Swapna Thacheril Sukumaran
Shiburaj Sugathan
Sabu Abdulhameed *Editors*

Plant Metabolites: Methods, Applications and Prospects

Springer

Plant Metabolites: Methods, Applications and Prospects

Swapna Thacheril Sukumaran •
Shiburaj Sugathan • Sabu Abdulhameed
Editors

Plant Metabolites: Methods, Applications and Prospects

Springer

Editors
Swapna Thacheril Sukumaran
Department of Botany
University of Kerala
Thiruvananthapuram, Kerala, India

Shiburaj Sugathan
Division of Microbiology
KSCSTE—Jawaharlal Nehru Tropical Botanic Garden and Research Institute
Thiruvananthapuram, Kerala, India

Sabu Abdulhameed
Department of Biotechnology and Microbiology
Kannur University
Kannur, Kerala, India

ISBN 978-981-15-5135-2 ISBN 978-981-15-5136-9 (eBook)
https://doi.org/10.1007/978-981-15-5136-9

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

We proudly dedicate this book to Prof. (Dr.) M Haridas, professor emeritus and founder head of the Department of Biotechnology and Microbiology, Kannur University, Kerala, India. Prof. Haridas epitomises the triumph of inventiveness over all other settings, which make ones contribution less significant, a triumph of genius over circumstances. For decades, he has pursued and achieved excellence in both research and teaching. His achievements are powerful and lasting in their impact on a global scale. What makes them so inspiring to younger generations of scientists and teachers is that they are more necessary in the present world where science only could solve the problems of environment and human society.

Foreword

Despite the many triumphs of modern medicine, there remain many diseases for which effective treatments are still not available, or for which modern medicines are not affordable, especially for patients living in resource-poor areas. However, many parts of the world have century-old traditions of using herbal medicines to treat a variety of medical conditions, including infections, cancer, and immune disorders.

While many of these treatments have not been rigorously tested in controlled clinical trials, challenges in clinical development consist of incomplete knowledge on their active ingredients and mechanisms of action. In addition to containing compounds to treat specific diseases currently already targeted, plants can also harbor many more metabolites that, if properly identified, could have promise for novel medical applications, or can be the basis for further chemical modification to improve efficacy and reduce the risk of adverse effects.

Further success in this research area is contingent on progress in the isolation, purification, in vitro and in vivo characterization, and production of these plant metabolites. This book provides the latest updates from world experts, investigating a variety of plant metabolites aimed at preventing or treating medical conditions. As the reader will conclude, much progress is being made, and this exciting research opens the door to further explore unchartered territories in future years.

University of California, Davis
Davis, CA, USA

Koen Van Rompay

Preface

Plants by way of their metabolites are a hot topic of research for all over the world. Both the types of metabolites—primary metabolites (produced for the existence) and secondary metabolites (produced as a result of metabolic errors or produced as non-essentials and accumulated as by-products)—find enormous relevance in pharmaceutical, food, agriculture and other sectors with commercial potential of billions of dollars. Plants ranging from the primitive to the highly evolved groups and those which survive in unique and diverse habitats are treasures of high-value molecules. Though there are concerted efforts taking place across the globe, the sector is yet to take a full bloom.

The book entitled "Plant Metabolites: Methods, Applications and Prospects" is compiled and edited by scientists from Kerala and most of the contributors of the different chapters are also from Kerala, India. Kerala is well known for its rich biodiversity and it has about 700 km long track of Western Ghats, the enlisted biodiversity hotspot. The state pioneers in sustainable utilization of biodiversity through novel practices, processes and products. Hence, the contribution to knowledge through this edited volume is directly from the players with expertise and experience evolved with a background of years. The background of herbal material utilization in Kerala goes remote in unrecorded history of folklore and in recorded history to the seventeenth century Latin classic, Hortus Malabaricus.

The book comprises twenty-two chapters with state-of-the-art knowledge on sources of plant metabolites ranging from small and primitive plant species to fully evolved trees, methods for extraction, purification and characterization of metabolites, applications of metabolites in various sectors, microbial biotransformation of metabolites for preparation of nutraceuticals, existing traditional knowledge-based practices for plant-based therapy and advanced techniques including that of gene technology for the enhanced and sustainable production of metabolites. While we have taken all the possible efforts to make the book a perfect one, there may be lapses which we are ready to accept and improve in future. We thank all who have

contributed to this book and all the well-wishers for extending their support. We sincerely wish that the scientific fraternity across the globe will make use of the information presented in this book published under the banner of Springer-Nature.

Thiruvananthapuram, Kerala, India — Swapna Thacheril Sukumaran
Thiruvananthapuram, Kerala, India — Shiburaj Sugathan
Kannur, Kerala, India — Sabu Abdulhameed

Contents

Editors and Contributors

About the Editors

Swapna Thacheril Sukumaran is a Professor at the Department of Botany, University of Kerala. She completed her PhD at the Department of Biotechnology, Cochin University of Science and Technology, India. Dr. Swapna has 25 years of teaching and research experience and has published nine books and over 100 research papers in journals and conference proceedings. Her research interests include *in vitro* secondary metabolite production, medicinal plant conservation and phytochemistry. Dr. Swapna is a recipient of research grants from the Department of Environment and Climate Change, Government of Kerala, Western Ghats Development Cell, Government of Kerala, University Grants Commission, Government of India and SERB, Department of Science and Technology, Government of India. She is currently serving as a Member of the Kerala State Biodiversity Board.

Shiburaj Sugathan has a background in Plant Sciences and has been working in Microbial Bioprospecting since 1995. He is currently a Senior Scientist and Head of the Division of Microbiology at Jawaharlal Nehru Tropical Botanic Garden and Research Institute, Thiruvananthapuram, India. He has over 25 years of experience in the area of microbial biotechnology and has published several papers in reputed national and international journals, three books, several book chapters and one Indian patent. Dr. Shiburaj received his PhD degree from the University of Kerala and completed his postdoctoral studies at Madurai Kamaraj University, India and Dalhousie

University, Canada. He is a recipient of a BOYCAST fellowship from the Department of Science & Technology, Government of India. Dr. Shiburaj is also a Visiting Professor at the University of Coahulia, Mexico. His current research interests include recombinant expression and scale up of industrial enzymes from Actinobacteria, and the characterization of antibiofilm molecules.

Sabu Abdulhameed is currently working as an Associate Professor at the Department of Biotechnology and Microbiology, Kannur University, India. He received his PhD from Cochin University of Science and Technology and joined the Biotechnology Division of CSIR-NIIST, Thiruvananthapuram as a Scientist Fellow. He has engaged in post-doctoral research work in microbial fermentation and biotransformation in the USA and France. His research interests are in microbial fermentation, therapeutic enzymes and the biotransformation of bioactive molecules. He has contributed extensively to fermentation and biotransformation in medicated wines and has three patents to his credit. Dr. Sabu has published five books and 70 research papers in international journals. He is a recipient of a postdoctoral fellowship from the Institute for Research and Development (IRD, France) and a Visiting Professorship from the Mexican Agency for International Cooperation and Development. He is also a member of many academic bodies and expert panels, and serves on the editorial boards of numerous research journals.

Contributors

M. Amruth Forestry and Human Dimension-Programme Division, KSCSTE—Kerala Forest Research Institute, Thrissur, Kerala, India

A. L. Aneeshkumar Ethnomedicine & Ethnopharmacology Division, KSCSTE—Jawaharlal Nehru Tropical Botanic Garden and Research Institute, Thiruvananthapuram, Kerala, India

R. Aswathi Nair Department of Biochemistry & Molecular Biology, Central University of Kerala, Kasaragod, Kerala, India

V. S. Athira Department of Botany, University of Kerala, Karyavattom Campus, Thiruvananthapuram, Kerala, India

Juan Roberto Banevente-Valdés Department of Chemical Engineering, Universidad Autónoma de Coahuila, Saltillo, Coahuila, México

Hari Bharathan Department of Zoology, Sree Narayana College, Kollam, Kerala, India

James P. Bhaskar ITC Life Sciences and Technology Centre, ITC Limited, Bangalore, Karnataka, India

M. Deepti Department of Biotechnology, University of Calicut, Malappuram, Kerala, India

K. K. Elyas Department of Biotechnology, University of Calicut, Malappuram, Kerala, India

Gayathri Valsala Division of Microbiology, KSCSTE—Jawaharlal Nehru Tropical Botanic Garden and Research Institute, Thiruvananthapuram, Kerala, India

Biju Haridas Division of Microbiology, KSCSTE—Jawaharlal Nehru Tropical Botanic Garden and Research Institute, Thiruvananthapuram, Kerala, India

M. Haridas Inter University Centre for Bioscience and Department of Biotechnology & Microbiology, Dr E K Janaki Ammal Campus, Kannur University, Kannur, Kerala, India

Shabeer Ali Hassan Mohammed Division of Molecular Parasitology & Immunology, CSIR-CDRI, Lucknow, Uttar Pradesh, India

T. K. Hrideek Department of Forest Genetics and Tree Breeding, KSCSTE—Kerala Forest Research Institute, Thrissur, Kerala, India

Anju V. Jalaj Department of Botany, University of Kerala, Thiruvananthapuram, Kerala, India

Sony Jayaraman Department of Biotechnology & Microbiology, Kannur University, Kannur, Kerala, India

R. G. Kala Advanced Centre for Molecular Biology and Biotechnology, Rubber Research Institute of India, Kottayam, Kerala, India

M. V. Lakshmi Department of Botany, University of Kerala, Thiruvananthapuram, Kerala, India

N. Lini Department of Biotechnology, Mar Ivanios College, Thiruvananthapuram, Kerala, India

Tom H. MacRae Department of Biology, Dalhousie University, Halifax, NS, Canada

Julio Montañez Department of Chemical Engineering, Universidad Autónoma de Coahuila, Saltillo, Coahuila, México

Lourdes Morales-Oyervides Department of Chemical Engineering, Universidad Autónoma de Coahuila, Saltillo, Coahuila, México

Sanjeeva Nayaka Lichenology Laboratory, CSIR—National Botanical Research Institute, Lucknow, Uttar Pradesh, India

O. Nikhita Surendran Department of Biotechnology & Microbiology, Inter University Centre for Bioscience, Kannur University, Dr E K Janaki Ammal Campus, Kannur, Kerala, India

Shunmugiah Karutha Pandian Department of Biotechnology, Science Campus, Alagappa University, Karaikudi, Tamil Nadu, India

Padmesh P. Pillai Department of Genomic Science, Central University of Kerala, Kasaragod, Kerala, India

R. Prakashkumar Ethnomedicine & Ethnopharmacology Division, KSCSTE—Jawaharlal Nehru Tropical Botanic Garden and Research Institute, Thiruvananthapuram, Kerala, India

Arumugam Priya Department of Biotechnology, Science Campus, Alagappa University, Karaikudi, Tamil Nadu, India

A. V. Raghu Kerala Forest Research Institute, Thrissur, Kerala, India

Resmi Ravindran Department of Pathology & Laboratory Medicine, University of California Davis Medical Center, Sacramento, CA, USA

K. P. Rini Vijayan Kerala Forest Research Institute, Thrissur, Kerala, India

Raúl Rodríguez-Herrera Food Research Department, Faculty of Chemistry, Autonomous University of Coahuila, Saltillo, Coahuila, Mexico

A. Sabu Inter University Centre for Bioscience, Department of Biotechnology and Microbiology, Dr E K Janaki Ammal Campus, Kannur University, Kannur, Kerala, India

S. R. Saranya Krishnan Department of Botany, University of Kerala, Thiruvananthapuram, Kerala, India

Praseetha Sarath Division of Microbiology, KSCSTE—Jawaharlal Nehru Tropical Botanic Garden and Research Institute, Thiruvananthapuram, Kerala, India

Lekshmi Sathyaseelan Department of Botany, University of Kerala, Thiruvananthapuram, Kerala, India

C. S. Sharanya Inter University Centre for Bioscience, Department of Biotechnology & Microbiology, Dr E K Janaki Ammal Campus, Kannur University, Kannur, Kerala, India

Crystel Aleyvick Sierra-Rivera Laboratory of Immunology and Toxicology, Faculty of Chemistry, Autonomous University of Coahuila, Saltillo, Coahuila, Mexico

E. A. Siril Department of Botany, University of Kerala, Thiruvananthapuram, Kerala, India

K. Sreejith Department of Biotechnology and Microbiology, Dr E K Janaki Ammal Campus, Kannur University, Kannur, Kerala, India

R. Sreelekshmi Department of Botany, University of Kerala, Thiruvananthapuram, Kerala, India

Suby Department of Forest Genetics and Tree Breeding, KSCSTE—Kerala Forest Research Institute, Thrissur, Kerala, India

Shiburaj Sugathan Division of Microbiology, KSCSTE—Jawaharlal Nehru Tropical Botanic Garden and Research Institute, Thiruvananthapuram, Kerala, India

S. R. Suja Ethnomedicine & Ethnopharmacology Division, KSCSTE—Jawaharlal Nehru Tropical Botanic Garden and Research Institute, Thiruvananthapuram, Kerala, India

Swapna Thacheril Sukumaran Department of Botany, University of Kerala, Thiruvananthapuram, Kerala, India

R. Supriya Advanced Centre for Molecular Biology and Biotechnology, Rubber Research Institute of India, Kottayam, Kerala, India

Karuna Surendran Department of Genomic Science, Central University of Kerala, Kasaragod, Kerala, India

M. K. Suresh Kumar Laboratory of Genetics and Genomics, National Cancer Institute—NIH, Bethesda, MD, USA

Ramesh Surianarayanan R&D Consultant, Chennai, Tamil Nadu, India

P. H. Surya Department of Biotechnology, University of Calicut, Malappuram, Kerala, India

Thirukannamangai Krishnan Swetha Department of Biotechnology, Science Campus, Alagappa University, Karaikudi, Tamil Nadu, India

George Szakacs Department of Applied Biotechnology and Food Science, Budapest University of Technology and Economics, Budapest, Hungary

Riyas Chakkinga Thodi Department of Botany, University of Kerala, Thiruvananthapuram, Kerala, India

A. Thulaseedharan Advanced Centre for Molecular Biology and Biotechnology, Rubber Research Institute of India, Kottayam, India

Renu Tripathi Division of Molecular Parasitology & Immunology, CSIR-CDRI, Lucknow, Uttar Pradesh, India

Jayadevi Variyar Department of Biotechnology & Microbiology, Kannur University, Dr. E K Janaki Ammal Campus, Kannur, Kerala, India

S. Vivek Department of Botany, University of Kerala, Thiruvananthapuram, Kerala, India

Alenjandro Zugasti-Cruz Laboratory of Immunology and Toxicology, Faculty of Chemistry, Autonomous University of Coahuila, Venustiano Carranza Blvd. and Jose Cardenas Valdes Street, Saltillo, Coahuila, Mexico

Zyju Damodharan Pillai Padmini Amma ThermoFisher Scientific, Dubai, UAE

Plant Metabolomics: Current Status and Prospects

1

C. S. Sharanya, A. Sabu, and M. Haridas

Abstract

Plant metabolomics deals with the interpretation of various metabolic pathways in contrast to other -omics technologies applied in systems biology. Metabolomics is a highly challenging field where the metabolite analysis is done by high-end technologies for proposing metabolic pathways. The high-throughput technologies utilized for these studies importantly include mass and nuclear magnetic resonance spectrometry. Both these techniques have their own specific features and it is usually difficult to interpret the data compared to genomics and transcriptomics data. The metabolic pathways in plants are changed during biotic and abiotic stresses. These changes can be noted at each step through metabolite analysis. Plants acclimatize to the changes during stress conditions by producing secondary metabolites by other mechanisms. We may analyse these changes by advanced techniques.

Keywords

Genomics · Transcriptomics · Proteomics · Metabolomics · GC MS · LC-MS

1.1 Introduction

In an era of developed high-throughput genomics (DNA sequencing), transcriptomics (gene expression analysis) and proteomics (protein analysis) metabolomics is the foremost of the '-omics' approaches emerging from the metabolic profiling. Thus, the significant results of all these technologies combine to become systems biology. The interactions among different organisms in the

C. S. Sharanya · A. Sabu · M. Haridas (✉)
Department of Biotechnology and Microbiology, Inter University Centre for Bioscience, Kannur University, Thalassery, Kerala, India

S. T. Sukumaran et al. (eds.), *Plant Metabolites: Methods, Applications and Prospects*, https://doi.org/10.1007/978-981-15-5136-9_1

environment are through natural products. This is particularly important in the communication among the members of the same species. Though metabolomics and natural products discovery evolved independently, they have a great amount of structure-function in common and to share. Though these two fields have different origins historically, they have overlapping objectives. The convergence of metabolomics and natural product discovery has been greatly enhanced by the databases devoted to structural parameters, particularly small-molecule databases. Metabolomics is a highly challenging field for the essentially inclusive, nonbiased, high-throughput screening of complicated metabolites present in the plant extracts used as traditional medicine. Furthermore, we can consider it as a complex framework built of all these emerging technologies for a whole, systems biology (Fig. 1.1).

1.1.1 Genomics

The first -omics phase is genomics, which is of the entire genome sequencing of an organism. Thus genome can be defined as a complete set of genes inside a cell. Once the sequencing of entire genome is completed, this sequence can be used to analyse functions of the genes (functional genomics), compare the genome with another genome (comparative genomics) or to generate 3D structure of proteins in a protein family, thus hinting to their function (structural genomics). DNA is the basis of every genome and its further processing leads to the molecular machinery. Prior to the many post-translational modifications, DNA is transcribed first into RNA and then translated into protein. In the case of plant species, the genetic background ensures its production and resistance to unfavourable conditions. It also ensures the scope of improving production and the potential of resistance to environmental risks. We can repair the dysfunctions and can generate a healthy condition by examining the genetic information of any deleterious mutations, insertions or deletions from the normal state. There may be many genetic variants present in an organism of either benign or protective in nature, having an advantage against the condition of no variants. So, the variants are classified as simple, nucleotide variations and structural variations. Single nucleotide variations (SNv) have single insertions or deletions whereas structural variations have large copy number variation and inversions. Moreover, we can analyse these variations by the relatively recent, advanced methods in the area of microarray technology. Microarrays measure the differences in the DNA sequence of individual plants and analyse the expression of thousands of genes at a time. This reveals the abnormalities in chromosomes by a process called comparative genomic hybridization. In the case of agriculture, the significant genes are identified for the nutritional purpose, production and showing relevance in resistance against pathogens (Horgan and Kenny 2011).

The first plant genome sequenced was that of *Arabidopsis thaliana* (Arabidopsis Genome Initiative 2000) and the first crop genome was of rice (Eckardt 2000). The International Rice Genome Sequencing Project (IRGSP) was initiated in 1997 at a workshop held in Singapore. Many countries from all over the world participated in the endeavour and guidelines were proposed (Table 1.1).

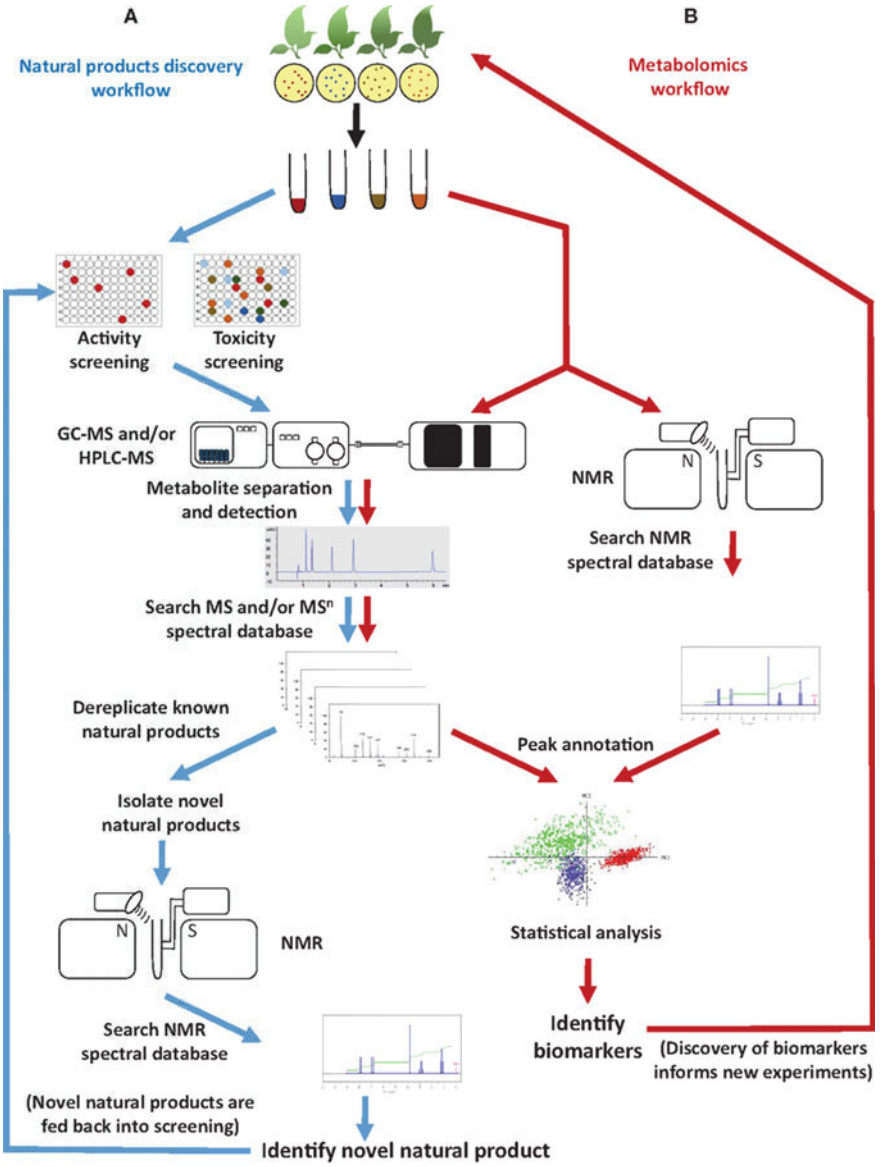

Fig. 1.1 Typical metabolomics work flow (Reproduced from Johnson and Lange 2015, under creative commons license https://doi.org/10.3389/fbioe.2015.00022)

Each country started working on different chromosomes, and finally the complete genome was published in the year 2000. Rice genome has a size of 400–430 million base pairs which is three times higher than *Arabidopsis thaliana* with a well-mapped genome, having 6000 markers and 40,000 expressed sequence tags, making rice a

Table 1.1 Countries involved in the International Rice Genome Sequencing Project

Research Institute	Chromosome
Rice Genome Research Program (RGP; Japan)	1, 6, 7, 8
Korea Rice Genome Research Program (Korea)	1
CCW (United States) CUGI (Clemson University) Cold Spring Harbor Laboratory Washington University Genome Sequencing Center	3, 10
TIGR (United States)	3, 10
PGIR (United States)	10
University of Wisconsin (United States)	11
National Center for Gene Research Chinese Academy of Sciences (China)	4
Indian Rice Genome Program (University of Delhi)	11
Academia Sinica Plant Genome Center (Taiwan)	5
Genoscope (France)	12
Universidad Federal de Pelotas (Brazil)	12
Kasetsart University (Thailand)	9
McGill University (Canada)	9
John Innes Centre (United Kingdom)	2

good candidate for sequencing. The IRGSP followed the strategy of map-based, clone-by-clone shotgun strategy for sequencing. This method was found to be efficient for exploiting the full potential of rice sequence and bacterial artificial chromosome libraries constructed from 'Nipponbare', a variety of *Oryza sativa* ssp. *japonica*. Molecular marker technology allows the identification of a gene that controls specific trait by genome having all genes, themselves acting as a marker and the process becomes more correct and effective. Hence, it becomes more significant while making allele-specific markers for the successive PCR reactions. Thus, end-sequencing, fingerprinting and marker-aided PCR screening are used to make sequence-ready contigs. After the completion of the project, all the data and annotated sequences are made publicly available at each IRGSP member's website. They maintain their sequence data with high quality and available publicly within a short period of time (Sasaki and Burr 2000).

1.1.2 Transcriptomics

Genome is the storage centre of all the biological data. However, it cannot transfer the data to cells directly, and for the transfer, a series of biochemical reactions are carried out with the help of enzymes and proteins. The entire process is referred to as gene expression. The sum of the results of genome expression is the transcriptome. The whole collection of transcripts in a species speaks to a key connection between the DNA and the phenotype whose biological data is required by the cell at a specific time. These RNA molecules are further translated into proteins, which dictate the nature of the biochemical reactions that the cell is able to carry out. Transcriptome

profile is widely used for analysing the genetic regulation of specific cell types (Frith et al. 2005).

Hence, transcriptomics is the study of relative RNA transcript abundances, using microarray technologies and sequencing-based methods. Microarray technology has a library of known origin transcripts specialized for each purpose. These RNA sequences are then reverse transcribed into its complementary DNA from two different samples, labelled with different dyes. And after the cDNA synthesis, the amount of expressed genes can be quantified in which RNAs are up-regulated or down-regulated in a normal sample versus abnormal sample. Moreover, the results from these studies can be highly varying because transcription levels are highly influenced by many factors including disease phenotypes. To be of general use, transcriptomics data must be generated under a wide variety of conditions and compared, to eliminate the trivial sources of variation (Dunwell et al. 2001).

Transcriptome studies were earlier based on the gene predictions and the transcripts were reverse transcribed to cDNA to generate expressed sequence tags (ESTs). These sequences were analysed by the automated Sanger sequencing method, similar to the human genome project (Adams et al. 1991). Later on, initial studies of Serial Analysis of Gene Expression (SAGE) were carried out in the human pancreas (Velculescu et al. 1995). In plants, first studies were carried out in Arabidopsis in which expression of 45 genes were analysed through microarrays containing cDNA spotted on to glass slides (Schena et al. 1995). In the case of plant research, gene expression during different stress conditions are important. Furthermore, a high-throughput screening in proteomics has been developed where the analysis of proteins at different conditions can be studied. The analyses of the products of transcripts were more prevalent than transcriptome analysis. Alteration in gene expression can be monitored through transcriptome or proteome to differentiate between the two biological states of a cell.

1.1.3 Proteomics

Proteomics, the study of a complete set of proteins, the effectors of the biological system and their levels, are reliant on comparing mRNA levels as well as post-translational mechanisms (Graves and Haystead 2002). In certain cases, mRNA level will be high and that cannot be measured by microarray chips in the transcriptome analysis. Therefore, the final product of the biological system, the proteins are profiled and this is considered to be the most significant data set to characterize a bio system. For characterizing proteins, we have to isolate the proteins initially using conventional chromatography techniques and different types of columns are available for this procedure. Columns used for isolation mainly include ion exchange, size exclusion, and affinity chromatography columns. For specific proteins, ELISA and western blotting can be used. But these are incapable of defining the protein expression levels. A complex protein mixture can be separated via SDS PAGE, 2D Gel electrophoresis, and 2D differential gel electrophoresis. In the cases of analysing protein expression, microarray chip-based high-throughput systems are available.

Functional analysis of a complete genome is challenging. So, mass spectrometry (MS) technique is considered better for scrutinizing complex protein mixtures with greater sensitivity. Amino acid sequencing of proteins can be achieved through the Edman degradation sequencing method and quantity of proteins analysed through stable isotope labelling with amino acids in cell culture (SILAC), isotope-coded affinity tag (ICAT) labelling and isobaric tag for relative and absolute quantitation (iTRAQ) techniques. The 3D structure of proteins can be analysed through X-ray crystallography and NMR spectroscopy. We can gain insights into their biological functions through structure data generated by these protocols. Furthermore, various bioinformatics tools are available for the 3D structure prediction, motif and domain analysis, protein-protein interactions and data analysis of MS. Evolutionary relationships can be studied through sequence and structure alignment tools. Thus, proteome investigation gives the total delineation of basic and structural functions of the cell, just as the reaction of the cell against different sorts of pressures/stresses and drugs, utilizing single or multiple proteomics strategies (Aslam et al. 2017).

1.1.4 Metabolomics

Current enrichment through the mass spectrometric study and by means of practical genomic research, the metabolomics area becomes strengthened in research. The foremost activities in this area entail setting up a multifaceted, wholly integrated strategy for the top line sample extraction, metabolite separation/detection/identification, computerized data procuring/handling/analysis, and ultimately quantification. Both analytical and computational trends are indispensable to acquire this goal (Hall et al. 2002).

The known profile of all compounds, synthesized by an organism, is known to be metabolome. The term was coined by Stephen Oliver (College of Manchester, UK) (Oliver et al. 1998). New definition of metabolomics considers it as all the metabolites of a living being gathered to be distinguished and quantified inside an organic framework under specific conditions (Fiehn 2002). A large variety of metabolites, with diverse physical properties, including ionic inorganic compounds, hydrophilic carbohydrates derived through biochemical means, and hydrophobic compounds, are coming under this metabolic pool. A wide variety of compounds are present in the plant kingdom. It appears that there is a complex network among these small molecules in plants, and by detecting the relation among these metabolites, metabolomic investigation contributes to the understanding of the connection essentially between genotype and metabolic outcome by handling key network components. Such a kind of metabolomic investigation, coordinating with transcriptome, has been effectively connected to explore the metabolite concentrations in plants (Hong et al. 2016).

Unlike transcriptome and proteome studies via genomic knowledge and central dogma, it is challenging to generate metabolomics data. In spite of the fact that metabolomics is downstream of the other genomics data (transcriptomics and proteomics), the metabolome of a species is not like transcriptome or proteome and

cannot be accessed specifically by known genomic data using central dogma. Subsequently, metabolomics is used to get a relevant massive data set for the disclosure of genes and pathways through accurate and high-throughput technologies. Owing to this diversity in metabolites, the present metabolomics studies combine large varieties of intricate analytical tools to combine all the varying data from complex biological samples. Among these, a prominent analytic tool used for the analysis is NMR. However, its lower sensitivity and dynamic range make NMR less prominent. Hence, the most powerful technique used is a liquid or gas chromatography coupled with mass spectrometry (Hall et al. 2002).

1.2 Plant Metabolites

There are mainly two types of metabolites present in plants, primary and secondary. Primary metabolites are required for the growth and development and found to be present in every plant and secondary metabolites derive from primary metabolites, which are considered to be important in plant defence mechanisms, and their functions are specific. These are low-molecular-weight compounds that are often used as drugs, dyes, insecticides, flavours and fragrances with high commercial value. They are species, organ, tissue and cell-specific and might be produced in certain conditions like defence against pathogens and herbivores, certain stress conditions, attracting insects and animals for fertilization and/or seed dispersal or repellence of unwanted feeders. They are distributed in different combinations in different parts of the plants like leaves stem, bark, roots, shoots, etc. at different stages of development like seeds, seedlings and plantlets under different ecological conditions in different classes of plants. In normal prospects, they are produced for survival and reproduction (Guerriero et al. 2018).

These metabolites are formed mainly via three pathways in plants, isoprenoid pathway, shikimate pathway, and polyketide pathway. Shikimate pathway is the major pathway for the synthesis of aromatic compounds and further modifications are based on the plant species. In the case of carbon flux, the phenylpropanoid pathway is significant and about 20% of cell metabolism work is based on these and chorismate mutase, found to be the target enzyme in this pathway. Shikimate pathway is prominent because it leads to the production of lignin, flavonoids and anthocyanin. Furthermore, the next relevant pathway is the isoprenoid pathway for the production of terpenoids. There are mainly three classes of metabolites, terpenoids, phenols, and nitrogen- and sulphur-containing components (Thirumurugan et al. 2018).

1.2.1 Phenols

Plants possess a wide range of phenolic compounds derived from phenylalanine or tyrosine by-products of the shikimate pathway. These include flavonoids, lignins, stilbenes, tannins and lignans along with components of suberin and cutin having

long-chain carboxylic acids. These components have been used as pesticides, antibiotics, UV protectants, for establishing symbiosis with rhizobia and so on.

Vanillin, the most important flavouring agent produced in vanilla orchid (*Vanilla planifolia*), constitutes about 1% only of the market share. Most of vanillin is produced by advanced synthetic biology technique for natural products, in laboratory by microorganisms. Hence, the microbial production of secondary metabolites ensures higher productivity and the metabolite can be analysed through mass spectrometry. Earlier, vanillin was produced from ferulic acid via biotransformation using *Streptomyces sannanensis, Escherichia coli* and *Pseudomonas fluorescens*. Later on, a combination of *Aspergillus niger* and *Pycnoporus cinnabarinus* is being used for the microbial catabolism of ferulic acid to vanillin (Pyne et al. 2019).

1.2.2 Terpenoids

Terpenoids, plant isoprenoids, form a major class of secondary metabolites derived from acetyl Co-A or glycolytic pathways and are commercially important components of aroma substance for cosmetics, food and beverages, vitamins (A, D, E) natural insecticides (pyrethrin), rubber and gutta-percha and solvents like turpentine. In the biosynthetic pathway of isoprenoid the starting precursor is acetyl CoA. Combination with two or three acetyl CoA molecules forms mevalonate. Mevlonic acid further gets transformed into isopentenyl pyrophosphate (IPP), the five-membered Skelton of terpenes via several reactions (pyrophosphorylation, decarboxylation, dehydration). IPP is also formed through glycolytic and photosynthetic reduction cycle via methylerythritol phosphate. Based on the number of five carbon units present terpenes are classified as monoterpene (2 C5 units), sesquiterpene (3 C5 units), di terpene (4 C5 units), triterpene (30 carbon), tetraterpene (40 carbon) and polyterpenes ([C5]n) (Kallscheuer et al. 2019).

Microbial production of terpenoids is also favoured globally. Linalool is a sesquiterpenoid produced via microbial system by expressing linalool synthase gene from *Lavandula angustifolia* to Sa*ccharomyces cerevisiae* and significant amount of linalool is produced by the genetically modified organism. In addition to this, monoterpene α-pinene production is also favoured by expressing geranyl diphosphate synthase and a pinene synthase *in Escherichia coli* and *Corynebacterium glutamicum* (Huccetogullari et al. 2019).

1.2.3 Secondary Metabolite with Sulphur Group

Secondary metabolites produced during sulphur metabolism are not usual. But they have a significant role in plant defence mechanism. The main two categories of sulphur-containing compounds are glucosinolates and the alliins. These compounds are usually inactive in plants and in case of any external stimuli like herbivore attack, the enzyme producing these compounds is activated and the products are released. The hydrolytic enzymes activating these compounds are myrosinases and alliinases.

They are usually stored in high amount and released in unfavourable situation leading to the production of metabolites. Alliins are found in Alliaceae family that include onion, garlic and leeks. It expresses many pharmacological activities like antibacterial, antifungal, anticancer and antiviral properties (Thirumurugan et al. 2018).

1.2.4 Secondary Metabolites with Nitrogen Group

Secondary metabolites having nitrogen in their heterocyclic ring are found to be alkaloids and are derived from aromatic amino acids likes tyrosine, tryptophan, aspartic acid and lysine. They are classified into three major categories, one with nitrogen-containing heterocyclic group, the other with non-cyclic biological amines and finally steroidal alkaloids, which are not derived from amino acids. Furthermore, they can block ion channels, hinder proteins, and act on nervous system and membrane transport (Wink 2018). Most of the alkaloids are found to be in its salt form, basic in nature and are soluble in hydro-ethanolic media. There are about 15,000 alkaloids reported from 20% of plants species. Nicotine and atropine alkaloids are true alkaloids with nitrogen in their heterocyclic ring and found to be basic in nature, while caffeine and solanidine are pseudo alkaloids not derived from any amino acid and exhibits basicity. Common ring structure for the true alkaloids includes piperidines, pyrrolidines, isoquinolines, indoles, pyridines and pyrroles. Another kind of alkaloids which are basic in nature but do not have nitrogen in their heterocyclic ring are termed as proto-alkaloids. Mescaline, a phenyl ethyl amine derived alkaloid, is an example for proto-alkaloids (Chomel et al. 2016).

Furthermore, presence of a large variety of these plant secondary metabolites and their impact on communication with environment and in plant defence mechanisms make them a crucial target in the plant metabolome studies. Biotic and abiotic stresses including salinity, temperature, CO_2 availability, and water scarcity challenge the plants to thrive under these conditions. The current scenario with extreme climate changes affects plants' biochemical and physiological mechanisms, and can be identified and studied by modern technologies.

1.3 Methods in Metabolomics

Metabolomics has seen an exponential development in the most recent decade, driven by significant applications spreading over a wide range of domains in life. MS in combination with chromatography and nuclear magnetic resonance are the two noteworthy analytical tools for investigations of metabolites as a mixture. Inferable from its innate and significantly quick information delivery, MS assumes an inexorably prevailing strategy in the metabolomics field. Accordingly, the metabolomics field has moved far. Subjective strategies and basic examples acknowledge ways to deal with a scope of worldwide and focused quantitative

methodologies. They are routinely utilized and give solid information, which bring in more prominent confidence in the derivations determined (Hong et al. 2016).

Metabolomics have effect on vast areas including pharmaceutics, drug development, toxicology, food and plant science, environmental science and medicine. Metabolic profiling studies were first started for analysing the human biological fluids to determine the variability in profiles across individuals. In 1940–1960s studies were carried out in paper chromatography and later on in 1970s, with the emergence of new techniques like GC/LC-MS, allowed the quantitative analysis of the metabolites. Earlier studies were done by Roger Williams and co-workers. Later on, Horning along with Pauling and Robinson teams started developing the metabolic profiling of body fluids by GC-MS (Lytovchenko et al. 2009).

Plants possess an invariant metabolism consisting of important metabolic pathways in various tissues, cells and subcellular organelles to produce metabolites essential for responses against infections, stress and the immediate environment. This coordinated action of all organelles helps in the synthesis of various chemical compounds with extraordinary roles, which make plants a tremendous model for metabolome analysis.

Plants are sessile life forms which cannot escape from the ecological conditions. They adjust themselves by changing metabolic conditions at disturbances in the system. The central metabolism in plants has changed with different developmental processes enabling plants to endure such natural dangers. The estimation of known essential metabolites reveals the plants' mechanism to adjust the fluctuating situations. These varying conditions can be assessed by MS-based technologies (Griffiths 2007; Sawada and Hirai 2013).

Metabolite analysis can be divided into numerous steps.

1. Experimental design.
2. Metabolites extraction, derivatization and analysis by GC/LC-MS.
3. MS files processing, interpretation and statistical analysis.

1.3.1 Designing an Experiment

For designing an experiment, the correct biological circumstances have to be chosen. One of the key factors in arranging metabolomics studies is to examine every conceivable source of variety that can apply credible effect on the theory (Jorge et al. 2016). The experimental conditions include environmental conditions under which the plant is developing, treatment undergone and the tissues used for further metabolomics analysis, and transfer the data into a biologically important information realm. It is required to consider the natural conditions and its impact on the metabolome when designing an experiment. The levels of multifaceted nature of those reactions are exposed to additional levels of troubles when the impacts of different abiotic or biotic stresses are examined. In regard to biotic stress, for instance, usually an outstanding arrangement of metabolites is shared between the

plant and the connecting accomplice, impeding the distinguishing proof of changes focused to every life form (Jorge and António 2017).

Determining the size of the experiment is also important. So the experimental size is based on certain variables, including variation occurring during developmental stages, genotypes, growth conditions and tissues analysed. Furthermore, there are intra- and inter-species variations in the metabolites. For keeping sample integrity in small- and large-scale experiments, randomization is critical. Indeed, keeping the plants in similar levels of variation during the examination is also important. Compared to nursery or field tests, plants are exposed to more prominent variations under natural conditions, which are unavoidable, yet can promptly be perceived. Randomization ensures that treatment and control tests are similar, as far as the watched and unobservable features are concerned, and reducing the contrasts in soil and environment heterogeneity.

1.3.2 Sample Preparation

First of all, the plant tissues for the analysis has to be selected and the extraction procedure follows. The extraction process, either hot extraction or cold extraction, is chosen based on the hypothesis of the metabolites to be analysed. If thermolabile compounds are present, then cold extraction is more preferable than hot extraction. So, at the initial stage a hypothesis is created, and based on that plants are selected and extraction procedures followed. The observation intervals are similarly significant since numerous metabolites, including essential and optional mixes, may change during the diurnal cycle. Abrupt changes in the natural conditions, such as covering the plants during examining, can likewise alter the metabolic production (Jin et al. 2017; Henion et al. 1998). Since the establishment of Ayurvedic practices of Charaka and Susruta, different medicinal plant collection protocols are implemented for optimum results (Tavhare and Nishteswar 2014 and references there in).

A large number of techniques could be employed to preserve the integrity of plant metabolites during harvesting. Most popular technique is storing of plant tissues in liquid nitrogen. While using liquid nitrogen, there is a possibility of conversion of chemical constituents, which is not favourable. An alternative method is freeze clamping; but it is also challenging to use in large-scale investigations. The tissue must be frozen in every step of reaping and sample preparation, if fresh materials are used. Every sample should be properly labelled during the reaping and sample preparation. It may be advisable to utilize best quality plastic tubes that will persevere through outrageous temperatures, for example, liquid nitrogen. Use of screw capped cryogenic micro tubes can avoid the chance of explosion and loss of materials during the thawing procedure. As indicated by the planned examinations, tests can be put away for new material solidified at −80 °C or if solidified at −20 °C or room temperature. In this context, use of fresh materials are essential for analysis of volatile and semi-volatile compounds (Harborne 1984).

It may be essential to granulate, filtrate, fractionate, or pre-concentrate the tissues according to the nature of plant tissues. The extraction efficiency of the tissues can be improved by tissue crushing, which permits sample homogenization and increase in contact surface area. Moreover, for handling a large number of samples, there is a wide range of crushing instruments available in the market (e.g. ball processors and robotized cryogenic processors). However, a less difficult, conceivable decision is to utilize a precooled mortar and pestle with liquid nitrogen. The containers for extraction procedure must be contamination free. A large variety of metabolites are found in plants and their physicochemical properties are also different. This will create problems for selecting extraction protocol due to the variations in plant components. Extraction should be carried out with less degradation or modification of components.

A better extraction method must be followed for analysing a large number of small molecules. A few parameters must be considered all through the extraction system to remove the metabolites productively from cells or tissues of intrigue. More often than not, the examination of metabolites pursues two potential approaches, directed or untargeted. In focused metabolomics, a characterized gathering of known and biochemically commented on highlights is recognized, while untargeted examinations are fair to give a review of all the quantifiable analytes in a sample, including obscure compounds. The decision on the extraction technique will improve the part of metabolites in focused investigations, with comparable chemical properties. In spite of a few endeavours to improve the metabolite inclusion required in untargeted investigations, the choice of an extraction method will support a particular class of compounds, presenting some sort of specificity. Hence, a mix of extraction techniques and diagnostic advances is recommended to intensify the quantity of recognizable metabolic highlights (Stein 1999).

Suitable solvents for the extraction purpose may have to be chosen. Selection of solvents is based on their different properties like polarity, selectivity, inertia, sample solvent ratio and the interaction between metabolites and biological matrix. The disintegration rate in the dissolvable extraction might be moderate, depending on the associations among metabolites and biological matrix. Despite the fact that there may be a problem of compound degradation, the time lag in extraction and temperature may support metabolite extraction. Hence, we can utilize ultra-sonication for cell wall breaking as an efficient method in extraction procedure. It is likewise critical to store the concentrates properly, according to the strength of the metabolites in focus. Also, the solvent purity, pH and its vaporization are also considered because these may change the metabolite concentration and texture. For example, plasticizers or surfactants can ruin the analysis. Hence, we have to use the extraction buffer within a short period of time.

The actual mode of extraction is based on the water content and texture of the compound being isolated. It is advisable to wash initially the desired plant part in boiling ethanol for killing the hydrolytic enzymes. Alcohol is a good solvent for extracting most of the plant components. If green tissues are used, alcohol extraction leads to the removal of chlorophyll and many of the low-molecular-weight compounds. Further extraction would be completely free of green colour. The

dried plant parts are usually extracted in organic solvents ranging from non-polar (ether, hexane, chloroform) to polar (ethyl acetate and alcohol) in a soxhlet apparatus. Other methods like sonication and refluxing are also used for extraction. Extractions with non-polar solvents extract lipids and terpenoids whereas polar solvents separate polar compounds. The physiological parameter, mainly temperature, has a profound effect on purifying the constituents from crude extract. The solubility of constituents is also an important factor, in which liquid–liquid extraction is most favoured. If non-polar constituents are needed, partitioning of crude extract with water would remove unwanted components. On the other hand, if polar compounds are needed, the crude extract may be extracted with suitable organic solvents after aqueous extraction. Further partition with acid or alkali would remove acidic and basic impurities (Jorge and António 2017). Acetonitrile is a better solvent for protein precipitation than methanol. In general, there is no 'generic protocol' for the extraction procedure. Any of the methods suitable to the component or metabolite that is being analysed need to be followed. By considering all these factors and after adopting suitable method for extraction, the concentrate should be perfect, with respect to the investigative procedure to be followed. Because of the highly sophisticated instruments (LC/MS) used for the further analysis of the analyte, an ideal concentrate of the sample is necessary. Otherwise, it will affect the ion suppression in instrument affecting the ionization efficiency of the compounds.

1.3.3 Metabolite Analysis via GC/LC-MS

Analysis of metabolites mainly depends upon the goal of the study. Mass spectrometry (MS) has been broadly utilized in plant metabolomics in combination with gas or liquid chromatography (GC and LC, respectively). The amount of metabolite measured is based on the specificity of the detection method and resolution of the chromatographic system. For the detection of volatile and semi-volatile compounds GC MS is more suitable (Miyagawa and Bamba 2019). It offers powerful measurement of a few metabolite classes in plants requiring a derivatization step for compounds with insufficient volatility in their normal state and thermal stability. There are a variety of methods for derivatization of compounds like silylation (adding silyl group by replacing hydrogen in the functional groups of metabolites), alkylation (using chloroformates, especially methyl chloroformate as the derivatizing agent), alkoxyamination and acylation. Alkylation cannot derivatize sugar and sugar alcohol. But silylation is broad and can generate variety of stable, less polar and volatile derivatives of parent compound. In plant metabolomics, the initial stage is the stabilization of carbonyl moiety in plant components by using methoxyamine in pyridine, thereby overwhelming the keto-enol tautomerism and the creation of multiple acetal or ketal structures. This process would aid to decrease the numbers of derivatives of reducing sugar and produce only two forms of –N=C< derivatives. In the next step, the functional groups like –OH, –COOH, –SH, and –NH groups are converted to TMS ethers, TMS esters, TMS sulphides or TMS amines, using a trimethyl silyl reagent usually BSTFA (N,O-bis(trimethylsilyl)

trifluroacetamide) or MSTFA (N-methyl-N-(trimethylsilyl) trifluoroacetamide). TMS derivatization has been thoroughly investigated and found to be very efficient. But in the case of certain compounds, multiple derivatives are formed after derivatization and it would make analysis difficult. In such cases, LC-MS is very useful where derivatization procedure is not used (Griffiths 2007).

The TMS derivatization procedure is generally used in the case of steroids. For the hydroxyl group derivatization of steroids, TMS ethers are used. But considering the oxo group in steroids, there may be certain problems as they react under forceful conditions and the molecular ion of poly hydroxyl steroid TMS ethers would be absent due to the extensive loss of trimethyl silanol to give ions. This can be overcome by preparation of TMS alkyl dimethylsilyl derivatives, which would help in the interpretation of spectra. Sterically congested alkylsilyl ethers are more stable than TMS ethers and can be purified by conventional chromatographic approaches (Zarate et al. 2016).

Other than the derivatization procedure, micro chemical or enzymatic responses can help in structure resolution or can be utilized to give appropriate mass spectrometric properties. Simple reactions include conversion of vicinal hydroxyl group into acetonides or boronates, periodate oxidation of vicinal hydroxyl group, oxidation with sodium bismuthate, selective or complete oxidation of –OH group with chromic acid in acetone, reduction of carbonyl group with sodium borohydride or lithium aluminium hydride, etc. (Griffiths 2007).

The analyst must be aware that, the trimethylsilylation process, even under the mild conditions generally employed in plant metabolite profiling, leads to by-products conversion. For example, arginine in reaction with BSTFA is converted to ornithine. Hence for interpreting the results, the analyst should take care of the intermediates formed that are reactive or unstable in the metabolic processes. If any problem occurs during interpretation, the analyst can perform another derivatization process with respect to the specific functional groups (Lytovchenko et al. 2009).

The trimethylsilyl subsidiaries of many plant metabolites, including sugars (e.g. glucose and fructose), have fundamentally the same EI mass spectra. So for recognizing them, GC retention constraints are especially significant. Since retention times fluctuate with the column dimensions, stationary stage and temperature, relative retention times may be incorporated with appropriate parameters of correlation, called retention indices (RI) (Halket et al. 2005). Relative retention times are basically the proportions of analyte times to the hour of a selected standard compound. An increasingly widespread arrangement depends on purported Kovats indices (Kováts 1958) which relate to the retention times of the metabolites to the retention times of *n*-alkanes examined under similar conditions, even by co-infusion. For instance, a sugar subsidiary having a similar retention time as the *n*-alkane, C20 would be allotted an index of 2000 or a ‘methylene unit’ estimation of 20.00. One showing up somewhere between C20 and C21, would be doled out a list of 2050, etc. Most manufacturers’ information frameworks don’t deal with RI well. Nevertheless, the AMDIS (Stein 1999) from the NIST has astounding RI abilities and is promptly accessible for download (http://chemdata.nist.gov/mass-spc/amdis/). The software can peruse most makers’ information documents and perform mass spectral