quicker, and automated techniques right from clone-by-clone and whole-genome shotgun approaches to a succession of second-generation sequencing methods. The development of software of different generations facilitated this genome sequencing. At the same time, newer concepts and strategies were emerging to handle sequencing of the complex genomes, particularly the polyploids.

It became a reality to chemically—and so directly—define plant genomes, popularly called whole-genome sequencing or simply genome sequencing.

The history of plant genome sequencing will always cite the sequencing of the genome of the model plant *Arabidopsis thaliana* in 2000 that was followed by sequencing the genome of the crop and model plant rice in 2002. Since then, the number of sequenced genomes of higher plants has been increasing exponentially, mainly due to the development of cheaper and quicker genomic techniques and, most importantly, the development of collaborative platforms such as national and international consortia involving partners from public and/or private agencies.

As I write this preface for the first volume of the new series “Compendium of Plant Genomes,” a net search tells me that complete or nearly complete whole-genome sequencing of 45 crop plants, eight crop and model plants, eight model plants, 15 crop progenitors and relatives, and 3 basal plants is accomplished, the majority of which are in the public domain. This means that we nowadays know many of our model and crop plants chemically, i.e., directly, and we may depict them and utilize them precisely better than ever. Genome sequencing has covered all groups of crop plants. Hence, information on the precise depiction of plant genomes and the scope of their utilization are growing rapidly every day. However, the information is scattered in research articles and review papers in journals and dedicated Web pages of the consortia and databases. There is no compilation of plant genomes and the opportunity of using the information in sequence-assisted breeding or further genomic studies. This is the underlying rationale for starting this book series, with each volume dedicated to a particular plant.

Plant genome science has emerged as an important subject in academia, and the present compendium of plant genomes will be highly useful to both students and teaching faculties. Most importantly, research scientists involved in genomics research will have access to systematic deliberations on the plant genomes of their interest. Elucidation of plant genomes is of interest not only for the geneticists and breeders, but also for practitioners of an array of plant science disciplines, such as taxonomy, evolution, cytology,

physiology, pathology, entomology, nematology, crop production, biochemistry, and obviously bioinformatics. It must be mentioned that information regarding each plant genome is ever-growing. The contents of the volumes of this compendium are, therefore, focusing on the basic aspects of the genomes and their utility. They include information on the academic and/or economic importance of the plants, description of their genomes from a molecular genetic and cytogenetic point of view, and the genomic resources developed. Detailed deliberations focus on the background history of the national and international genome initiatives, public and private partners involved, strategies and genomic resources and tools utilized, enumeration on the sequences and their assembly, repetitive sequences, gene annotation, and genome duplication. In addition, synteny with other sequences, comparison of gene families, and, most importantly, the potential of the genome sequence information for gene pool characterization through genotyping by sequencing (GBS) and genetic improvement of crop plants have been described. As expected, there is a lot of variation of these topics in the volumes based on the information available on the crop, model, or reference plants.

I must confess that as the series editor, it has been a daunting task for me to work on such a huge and broad knowledge base that spans so many diverse plant species. However, pioneering scientists with lifetime experience and expertise on the particular crops did excellent jobs editing the respective volumes. I myself have been a small science worker on plant genomes since the mid-1980s and that provided me the opportunity to personally know several stalwarts of plant genomics from all over the globe. Most, if not all, of the volume editors are my longtime friends and colleagues. It has been highly comfortable and enriching for me to work with them on this book series. To be honest, while working on this series I have been and will remain a student first, a science worker second, and a series editor last. And I must express my gratitude to the volume editors and the chapter authors for providing me the opportunity to work with them on this compendium.

I also wish to mention here my thanks and gratitude to the Springer staff particularly Dr. Christina Eckey and Dr. Jutta Lindenborn for the earlier set of volumes and presently Ing. Zuzana Bernhart and Dr. Anette Lindqvist for all their timely help and support.

I always had to set aside additional hours to edit books besides my professional and personal commitments—hours I could and should have given to my wife, Phullara, and our kids, Sourav and Devleena. I must mention that they not only allowed me the freedom to take away those hours from them but also offered their support in the editing job itself. I am really not sure whether my dedication of this compendium to them will suffice to do justice to their sacrifices for the interest of science and the science community.

New Delhi, India

Prof. Chittaranjan Kole

Preface to the Volume

Neem (*Azadirachta indica* A. Juss) is a member of the mahogany family, Meliaceae. Neem trees are attractive broad-leaved evergreens that can grow up to 30 m tall and 2.5 m in girth. A Neem tree normally begins bearing fruit after 3–5 years and becomes fully productive in 10 years and it may live for more than two centuries. The tree is easily propagated—both sexually and vegetatively. It can be planted using seeds, seedlings, saplings, root suckers, or tissue culture. The Neem tree was first discovered in India about 4500 years ago. Neem is native to India and Burma, and also found in Bangladesh, Pakistan and Nepal. The tree is said to grow “almost anywhere” in the lowland tropics. However, it generally performs best in areas with annual rainfalls of 400–1200 mm. It thrives under the hottest conditions, where maximum shade temperature may soar past 50 °C, but it will not withstand freezing or extended cold. The taproot (at least in young specimens) may be as much as twice the height of the tree.

The Neem research was started in the 1920s by the Indian scientists; Neem’s ability to repel insects was reported in the scientific literature in 1926–1929 (National Research Council 1992). Neem ingredients are applied in Ayurveda, Unani, homeopathy, and modern medicine for the treatment of many infectious, metabolic, or cancer diseases (Mohammad 2016). Neem products have been used for centuries in the field of agriculture and medicine, and the plant extracts of Neem have enormous potential to influence modern agrochemical research. Neem has been well known for various medicinal properties like anti-inflammatory, antipyretic, antihistamine, antifungal, anti-bacterial, anti-ulcer, analgesic, antiarrhythmic, anti-tubercular, antimalarial, diuretic, spermicide, anti-arthritis, anti-protozoal, insect repellent, anti-feedant, and anti-hormonal properties. Neem has one of the richest sources of secondary metabolites in nature, and over 300 compounds have been isolated from different parts of the Neem plant.

In this context, this book compiles up-to-date information on research and development related to the Neem plant. This book comprises a total of 12 chapters, which starts with the traditional practices and knowledge of Neem (Chap. 1) and ends with the Neem microbiome (Chap. 12). The first two chapters (1 and 2) provide a general introduction to the Neem, traditional practices related to Neem as documented in Ayurveda and India’s mythology. Chapter 3 enumerates various secondary metabolites produced by Neem plant, which are crucial for multifunctional properties like anti-oxidation, anti-inflammation, antimalarial, and anti-carcinogenic activities. Chapter 4

describes application of Neem and its products in agriculture. Chapter 5 provides the information about phylogeny of Neem and related species in the Meliaceae family based on genomic resources. Chapter 6 summarizes various strategies and tools for sequencing the genome and transcriptome of Neem. Chapter 7 provides a comprehensive knowledge on assembly of nuclear genome, organelle genome (chloroplast and mitochondria), and transcriptome of Neem plant. Chapters 8 and 9 provide information regarding the presence of the repetitive sequences in the Neem genome and annotation of genome. Chapter 10 elucidates on the synteny analysis and comparison of gene families with related plant species for better understanding of gene evolution in Neem. Chapter 11 gives the detailed information about the tissue culture techniques to produce a triploid Neem plant. The content of this book ends with Chap. 12 which describes various endophytes inhabiting in Neem plant and producing the similar secondary metabolites as Neem plant. Altogether, this book will serve as a primary resource material for researchers, breeders, and students working on Neem plant.

These chapters have been authored by 13 scientists from different research institutes and universities in India. We express our thanks to them for their contributions and cooperation for this book project. Experts in the field reviewed each chapter, and thus we are thankful for their efforts to improve the quality of this compilation. Dr. Malali Gowda and Dr. Ambardar Sheetal express their thanks and high gratitude to Prof. Chittaranjan Kole, Series Editor of the “Compendium of Plant Genomes,” for giving the opportunity to co-edit this book and for his constant support and guidance right from the inception till publication of this book on *The Neem Genome*. The editors also acknowledge the help from all the staff of Springer Nature at all the stages.

Bengaluru, India
Bengaluru, India
New Delhi, India

Prof. Malali Gowda
Dr. Ambardar Sheetal
Prof. Chittaranjan Kole

References

- National Research Council (1992) *Neem: a tree for solving global problems*. The National Academies Press, Washington, DC. <https://doi.org/10.17226/1924>
- Mohammad AA (2016) Therapeutics role of *Azadirachta indica* (Neem) and their active constituents in diseases prevention and treatment. *Evid-Based Complement Altern Med* 2016, Article ID 7382506, 11 p. <http://dx.doi.org/10.1155/2016/7382506>

Contents

1	Neem: Traditional Knowledge from Ayurveda	1
	S. N. Venugopalan Nair, Naveen Shilpa, Thomas Vargheese and I. F. Tabassum	
2	Heritage of Neem–Peepal Tree Resides a Profound Scientific Facts	13
	K. L. Ashalatha and Malali Gowda	
3	Method to Quantify Plant Secondary Metabolites: Quantification of Neem Metabolites from Leaf, Bark, and Seed Extracts as an Example	21
	Kannan Rangiah and Malali Gowda	
4	Utilization of Neem and Neem Products in Agriculture	31
	Rishu Sharma and Chittaranjan Kole	
5	Phylogeny of Neem and Related Species in the Meliaceae Family	49
	Nagesh A. Kuravadi and Malali Gowda	
6	Strategies and Tools for Next Generation Sequencing	53
	Nagesh A. Kuravadi and Malali Gowda	
7	Neem Genome Assembly	59
	Nagesh A. Kuravadi and Malali Gowda	
8	Repetitive Sequences	65
	Nagesh A. Kuravadi and Malali Gowda	
9	Neem Genome Annotation	73
	Nagesh A. Kuravadi and Malali Gowda	
10	Comparison of Gene Families and Synteny Analysis from Neem Genome	93
	Nagesh A. Kuravadi and Malali Gowda	

- 11 Neem Tissue Culture**..... 99
Divya Mohan, Ashmita J. Tontanahal, B. N. Sathyanarayana
and Malali Gowda
- 12 Neem Microbiome**..... 111
Varalaxmi B. Agasimundin, Kannan Rangiah,
Ambardar Sheetal and Malali Gowda

Contributors

Varalaxmi B. Agasimundin Centre for Functional Genomics and Bioinformatics, TransDisciplinary University, Bengaluru, India;
Centre for Cellular and Molecular Platforms, Bengaluru, India

K. L. Ashalatha Centre for Functional Genomics and Bioinformatics, The University of TransDisciplinary Health Sciences and Technology, Bengaluru, India

Malali Gowda Center for Functional Genomics and Bio-Informatics, The University of TransDisciplinary Health Sciences and Technology, Foundation for Revitalization of Local Health Traditions, Bengaluru, India

Chittaranjan Kole National Research Centre on Plant Biotechnology, New Delhi, India

Nagesh A. Kuravadi Centre for Cellular and Molecular Platforms, National Centre for Biological Sciences, Bengaluru, Karnataka, India

Divya Mohan The University of TransDisciplinary and Health Sciences, Bengaluru, Karnataka, India;
Department of Horticulture, University of Agricultural Sciences, Bengaluru, Karnataka, India

Kannan Rangiah Centre for Cellular and Molecular Platforms, Bengaluru, India;
Food Safety and Analytical Quality Control Laboratory, CSIR-Central Food Technological Research Institute (CFTRI), Mysore, Karnataka, India

B. N. Sathyanarayana Department of Horticulture, University of Agricultural Sciences, Bengaluru, Karnataka, India

Rishu Sharma Department of Plant Pathology, BC Agricultural University, Mohanpur, West Bengal, India

Ambardar Sheetal Centre for Functional Genomics and Bioinformatics, TransDisciplinary University, Bengaluru, India;
Centre for Cellular and Molecular Platforms, Bengaluru, India

Naveen Shilpa The University of TransDisciplinary Health Sciences and Technology (TDU), The Center for Traditional Knowledge Informatics and Data Sciences, Yelahanka, Bengaluru, India