Debmalya Barh · Muhammad Sarwar Khan
Eric Davies *Editors*

PlantOmics: The Omics of Plant Science

Springer

Debmalya Barh
Muhammad Sarwar Khan • Eric Davies
Editors

PlantOmics: The Omics of Plant Science

Springer

Editors
Debmalya Barh
Department of Genomics
Institute of Integrative Omics and Applied Biotechnology (IIOAB)
Nonakuri, West Bengal, India

Muhammad Sarwar Khan
Centre of Biochemistry and Biotechnology
University of Agriculture
Faisalabad, Pakistan

Eric Davies
Department of Plant Biology
North Carolina State University
Raleigh, NC, USA

ISBN 978-81-322-2171-5 ISBN 978-81-322-2172-2 (eBook)
DOI 10.1007/978-81-322-2172-2

Library of Congress Control Number: 2015930013

Springer New Delhi Heidelberg New York Dordrecht London

Printed on acid-free paper

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

My loving Gul: Shaurya Shree

Foreword

Plants are the primary source of "most of our food, fuel, fibers, fabrics, and pharmaceuticals". Stresses (biotic and abiotic) are the major threats to plants, being the primary cause of crop yield losses worldwide. On the other hand, with the global population expected to reach nine billion by 2050, an increase in crop productivity and quality will be needed to meet the requirements. Each of the 29 chapters of this *PlantOmics: The Omics of Plant Science* book opens a door to exciting cutting-edge omics approaches and their applications to meet the future demands.

The flow of the chapters in the book is highly scientific and strategically organized to be easy to go. It starts with the topic omics approaches in model plants and their applications in improvement of maize and rice like major cereal crops. Chapters 2, 3, and 4 describe very important technologies such as spectroscopy (NIR, MIR, Raman), next generation sequencing (NGS), and functional genomics and their applications in current plant science. Chapters 5 and 6 deal with technical advancements and applications of cyto-mutagenomics and epigenomics in crop improvement. Chapter 7 gives a detailed account on plant miRNA biology, associated technologies, and their tailor-made applications to improve plant stress response.

Each topic dealt in Chaps. 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, and 25 is a unique imprint of this book. These chapters cover well established and several budding omics areas in plant science such as Plant Proteomics, Metabolomics, Glycomics, Lipidomics, Secretomics, Phenomics, Cytomics, Physiomics, Signalomics, Thiolomics, Organelle Omics, Micromorphomics, Microbiomics, Cryobionomics, Nanobiotechnology, and Plant Pharmacogenomics. Each of these chapters describes the latest technologies and applications of the respective omics in a very comprehensive way; therefore, they are up to date, easy to understand, and can be spontaneously adopted to expand the area of our research and development.

Chapters 26, 27, and 28 deal with computational and systems biology approaches in plant science making the book more useful to any kind of plant biology research, whether in a wet lab or *in silico*. The last chapter (Chap. 29) is very brief but interesting where the editors have provided valuable insights on the future directions of omics and plantomics. They have proposed several new areas in omics which we must explore towards development of an integrated meta-omics strategy to ensure the world and earth's health and related issues.

Overall, it is a great effort by Dr. Barh, his editorial team, and 90 expert contributors from 15 countries to make this highly resourceful, up-to-date, thought provoking, and worthwhile unique book for students and researchers in the field of cutting-edge plant omics sciences. I highly recommend the book for keeping you up to date in the field.

Professor Ana Paulina Barba de la Rosa, PhD
Molecular Biology Division
Head, Unit of Proteomics and Molecular Biomedicine
Instituto Potosino de Investigación
Científica y Tecnológica (IPICyT), Mexico
President, Mexican Proteomics Society, Mexico

Preface

The term "*omics*" depicts completeness. In the last two decades, the term has been suffixed with several biological topics to provide complete information on the subject. With the advent of new technologies, the arena of "*omics*" is increasing rapidly. However, most of the currently available books that deal with omics technologies and their applications are mainly focused on animal system. To fill this gap, we have introduced this *PlantOmics: The Omics of Plant Science* book to provide a complete spectrum of plant related omics to the students and researchers working in the field of cutting-edge plant molecular biotechnology. Equal focus has been given to the technological advancements as well as their specific applications. Therefore, the book provides a comprehensive account of the state-of-the-art latest developments and trends of *omics* approaches in plant science. Several topics have also highlighted the integrative omics strategies enabling the cost-effective development of superior plants for various purposes.

The book consists of *29 chapters* written by *90 experts* from *15 countries* that represent three-fourths of the globe. In the introductory chapter (Chap. 1), Dr. Agrawal and colleagues have described the omics of model plants where genomics, proteomics, transcriptomics, and metabolomics of model plants such as *Arabidopsis*, rice, and maize are dealt in detail. Further, this chapter also provides how these technology derived knowledge can be used for transgenomics, mapping for biotic and abiotic stresses, and marker assisted selection for crop improvement. In Chap. 2, Dr. Cozzolino's group has given a nice overview on the most commonly used spectroscopy techniques such as NIR, MIR, and Raman in plant omic analysis. To make the chapter more resourceful, Dr. Cozzolino has also demonstrated instrumentations and analytic software for these spectroscopy techniques. The hot topic, next generation sequencing (NGS), its technologies, various platforms, algorithms, and *de novo* assembly, annotation, and analysis of plant genome are given by Dr. Tiwary in Chap. 3. Chapter 4, by Dr. Jha and his colleagues, provides a comprehensive account of techniques associated with plant functional genomics and their applications. Drs. Talukdar and Sinjushin in Chap. 5 have described various techniques of cytogenomics and mutagenomics and their cost effective applications in plant breeding and biology. This chapter has also highlighted the mutations that cause alterations in antioxidant defense response to withstand diverse abiotic stresses to reveal intrinsic cellular and metabolic events towards sensitivity of seed plants to salinity, drought, metal toxicity

and other stresses, prospecting to formulate effective breeding strategies in different agro-climatic conditions. Epigenomics technologies and their potential applications in crop improvement are summarized in Chap. 6 by Drs. Shafiq and Khan. Especially, this chapter highlights the roles of chromatin remodeling mechanisms in response to environmental stimuli and their role in crop improvement. Chapter 7, by Dr. Boopathi, on Plant miRNomics gives a comprehensive account to explain how the miRNAs fine tune the gene expression and play key roles in developmental timing and patterning of structures in response to external and internal stimuli in plants. This chapter also provides how the miRNAs can be used to improve plant stress responses. Chapter 8, by Dr. Agrawal and his group, describes the recent technological progresses in plant proteomics and highlights the achievements made in understanding the plant proteomes and their applications. In Chap. 9, Dr. Sangwan and colleagues explain various technology platforms in plant metabolomics research and how the metabolomics is used in monitoring and assessing gene functions, stress responses, and to characterize post-genomic processes from a broad perspective along with the challenges the domain is facing. Dr. Khurana's group in Chap. 10 overviews the chemistry and technologies in plant glycomics. This chapter also gives summary of applications of glycomics in biopharming and several biological processes such as plant signaling, stress responses, and immunity. In the next chapter (Chap. 11), Dr. Namasivayam elucidates the chemistry and analytic technologies, lipid signaling in plants, lipidomes in plant defense mechanisms, and several other aspects of plant lipidomics. The comprehensive mechanisms regulating constitutive and induced secretome of diverse plants and their habitat along with technological approaches are discussed by Dr. Yadav and her group in Chap. 12. In Chap. 13, Dr. Rahman and colleagues give a detailed account on integrated-omics approaches in phenomics and its applications in plant and agriculture. Chapter 14, by Drs. Davies and Stankovic, describes how novel methods based on super-fast and super-resolution microscopy can be used in describing proteins, nucleic acids, cytoskeleton, and small molecules of major interest to plants. In Chapter 15, Dr. Karpiński and colleagues educate us on plant physiomics. The chapter provides insights on how the combined molecular-physiological events drive plant growth, development, acclimatization, and defense responses. Dr. Vian et al., in Chap. 16, have introduced the term "Signalomics" and have shown how novel methods can be used to analyze systemic signals including electrical and hydraulic signals in plants. In Chap. 17, Dr. Talukdar and colleagues elucidate the use of latest cutting-edge functional genomics tools to understand the plant thiol metabolism from source (soil) to sink (grains) in diverse arenas of "thiolomics". The next three chapters (Chaps. 18, 19, 20) are dedicated to organelle omics. Chapter 18, by Dr. de Luna Valdez et al., explores how chloroplasts organize their genomes and regulate their transcriptomes, proteomes, and metabolomes, trying to focus on classical knowledge and reviewing new datasets obtained through large-scale research projects and systems approaches that shed light on chloroplast functionality under the chloroplast omics chapter. In Chap. 19, Dr. Khan summarizes the developments from plastid genomics to gene expression and briefly describes how transplastome facilitates expression of

vaccines, therapeutics, and plantibodies, in addition to tailoring agronomic traits in plants. Plant mitochondrial omics (Chap. 20), by Dr. Mustafa and his colleagues, describes a detailed account on regulation of mitochondrial genes at transcriptional, post-transcriptional (splicing and RNA editing), translational, and post-translational levels in omics perspective. Chapter 21 describes "Micromorphomics", a term coined by Dr. Tulika Talukdar to explain how plants combat environmental stresses through collective morphological manifestations in their organs architectures. Chapter 21 is dedicated to microbiomics. In this chapter, Dr. Sharma's team has discussed technologies to identify new groups of microorganisms involved in plant diseases from microbiome of rhizosphere and roles of microbiome in plant health and related areas. Drs. Martinez-Montero and Harding in Chap. 22 (Cryobionomics) intend to explore the connections between stability and cryogenic/non-cryogenic stress factors with a view to aiding protocol improvement, optimization, and validation for plant genetic resources conservation with several examples. Chapter 24, by Dr. Kazi and colleagues, focuses on the development and use of "nanotechnology" for formulating agriculturally important chemicals (fertilizers) with more useful properties and their direct delivery as well as their applications in various agricultural sectors. Chapter 25, by the same group, systemically analyzes the recent developments in plant pharmacogenomics and its contributions in the field of molecular and pharmaceutical sciences. Dr. Somvanshi and colleagues in Chap. 26 have attempted to describe several machine learning approaches and their applications in plant biology in a very simple way. Similarly, in Chap. 27, Dr. Sarika's team has emphasized on a number of applications of bioinformatics in agriculture in view of functional genomics, data mining techniques, genome-wide association studies, high-performance computing facilities in agriculture, and various bioinformatics tools/databases important for breeders, biotechnologists, and pathologists. Chapter 28 (Plant systems biology), by Drs. Bhardwaj and Somvanshi, describes recent insights and advancements in systems biology approaches in order to understand how plant systems work. In the brief concluding chapter (Chap. 29), we, the editors, have proposed several omics terms under "*Futuromics*" centraling Plantomics to direct the future perspectives of plant omics in meta-omics era.

We believe that this book will be a valuable resource to all who are working on cutting-edge plant omics. We appreciate your comments and suggestions to improve the next edition.

Nonakuri, India — Debmalya Barh
Faisalabad, Pakistan — Muhammad Sarwar Khan
Raleigh, NC, USA — Eric Davies

Contents

Omics of Model Plants 1
Pawan Kumar Agrawal, B. Kalyana Babu, and Navinder Saini

Instrumental Techniques and Methods: Their Role in Plant Omics 33
Daniel Cozzolino, Alberto Fassio, Ernesto Restaino, and Esteban Vicente

Next-Generation Sequencing and Assembly of Plant Genomes 53
Basant K. Tiwary

Functional Genomics: Applications in Plant Science 65
Uday Chand Jha, Jayant S. Bhat, Basavanagouda S. Patil, Firoz Hossain, and Debmalya Barh

Cytogenomics and Mutagenomics in Plant Functional Biology and Breeding 113
Dibyendu Talukdar and Andrey Sinjushin

Plant Epigenetics and Crop Improvement 157
Sarfraz Shafiq and Abdul Rehman Khan

Plant miRNomics: Novel Insights in Gene Expression and Regulation 181
N. Manikanda Boopathi

Plant Proteomics: Technologies and Applications 213
Deepti Bhushan Gupta, Shubhendu Shekhar, and Lalit Agrawal

Plant Metabolomics: An Overview of Technology Platforms for Applications in Metabolism 257
Neelam S. Sangwan, Pragya Tiwari, Siddhartha Kumar Mishra, Ritesh K. Yadav, Swati Tripathi, Amit K. Kushwaha, and Rajender Singh Sangwan

Plant Glycomics: Advances and Applications 299
Sarika Yadav, Dinesh K. Yadav, Neelam Yadav, and S.M. Paul Khurana

Plant Lipidomics: Signalling and Analytical Strategies 331
Elangovan Namasivayam, R. Kowsalya, Pavan Kumar Padarthi, K. Manigandan, Richard L. Jayaraj, Johnravindar D and Kaliaperumal Jagatheesh

Plant Secretomics: Unique Initiatives 357
Neelam Yadav, S.M. Paul Khurana, and Dinesh K. Yadav

Phenomics: Technologies and Applications in Plant and Agriculture 385
Hifzur Rahman, Valarmathi Ramanathan, N. Jagadeeshselvam, Sasikala Ramasamy, Sathishraj Rajendran, Mahendran Ramachandran, Pamidimarri D.V.N. Sudheer, Sushma Chauhan, Senthil Natesan, and Raveendran Muthurajan

Plant Cytomics: Novel Methods to View Molecules on the Move 413
Eric Davies and Bratislav Stankovic

Plant Physiomics: Photoelectrochemical and Molecular Retrograde Signalling in Plant Acclimatory and Defence Responses 439
Magdalena Szechyńska-Hebda, Paweł Budiak, Piotr Gawroński, Magdalena Górecka, Milena Kulasek, and Stanisław Karpiński

Signalomics: Diversity and Methods of Analysis of Systemic Signals in Plants 459
Alain Vian, Bratislav Stankovic, and Eric Davies

Thiolomics: Molecular Mechanisms of Thiol-Cascade in Plant Growth and Nutrition 491
Dibyendu Talukdar and Tulika Talukdar

Chloroplast Omics 533
L.A. de Luna-Valdez, P. León-Mejía, S. Encarnación-Guevara, and A.A. Guevara-García

Transplastomics: A Convergence of Genomics and Biotechnology 559
Muhammad Sarwar Khan

Plant Mitochondrial Omics: State-of-the-Art Knowledge 573
Mustafa Malik Ghulam, Sumaira Kousar, and Harsh Vardhan

Micromorphomics: A Morphological Dissection to Unveil Environmental Stress 615
Tulika Talukdar

Microbiomics: An Approach to Community Microbiology 633
Pankaj Sharma, Vijaya Brahma, Anamika Sharma, R.K. Dubey, G.S. Sidhu, and P.K. Malhotra

Cryobionomics: Evaluating the Concept in Plant Cryopreservation 655
Marcos E. Martinez-Montero and Keith Harding

Nanobiotechnology in Agricultural Development 683
Saleha Resham, Maria Khalid, and Alvina Gul Kazi

Plant Pharmacogenomics: From Drug Discovery to Personalized Ethnomedicine 699
Mustafeez Mujtaba Babar, Najam us Sahar Sadaf Zaidi, and Alvina Gul Kazi

Machine Learning Techniques in Plant Biology 731
Khwaja Osama, Bhartendu Nath Mishra, and Pallavi Somvanshi

Applications of Bioinformatics in Plant and Agriculture 755
M.A. Iquebal, Sarika Jaiswal, C.S. Mukhopadhyay, Chiranjib Sarkar, Anil Rai, and Dinesh Kumar

Plant Systems Biology: Insights and Advancements 791
Tulika Bhardwaj and Pallavi Somvanshi

Plantomics and Futuromics 821
Eric Davies and Debmalya Barh

Contributors

Lalit Agrawal, Ph.D. National Botanical Research Institute, Lucknow, India

Pawan Kumar Agrawal, Ph.D. Vivekananda Parvatiya Krishi Anusandhan Sansthan (VPKAS), Almora, India

Mustafeez Mujtaba Babar, Ph.D. Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan

B. Kalyana Babu, Ph.D. Vivekananda Parvatiya Krishi Anusandhan Sansthan (VPKAS), Almora, India

Debmalya Barh, Ph.D. Department of Genomics, Institute of Integrative Omics and Applied Biotechnology (IIOAB), Nonakuri, West Bengal, India

Tulika Bhardwaj, Ph.D. Department of Biotechnology, Teri University, New Delhi, India

Jayant S. Bhat, Ph.D. Regional Research Centre, Indian Agriculture Research Institute, Dharwad, Karnataka, India

N. Manikanda Boopathi, Ph.D. Department of Plant Molecular Biology and Bioinformatics, Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Vijaya Brahma, Ph.D. Toronto General Research Institute, University Health Network, Toronto, ON, Canada

Paweł Budiak, Ph.D. Department of Genetics, Breeding and Biotechnology, Warsaw University of Life Sciences, Warszawa, Poland

Sushma Chauhan, Ph.D. Department of Chemical and Biochemical Engineering, Dongguk University, Seoul, South Korea

Daniel Cozzolino, Ph.D. School of Agriculture, Food and Wine, The University of Adelaide, Osmond, SA, Australia

Eric Davies, Ph.D. Department of Plant Biology, North Carolina State University, Raleigh, NC, USA

L.A. de Luna-Valdez, Ph.D. Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

R.K. Dubey, Ph.D. Department of Floriculture and Landscaping, Punjab Agricultural University, Ludhiana, Punjab, India

S. Encarnación-Guevara, Ph.D. Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

Alberto Fassio, Ph.D. National Institute for Agricultural Research, INIA La Estanzuela, Colonia, Uruguay

Piotr Gawroński, Ph.D. Department of Genetics, Breeding and Biotechnology, Warsaw University of Life Sciences, Warszawa, Poland

Mustafa Malik Ghulam, Ph.D. Agriculture Biotechnology, National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan

Magdalena Górecka, Ph.D. Department of Genetics, Breeding and Biotechnology, Warsaw University of Life Sciences, Warszawa, Poland

A.A. Guevara-García, Ph.D. Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

Deepti Bhushan Gupta, Ph.D. TERI University, Vasant Kunj, New Delhi, India

Keith Harding, Ph.D. Damar Research Scientists, Damar, Cuparmuir, Fife, Scotland, UK

Firoz Hossain, Ph.D. Division of Genetics, Indian Agriculture Research Institute, New Delhi, India

M.A. Iquebal, Ph.D. Centre for Agricultural Bioinformatics, Indian Agricultural Statistics Research Institute, New Delhi, India

N. Jagadeeshselvam, Ph.D. Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Kaliaperumal Jagatheesh, Ph.D. Department of Biotechnology, Periyar University, Salem, India

Sarika Jaiswal, Ph.D. Centre for Agricultural Bioinformatics, Indian Agricultural Statistics Research Institute, New Delhi, India

Johnravindar D, Ph.D. Department of Biotechnology, Periyar University, Salem, India

Richard L. Jayaraj, Ph.D. Department of Biotechnology, Periyar University, Salem, India

Uday Chand Jha, Ph.D. Indian Institute of Pulses Research (IIPR), Kanpur, India

Stanisław Karpiński, Ph.D. Department of Genetics, Breeding and Biotechnology, Warsaw University of Life Sciences, Warszawa, Poland

Alvina Gul Kazi, Ph.D. Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan

Maria Khalid, Ph.D. Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan

Abdul Rehman Khan, Ph.D. Department of Environmental Sciences, COMSATS I.I.T., Abbottabad, Pakistan

Muhammad Sarwar Khan, Ph.D. Centre for Agricultural Biochemistry and Biotechnology (CABB), University of Agriculture, Faisalabad, Pakistan

S.M. Paul Khurana, Ph.D. Amity Institute of Biotechnology, Amity University Harayana, Gurgaon, Haryana, India

Sumaira Kousar, Ph.D. Industrial Biotechnology, National Institute for Biotechnology and Genetic Engineering (NIBGE), Faisalabad, Pakistan

R. Kowsalya, Ph.D. Department of Biotechnology, Periyar University, Salem, India

Milena Kulasek, Ph.D. Department of Genetics, Breeding and Biotechnology, Warsaw University of Life Sciences, Warszawa, Poland

Dinesh Kumar, Ph.D. Centre for Agricultural Bioinformatics, Indian Agricultural Statistics Research Institute, New Delhi, India

Amit K. Kushwaha, Ph.D. Department of Metabolic and Structural Biology, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

P. León-Mejía, Ph.D. Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico

P.K. Malhotra, Ph.D. Schools of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, Punjab, India

K. Manigandan, Ph.D. Department of Biotechnology, Periyar University, Salem, India

Marcos E. Martinez-Montero, Ph.D. Bioplantas Center, Plant Breeding Laboratory, University of Ciego de Ávila, Ciego de Avila, Cuba

Bhartendu Nath Mishra, Ph.D. Department of Biotechnology, Institute of Engineering and Technology, G.B. Technical University, Lucknow, India

Siddhartha Kumar Mishra, Ph.D. School of Biological Sciences, Dr. Harisingh Gour Central University, Sagar, India

C.S. Mukhopadhyay, Ph.D. School of Animal Biotechnology, Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana, Punjab, India

Raveendran Muthurajan, Ph.D. Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Elangovan Namasivayam, Ph.D. Department of Biotechnology, Periyar University, Salem, India

Senthil Natesan, Ph.D. Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Khwaja Osama, Ph.D. Department of Biotechnology, Institute of Engineering and Technology, G.B. Technical University, Lucknow, India

Pavan Kumar Padarthi, Ph.D. Department of Biotechnology, Periyar University, Salem, India

Basavanagouda S. Patil, Ph.D. Regional Research Centre, Indian Agriculture Research Institute, Dharwad, Karnataka, India

Hifzur Rahman, Ph.D. Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Anil Rai, Ph.D. Centre for Agricultural Bioinformatics, Indian Agricultural Statistics Research Institute, New Delhi, India

Sathishraj Rajendran, Ph.D. Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Mahendran Ramachandran, Ph.D. Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Valarmathi Ramanathan, Ph.D. Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Sasikala Ramasamy, Ph.D. Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

Saleha Resham, Ph.D. Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan

Ernesto Restaino, Ph.D. National Institute for Agricultural Research, INIA La Estanzuela, Colonia, Uruguay

Navinder Saini, Ph.D. Vivekananda Parvatiya Krishi Anusandhan Sansthan (VPKAS), Almora, India

Neelam S. Sangwan, Ph.D. Department of Metabolic and Structural Biology, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

Rajender Singh Sangwan, Ph.D. Department of Metabolic and Structural Biology, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

Center for Innovative and Applied Bioprocessing (formerly BioProcessing Unit), (An Autonomous Institute under Department of Biotechnology, Govt. of India), Mohali, Punjab, India

Chiranjib Sarkar, Ph.D. Centre for Agricultural Bioinformatics, Indian Agricultural Statistics Research Institute, New Delhi, India

Sarfraz Shafiq, Ph.D. Department of Environmental Sciences, COMSATS I.I.T., Abbottabad, Pakistan

Anamika Sharma, Ph.D. Department of Dental Microbiology, Theerthankar Mahaveer University, Moradabad, Uttar Pradesh, India

Pankaj Sharma, Ph.D. Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, Punjab, India

Shubhendu Shekhar, Ph.D. National Institute of Plant Genome Research, New Delhi, India

G.S. Sidhu, Ph.D. Schools of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, Punjab, India

Andrey Sinjushin, Ph.D. Department of Genetics, M.V. Lomonosov Moscow State University, Moscow, Russia

Pallavi Somvanshi, Ph.D. Department of Biotechnology, Teri University, New Delhi, India

Bratislav Stankovic, Ph.D. University for Information Science and Technology "St. Paul the Apostle", Ohrid, Republic of Macedonia

Pamidimarri D.V.N. Sudheer, Ph.D. Department of Chemical and Biochemical Engineering, Dongguk University, Seoul, South Korea

Magdalena Szechyńska-Hebda, Ph.D. Department of Genetics, Breeding and Biotechnology, Warsaw University of Life Sciences, Warszawa, Poland

The Franciszek Górski Institute of Plant Physiology, Polish Academy of Sciences, Kraków, Poland

Dibyendu Talukdar, Ph.D. Department of Botany, R.P.M. College, University of Calcutta, Uttarpara, Hooghly, West Bengal, India

Tulika Talukdar, Ph.D. Department of Botany, Krishnagar Government College, Nadia, Krishnagar, West Bengal, India

Department of Botany, APC Roy Govt. College, Siliguri, Darjeeling, West Bengal, India

Pragya Tiwari, Ph.D. Department of Metabolic and Structural Biology, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

Basant K. Tiwary, Ph.D. Centre for Bioinformatics, Pondicherry University, Pondicherry, India

Swati Tripathi, Ph.D. Vegetable Research Division, National Institute of Horticultural and Herbal Science, Rural Development Administration, Suwon, Republic of Korea

Najam us Sahar Sadaf Zaidi, Ph.D. Atta-ur-Rahman School of Applied Biosciences (ASAB), National University of Sciences and Technology (NUST), Islamabad, Pakistan

Harsh Vardhan, Ph.D. Faculty of Medicine, Department of Immunology/ Service of pneumology, University of Sherbrooke, Sherbrooke, QC, Canada

Alain Vian, Ph.D. UMR 1345 IRHS (Université d'Angers, Agrocampus Ouest, INRA), SFR 4207 Quasav, Faculté des Sciences, Angers cedex 01, France

Esteban Vicente, Ph.D. National Institute for Agricultural Research, INIA Salto Grande, Salto, Uruguay

Dinesh K. Yadav, Ph.D. Amity Institute of Biotechnology, Amity University Harayana, Gurgaon, Haryana, India

Neelam Yadav, Ph.D. Amity Institute of Biotechnology, Amity University Harayana, Gurgaon, Haryana, India

Ritesh K. Yadav, Ph.D. Department of Metabolic and Structural Biology, CSIR-Central Institute of Medicinal and Aromatic Plants, Lucknow, India

Sarika Yadav, Ph.D. Amity Institute of Biotechnology, Amity University Harayana, Gurgaon, Haryana, India

About the Editors

Debmalya Barh (M.Sc., M.Tech., M.Phil., Ph.D., PGDM) is a well-known molecular biotechnologist who is an expert in integrative *omics*–based cutting-edge R&D. He is the founder of the Institute of Integrative Omics and Applied Biotechnology (IIOAB) – a first-of-its-kind research organization in India that provides a global platform for multidisciplinary research and advocacy. He works with 400+ well-regarded researchers from 40+ countries and has 150+ high-impact international publications, several books, and book chapters in the field of biomedical and agricultural omics. Dr. Barh is a globally branded editor for editing omics related research reference books from Springer and Taylor & Francis and an editorial and review board member for a number of highly reputed international journals. Due to his significant contributions in the field and in promoting applied biological, agricultural, and bio-medical sciences using unique research strategies, in the year 2010 he has been recognized by *Who's Who in the World* and in 2014 he has been entered into the *Limca Book of Records,* "the Indian equivalent to the Guinness Book of World Records".

Muhammad Sarwar Khan, a Doctorate from the University of Cambridge, is a highly regarded Molecular Biologist from Pakistan who was a Founding Head of Biotech Interdisciplinary Division at NIBGE and is currently serving as the Director of Center of Agricultural Biochemistry and Biotechnology (CABB), University of Agriculture, Faisalabad, Pakistan. He has several awards to his credit, including Civil Award (Presidential Medal for Technology), Gold Medal in Agriculture (Pakistan Academy of Sciences), Performance Gold Medal (Pakistan Atomic Energy Commission), and Biotechnologist Award (National Commission on Biotechnology), and is a Life Fellow of Cambridge Commonwealth Society. He has several high-impact publications in scientific journals including *Nature* and *Nature Biotechnology*, and is an author of a number of books and book chapters. Dr. Khan has made colossal contributions in the field of chloroplast genetic engineering and is pioneer in expressing GFP in plant chloroplasts, developing transplastomic rice and sugarcane. His research focus also includes expression of foreign genes in the chloroplasts to confer agronomic traits such as insect-pest resistance, salinity and herbicide tolerance, and overexpression of antigenic and therapeutic proteins in chloroplasts to develop cost-effective therapeutics and vaccines.

Eric Davies, Ph.D. is currently an Emeritus Professor in the Department of Plant and Microbial Biology, North Carolina State University (NCSU), USA. He is the former Head of the Botany Department at NCSU and the Director of the NASA Specialized Center of Research and Training (NSCORT) in Gravitational Biology. Professor Davies was also a former Visiting Professor at Purdue University, University of Arizona, Ehime University in Japan, University of Brussels in Belgium, Universite Blaise Pascal and Universite d'Angers in France. He is an Honorary Professor at the University of Warmia and Mazuryi in Poland, and has an Honorary Doctorate from Universite Blaise Pascal. He has several refereed publications, book chapters, and books. Professor Davies is an expert in plant wound signaling, especially the rapidly-transmitted physical (turgor and electrical) signals, local changes in ions and hormones, and ultra-rapid, systemic transcript accumulation and degradation. His other research areas are the role of the cytoskeleton in mRNAs attachment to enhance translation in specific sub-cellular regions of the cell; gravity regulated cell signaling and gene expression, especially the response in the maize pulvinus; microwave irradiation; and gene expression in plants.

Omics of Model Plants

Pawan Kumar Agrawal, B. Kalyana Babu,
and Navinder Saini

Contents

Introduction 2

Arabidopsis thaliana 2
Genomics 2
Functional Genomics 4
Proteomics 5
Metabolomics 6

Rice 8
Genomics and Synteny 9
Genetic and Physical Mapping 9
BAC Library 9
Synteny 10
Gene Expression 10
QTL Mapping and Tagging of Useful Genes 11
Marker-Assisted Selection 15
Next-Generation Sequencing 16
Expressed Sequence Tags (ESTs)
and Flanking Sequence Tags (FSTs) 17
Functional Genomics 17
Transgenic for Genomic Studies 18
T-DNA Insertional Mutagenesis
for Functional Genomics in Rice 18
Proteomics 19
Abiotic Stresses 19
Biotic Stresses 20
Metabolomics 20

Maize 22
Genomics 22
Functional Genomics 23
Proteomics 24
Metabolomics 24

Conclusions 25

References 25

Abstract

The multiple omics tools and strategies like high-throughput genome-scale genotyping platforms such as whole-genome re-sequencing, proteomics, and metabolomics provide greater opportunities to dissect molecular mechanisms and the discovery of key genes in developing ideal genotypes in the changing climate scenario. The last decade has seen rapid advances in functional genomic research globally. Most of the efforts involve construction of technological and resource platforms for high-throughput DNA sequencing, gene identification, and physical and genetic mapping; functional analysis of genomes for agronomic traits and biological processes; and identification and isolation of functional genes. The functional genomic research aims to understand how the genome functions at the whole-genome level, whereas proteomics looks for the systematic analysis of the protein population in a tissue, cell, or subcellular compartment. Metabolites are the end products of cellular process, and they show the response of biological systems to environmental changes. The current trend in metabolomic

P.K. Agrawal, Ph.D. (✉) • B.K. Babu, Ph.D.
N. Saini, Ph.D.
Vivekananda Parvatiya Krishi Anusandhan
Sansthan (VPKAS), Almora, India
e-mail: pawankagrawal@hotmail.com

D. Barh et al. (eds.), *PlantOmics: The Omics of Plant Science*,
DOI 10.1007/978-81-322-2172-2_1, © Springer India 2015

studies is to define the cellular status at a particular time point of development or physiological status. These techniques complement other techniques such as transcriptomics and proteomics and depict precise pictures of the whole cellular process. The growing number of sequenced plant genomes has opened up immense opportunities to study biological processes related to physiology, growth and development, and tolerance to biotic and abiotic stresses at the cellular and whole plant level using a novel systems-level approach. The "omics" approach integrates genome, proteome, transcriptome, and metabolome data into a single data set and can lead to the identification of unknown genes and their regulatory networks involved in metabolic pathways of interest. This will also help in understanding the genotype–phenotype relationship and consequently help to improve the quality and productivity of crop plants for the food and nutritional security of millions of human populations.

Keywords

Arabidopsis • Maize • Rice • *Zea mays* • *Oryza sativa* • Functional genomics • Phenomics • Transcriptomics • Proteomics • Metabolomics • MAS • QTL • Transgenics

Introduction

The last decade has seen rapid advances in functional genomic research globally. Most of the efforts involve construction of technological and resource platforms for high-throughput DNA sequencing, gene identification, and physical and genetic mapping; functional analysis of genomes for agronomic traits and biological processes; and identification and isolation of functional genes. The overall goal of functional genomic research is to understand how the genome functions at the whole-genome level. Similarly, proteomics looks for the systematic analysis of the protein population in a tissue, cell, or subcellular compartment. It enables correlations to be drawn between the range of proteins produced by a cell or tissue and the initiation or progression of a stress or normal metabolism. Metabolites are the end products of cellular process, and they show the response of biological systems to environmental changes. The current trend in metabolomic studies is to define the cellular status at a particular time point of development or physiological status. These techniques complement other techniques such as transcriptomics and proteomics and depict precise pictures of the whole cellular process. The growing number of sequenced plant genomes has opened up immense opportunities to study biological processes related to physiology, growth and development, and tolerance to biotic and abiotic stresses at the cellular and whole plant level using a novel systems-level approach. The "omics" approach integrates genome, proteome, transcriptome, and metabolome data into a single data set and can lead to the identification of unknown genes and their regulatory networks involved in metabolic pathways of interest. This will also help in understanding the genotype–phenotype relationship.

Arabidopsis thaliana

Genomics

Arabidopsis thaliana is an excellent model organism for the analysis of complex biological processes in plants using molecular and biotechnological techniques. The frontiers of plant science, like other branches of the life sciences, have been dominated by genomics over the past 25 years. Many research laboratories are currently putting intensive efforts to isolate *Arabidopsis* genes of biological importance using map-based cloning strategy. Although genetic linkage (Koornneef et al. 1983) and recombinant inbred line (RIL) maps (Lister and Dean 1993) have been reported, the construction of accurate physical maps of the chromosomes will be highly advantageous not only for the genomic sequencing but also for map-based gene cloning (Ward and Jen 1990). Hence, a complete physical map of the *Arabidopsis* genome should be greatly

advantageous for cloning the genetic loci of interest as well as sequencing the entire genome. Hence, yeast artificial chromosome (YAC)-based physical maps of chromosomes 2 (Zachgo et al. 1996) and 4 (Schmidt et al. 1995) of *A. thaliana* have been constructed by several workers. Based on the sizes of the YACs and their coverage of the chromosome, the length of chromosome 2 was estimated to be at least 18 Mb. Sato et al. (1998) presented physical map of the entire chromosome 3 which was constructed by ordering the clones from YAC, PI, TAC, and BAC libraries using the information from the sequences of various DNA markers and the terminal sequences of the clones. The sizes of the centromeric regions of *Arabidopsis thaliana* chromosomes 1, 2, and 3 were determined by construction of their physical maps on the basis of restriction analysis (Hosouchi et al. 2002). The sizes of the genetically defined centromeric regions were deduced to be 9 megabases (Mb), 4.2 Mb, and 4.1 Mb, respectively (chromosome 1, from markers T22C23-t7 to T3P8-sp6; chromosome 2, from F5J15-sp6 to T15D9; chromosome 3, from T9G9-sp6 to T15M14) (Copenhaver et al. 1998). Mitochondrial genomes in higher plants are characterized by their high flexibility and variation in size and structure. The mitochondrial genome of *A. thaliana* was physically mapped using cosmid and YAC clones and was found to contain 372 kb size which was relatively large (Klein et al. 1994). The presence of this comparatively large mitochondrial genome in a plant with one of the smallest nuclear genome showed that different size constraints act upon the different genomes in plant cells. *A. thaliana* is known to contain approximately 1,000 copies of 5S rDNA per haploid genome, and they occur in tandem arrays (Campell et al. 1992). The 5S ribosomal RNA genes were mapped to mitotic chromosomes of *Arabidopsis thaliana* by fluorescence in situ hybridization (FISH) by Murata et al. (1997).

Arabidopsis thaliana is widely used as a model for the study of many aspects of plant biology. Because of its small genome size (125 Mb), it was chosen as the subject of the first plant genome sequencing project, an effort that was completed. *Arabidopsis thaliana* was the first plant, and the third multicellular organism after *Caenorhabditis elegans* (The *C. elegans* Sequencing Consortium 1998) and *Drosophila melanogaster* (Adams et al. 2000), to be completely sequenced (The *Arabidopsis* Genome Initiative 2000). Since systematic sequencing was completed in late 2000, the genome sequence has undergone several rounds of reassembly, hole patching, and extension into un-sequenced regions. One of the major features of the *Arabidopsis* genome revealed by the genome sequence was the extent of gene duplication and segmental duplications, which was surprising given the expectation of a functionally compact genome. Approximately 60 % of the genome was thought to be derived from a single duplication event, possibly of the entire genome (The *Arabidopsis* Genome Initiative 2000). The extensive work carried out based on the *Arabidopsis* genome sequence also supports interpretations of the evolution of the vertebrate lineage that propose a central role for genome duplications (Wolfe 2001). Comparison of *Arabidopsis* sequences with genomic sequence from the closely related *Brassica oleracea* (Chinese cabbage) identified regions of high similarity that either identified putative new genes or extended existing gene models. About 30 % of these new genes encoded a transcript. About 25 % of the originally predicted genes had no supporting evidence such as an EST match or reasonable similarity of their putative peptide sequence to any other protein. The decreasing cost along with rapid progress in next-generation sequencing and related bioinformatics computing resources has facilitated large-scale discovery of SNPs in *Arabidopsis* species. Large numbers and genome-wide availability of SNPs make them the marker of choice in partially or completely sequenced genomes. The complete nucleotide sequence of the chloroplast genome of *Arabidopsis thaliana* has been determined (Sato et al. 1999). The genome as a circular DNA composed of 154,478 bp containing a pair of inverted repeats of 26,264 bp, which are separated by small and large single copy regions of 17,780 bp and 84,170 bp, respectively. Cao et al. (2011) presented the first phase of the project, based on population-scale sequenc-

ing of 80 strains of *A. thaliana* populations drawn from eight regions throughout the species' native range. They found common small-scale polymorphisms as well as many larger insertions and deletions in the *A. thaliana* pan-genome.

A major goal in evolutionary biology is to identify the genetic basis of adaptive trait variation. In the model plant species *Arabidopsis thaliana*, studies are now being performed exploiting natural variation as a powerful alternative to classical mutant genetics (Koornneef et al. 2004), in particular to identify genes underlying important quantitative trait variation. Benjamin et al. (2010) studied combined analysis of genome-wide association (GWA) study with traditional linkage mapping in order to detect the genetic bases underlying natural variation in flowering time in ecologically realistic conditions in the plant *Arabidopsis thaliana*. It involved phenotyping of nearly 20,000 plants over 2 winters under field conditions in a temperate climate. Simon et al. (2008) studied phenotyping of nearly 20,000 plants over 2 winters under field conditions, including 184 worldwide natural accessions genotyped for 216,509 SNPs and 4,366 RILs derived from 13 independent crosses chosen to maximize genetic and phenotypic diversity. The results showed that combined linkage and association mapping clearly outperforms each method alone when it comes to identifying true associations. Kuittinen et al. (1997) described a quantitative trait locus (QTL) mapping experiment for flowering time in *Arabidopsis*. Five to seven QTLs affecting flowering time were found in a BC_1 population derived from the Finnish Naantali genotype and the German strain Li-5. In a different population, consisting of 165 RILs, Alonso-Blanco et al. (1997) found four QTLs affecting the flowering time. Several loci exhibiting variation in complex traits (quantitative trait loci or QTLs) have been cloned. Examples include using linkage disequilibrium (LD) to fine map the *FRI* and *FLC* loci controlling flowering time (Hagenblad et al. 2004). Natural variation in hypocotyl responses to light was shown to be due to polymorphisms in phytochrome light receptors. Affymetrix expression arrays have also been used for genotyping; total genomic DNA from recombinant inbred lines (RILs) made from a cross of Col and Ler was hybridized to the ATH1 Affymetrix array, and recombination events were identified. Marker and QTL information obtained from a segregating population can be used for the design of efficient breeding strategies. Marker-assisted selection (MAS) has been advocated as a useful tool for rapid genetic advance in the case of quantitative traits (Lande and Thompson 1990; Knapp 1994, 1998). Berloo and Stam (1999) described an experiment using RILs of *A. thaliana* with an objective to compare an MAS breeding strategy, using molecular marker and QTL information, with conventional breeding methods, based on phenotype only. Selection based on marker and QTL information gave approximately the same result as selection based on phenotype. The relative high heritability of flowering time in *Arabidopsis* facilitated successful phenotypic selection. The difference in selection result that was anticipated to be in favor of the marker-assisted approach was therefore not observed.

Functional Genomics

With the availability of complete genome sequences of several organisms, the focus has shifted from structural genomics to functional genomics, specifically in plants where the complete genomic sequences are becoming available (*Arabidopsis* and rice). A variety of approaches are used to clone and gather information about the function(s) of gene(s). Among these, insertional mutagenesis has been extensively used for cloning genes, promoters, enhancers, and other regulatory sequences from *Arabidopsis*. Strategies used for cloning and characterization depend upon the information available about the gene or its product. Expressed sequence tags (ESTs) and microarray-based techniques are some of the powerful approaches in this direction.

A comprehensive molecular-marker-based linkage map exists for *Arabidopsis*, and the map-based cloning of genes conferring specific phenotypes will become even easier with the availability of genomic sequence information. Jun et al. (2011) conducted whole-genome