

# Biotechnology in Agriculture and Forestry

---

Edited by

T. Nagata (Managing Editor)

H. Lörz

J. M. Widholm

# Biotechnology in Agriculture and Forestry

---

Volumes already published and in preparation are listed at the end of this book.

---

# Biotechnology in Agriculture and Forestry 57

---

## *Plant Metabolomics*

Edited by  
K. Saito, R.A. Dixon, and L. Willmitzer

With 96 Figures, 29 in Color, and 10 Tables

 Springer

*Series Editors*

Professor Dr. TOSHIYUKI NAGATA  
University of Tokyo  
Graduate School of Science  
Department of Biological Sciences  
7-3-1 Hongo, Bunkyo-ku  
Tokyo 113-0033, Japan

Professor Dr. HORST LÖRZ  
Universität Hamburg  
Biozentrum Klein Flottbek  
Zentrum für Angewandte Molekularbiologie  
der Pflanzen (AMP II)  
Ohnhorststraße 18  
22609 Hamburg, Germany

Professor Dr. JACK M. WIDHOLM  
University of Illinois  
285A E.R. Madigan Laboratory  
Department of Crop Sciences  
1201 W. Gregory  
Urbana, IL 61801, USA

*Volume Editors*

Professor Dr. KAZUKI SAITO  
Chiba University  
Graduate School of Pharmaceutical Sciences  
Yayoi-cho 1-33, Inage-ku  
Chiba 263-8522, Japan;  
RIKEN Plant Science Center  
Yokohama 230-0045, Japan

Professor Dr. RICHARD A. DIXON  
Plant Biology Division  
Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway  
Ardmore, OK 73401, USA

Professor Dr. LOTHAR WILLMITZER  
Max Planck Institute  
of Molecular Plant Physiology  
Am Mühlberg 1  
14476 Golm, Germany

Library of Congress Control Number: 2005936763

ISSN 0934-943X

ISBN-10 3-540-29781-2 Springer Berlin Heidelberg New York

ISBN-13 978-3-540-29781-9 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable for prosecution under the German Copyright Law.

Springer is a part of Springer Science + Business Media  
springer.com

© Springer-Verlag Berlin Heidelberg 2006  
Printed in Germany

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

Editor: Dr. Dieter Czeschlik, Heidelberg, Germany  
Desk Editor: Dr. Andrea Schlitzberger, Heidelberg, Germany  
Cover design: *design&production* GmbH, Heidelberg, Germany  
Typesetting and production: LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig, Germany  
Printed on acid-free paper 31/3152 5 4 3 2 1 0

## Preface

Metabolomics is a rapidly-emerging sector of post-genome research. The metabolome (a set of all metabolites of an organism) represents not only the ultimate phenotype of cells by the perturbation of gene expression and the modulation of protein functions caused by the environment or mutations, but the metabolome can also feed back on gene expression and protein function. Therefore, metabolomics plays a key role for understanding cellular systems. Metabolomics is applied to a variety of biological fields from medical science to agriculture. Nevertheless, metabolomics research is particularly important in the plant field because plants collectively produce a huge variety of chemical compounds, far more than animals and even microorganisms. The number of all metabolites in the plant kingdom is estimated at 200,000 or more. In addition, most of the human-beneficial properties of plants, be they foods, medicinal resources, or industrial raw materials, are ascribed to plant metabolites.

This book aims to review the current status of plant metabolomics research. Since metabolomics itself is a new field, no such comprehensive book has yet been published. The chapters are divided into three sections: analytical technology, bioinformatics, and applications. These represent three major elements of metabolomics research. Each chapter provides cutting-edge information contributed by leading researchers from throughout the world.

We hope that this book will be a landmark for plant metabolomics research into the future and will give beneficial guidance to graduate students and researchers in academia, industry, and technology transfer organizations. Since metabolomics is still a growing discipline, further technology development in chemical analysis and bioinformatics will be required. We look forward to breakthrough technology innovations in metabolomics, and yet unforeseen findings and applications in plant science.

Finally, we would like to acknowledge our contributors who have enthusiastically put their efforts to ensure the high scientific quality of this volume. We also would like to thank our colleagues at Springer.

January 2006

Kazuki Saito,  
Richard A. Dixon,  
and Lothar Willmitzer

# Contents

## Section I Analytical Technology

I.1	Gas Chromatography Mass Spectrometry . . . . .	3
	J. KOPKA	
1	Introduction . . . . .	3
2	GC-MS Profiling Technology in a Nutshell . . . . .	5
3	Short Excursion into Nomenclature and Definitions . . . . .	10
4	Present Challenges of GC-MS Profiling . . . . .	13
	References . . . . .	17
I.2	Current Status and Forward Looking Thoughts on LC/MS Metabolomics . . . . .	21
	L.W. SUMNER	
1	Introduction . . . . .	21
2	Chromatography Theory . . . . .	24
3	Limitations of Current Metabolic Profiling Approaches and Proposed Solutions to Advance Metabolomics . . . . .	25
4	Future Directions and Forward-Looking Thoughts . . . . .	28
	References . . . . .	30
I.3	Plant Metabolomics Strategies Based upon Quadrupole Time of Flight Mass Spectrometry (QTOF-MS) . . . . .	33
	H.A. VERHOEVEN, C.H. RIC DE VOS, R.J. BINO, and R.D. HALL	
1	Introduction . . . . .	33
2	The Technology . . . . .	34
3	Data Analysis . . . . .	37
4	Application of QTOF MS-based Plant Metabolomics Analyses . . . .	38
5	Conclusions and Future Prospects . . . . .	46
	References . . . . .	46

I.4	Capillary HPLC . . . . .	49
	T. IKEGAMI, E. FUKUSAKI, and N. TANAKA	
1	Introduction . . . . .	49
2	Monolithic Silica Columns for Micro HPLC . . . . .	49
3	Applications of Monolithic Silica Columns to Metabolomics . . . . .	54
4	Two-Dimensional HPLC . . . . .	55
5	Combination of Reversed-Phase HPLC and Other Separation Modes . . . . .	59
6	Outlook . . . . .	61
	References . . . . .	61
I.5	Capillary HPLC Coupled to Electrospray Ionization Quadrupole Time-of-flight Mass Spectrometry . . . . .	65
	S. CLEMENS, C. BÖTTCHER, M. FRANZ, E. WILLSCHER, E. V. ROEPENACK-LAHAYE, and D. SCHEEL	
1	Introduction . . . . .	65
2	Extraction, Chromatography and Mass Spectrometry . . . . .	67
3	Potential and Limitations . . . . .	72
4	Conclusions and Outlook . . . . .	77
	References . . . . .	78
I.6	NMR Spectroscopy in Plant Metabolomics . . . . .	81
	J.L. WARD and M.H. BEALE	
1	Introduction . . . . .	81
2	High-throughput Screening by 1D <sup>1</sup> H-NMR . . . . .	82
3	Data Analysis . . . . .	84
4	Two-dimensional NMR . . . . .	85
5	Stable Isotope Labelling . . . . .	86
6	Hyphenated NMR . . . . .	87
7	Discussion: Applying NMR to Plant Metabolomics . . . . .	88
	References . . . . .	89
I.7	Hetero-nuclear NMR-based Metabolomics . . . . .	93
	J. KIKUCHI and T. HIRAYAMA	
1	Introduction . . . . .	93
2	Historical Aspects of NMR Studies of Plant Metabolism . . . . .	93
3	<sup>1</sup> H-NMR-based Metabolomics . . . . .	94
4	Use of Stable Isotope Labeling Technique to Enable Monitoring of the Dynamic Movement of Metabolites . . . . .	94
5	Approach for Hetero-nuclear NMR-based Metabolomics . . . . .	95
6	Prospects for the Future . . . . .	98
	References . . . . .	99

**Section II Bioinformatics**

II.1	Bioinformatics Approaches to Integrate Metabolomics and Other Systems Biology Data . . . . .	105
	B. MEHROTRA and P. MENDES	
1	Introduction . . . . .	105
2	Databases . . . . .	107
3	Data Visualization . . . . .	110
4	Data Analysis . . . . .	111
5	Conclusion . . . . .	112
	References . . . . .	113
II.2	Chemometrics in Metabolomics – An Introduction . . . . .	117
	J. TRYGG, J. GULLBERG, A.I. JOHANSSON, P. JONSSON, and T. MORITZ	
1	Introduction . . . . .	117
2	Theory and Methods . . . . .	118
3	Example: Metabolomics Study on Arabidopsis Mutants . . . . .	125
4	Summary and Future Prospectives . . . . .	126
	References . . . . .	127
II.3	Map Editor for the Atomic Reconstruction of Metabolism (ARM) . . . . .	129
	M. ARITA, Y. FUJIWARA, and Y. NAKANISHI	
1	Introduction . . . . .	129
2	Definition of Metabolic Information . . . . .	131
3	Metabolic Map Editor . . . . .	133
4	Applications . . . . .	137
5	Conclusions . . . . .	139
	References . . . . .	139
II.4	AraCyc: Overview of an Arabidopsis Metabolism Database and its Applications for Plant Research . . . . .	141
	S.Y. RHEE, P. ZHANG, H. FOERSTER, and C. TISSIER	
1	Introduction . . . . .	141
2	Database Content . . . . .	142
3	Search, Browse, and Analyze Functionalities . . . . .	145
4	Applications of AraCyc . . . . .	149
5	Current Issues and Future Directions . . . . .	152
6	Conclusions . . . . .	152
	References . . . . .	153
II.5	KaPPA-View: A Tool for Integrating Transcriptomic and Metabolomic Data on Plant Metabolic Pathway Maps . . . . .	155
	T. TOKIMATSU, N. SAKURAI, H. SUZUKI, and D. SHIBATA	
1	Introduction . . . . .	155
2	General Features of the KaPPA-View Tool . . . . .	155



3	Plant Metabolic Pathway Maps . . . . .	158
4	Integration of Transcriptomic and Metabolomic Data on Pathway Maps . . . . .	159
5	Comparison with Other Databases and Tools . . . . .	159
6	Limitations and Future Improvements . . . . .	160
7	Conclusions . . . . .	162
	References . . . . .	163
II.6	KNApSACk: A Comprehensive Species-Metabolite Relationship Database . . . . .	165
	Y. SHINBO, Y. NAKAMURA, M. ALTAF-UL-AMIN, H. ASAHI, K. KUROKAWA, M. ARITA, K. SAITO, D. OHTA, D. SHIBATA, and S. KANAYA	
1	Introduction . . . . .	165
2	Search Options of the KNApSACk Database . . . . .	166
3	Statistics of the Database . . . . .	172
4	Classification Based on Common Metabolites . . . . .	177
5	Conclusion and Remarks . . . . .	179
6	Access to KNApSACk . . . . .	179
	References . . . . .	180

### Section III Applications

III.1	Systems Biology: A Renaissance of the Top-down Approach for Plant Analysis . . . . .	185
	F. CARRARI, N. SCHAUER, L. WILLMITZER, and A.R. FERNIE	
1	Introduction . . . . .	185
2	Re-emergence of Top-down Thinking . . . . .	186
3	Systems Biology in Non-plant Systems . . . . .	186
4	Systems Biology in Plant Systems . . . . .	188
5	Dynamic Profiling in Plant Cells . . . . .	192
6	Conclusions and Future Perspectives . . . . .	195
	References . . . . .	195
III.2	Systems-based Analysis of Plant Metabolism by Integration of Metabolomics with Transcriptomics . . . . .	199
	M.Y. HIRAI, T. TOHGE, and K. SAITO	
1	Introduction . . . . .	199
2	Understanding Whole Plant Metabolism – Our Aims and Strategy .	199
3	Metabolome and Transcriptome Analyses . . . . .	200
4	Studies on Sulfur Metabolism . . . . .	201
5	Studies on Anthocyanin Metabolism . . . . .	206

6	Conclusions . . . . .	208
	References . . . . .	209
III.3	Targeted Profiling of Fatty Acids and Related Metabolites . . . . .	211
	T.R. LARSON and I.A. GRAHAM	
1	Introduction . . . . .	211
2	Metabolite Profiling Techniques Used to Study Plant Lipid Metabolism . . . . .	213
3	Future Developments . . . . .	223
	References . . . . .	224
III.4	Metabolic Profiling and Quantification of Carotenoids and Related Isoprenoids in Crop Plants . . . . .	229
	P.D. FRASER and P.M. BRAMLEY	
1	Introduction . . . . .	229
2	Analytical Methodologies Employed in the Analysis of Carotenoids . . . . .	233
3	Examples of Carotenoid/isoprenoid Profiling . . . . .	237
4	Conclusions . . . . .	240
	References . . . . .	240
III.5	Metabolomics and Gene Identification in Plant Natural Product Pathways . . . . .	243
	R.A. DIXON, L. ACHNINE, B.E. DEAVOURS, and M. NAOUMKINA	
1	Introduction . . . . .	243
2	Gene Discovery – Past and Present Strategies . . . . .	243
3	Enzyme Promiscuity in Natural Product Pathways . . . . .	246
4	Examples of the Use of Metabolomics in the Elucidation of Gene Function . . . . .	247
5	Single Cell or Isolated Tissue Metabolomics . . . . .	253
6	Concluding Remarks . . . . .	256
	References . . . . .	256
III.6	Metabolomic Analysis of <i>Catharanthus roseus</i> Using NMR and Principal Component Analysis . . . . .	261
	H.K. KIM, Y.H. CHOI, and R. VERPOORTE	
1	Introduction . . . . .	261
2	Experimental Consideration for Metabolomics Using NMR . . . . .	262
3	Application of NMR for Plant Metabolome . . . . .	266
4	Principal Component Analysis . . . . .	273
5	Concluding Remarks . . . . .	275
	References . . . . .	275

III.7	Metabolomics of Plant Secondary Compounds: Profiling of <i>Catharanthus</i> Cell Cultures . . . . .	277
	M. OREŠIČ, H. RISCHER, and K.-M. OKSMAN-CALDENTY	
1	Introduction . . . . .	277
2	Metabolomics as a Platform to Study Plant Secondary Metabolites .	278
3	Case Study: Metabolic Profiling of <i>Catharanthus roseus</i> Cells . . . . .	280
4	Protocol . . . . .	285
5	Perspectives . . . . .	286
	References . . . . .	287
III.8	The <i>Taxus</i> Metabolome and the Elucidation of the Taxol® Biosynthetic Pathway in Cell Suspension Cultures . . .	291
	R.E.B. KETCHUM and R.B. CROTEAU	
1	Introduction . . . . .	291
2	Results and Discussion . . . . .	294
3	Protocol . . . . .	306
4	Conclusion . . . . .	307
	References . . . . .	308
III.9	The Use of Non-targeted Metabolomics in Plant Science . . . . .	311
	T. DASKALCHUK, P. AHIAHONU, D. HEATH, and Y. YAMAZAKI	
1	Introduction . . . . .	311
2	Fundamental Investigations into Plant Metabolomics . . . . .	313
3	Conclusion . . . . .	324
	References . . . . .	324
III.10	Plant Metabolite Profiling for Industrial Applications . . . . .	327
	R.N. TRETHERWEY	
1	Introduction . . . . .	327
2	The Metabolome . . . . .	327
3	Profiling Technologies . . . . .	328
4	High Throughput Metabolite Profiling . . . . .	332
5	Industrial Applications . . . . .	335
6	Outlook . . . . .	338
	References . . . . .	338
	Subject Index . . . . .	341

# List of Contributors

L. ACHNINE

Plant Biology Division, Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA

P. AHIAHONU

Phenomenome Discoveries Inc., 204–407 Downey Road, Saskatoon, Saskatchewan, Canada S7N 4L8

M. ALTAF-UL-AMIN

Department of Bioinformatics and Genomics, Graduate School of Information Science, Nara Institute of Science and Technology (NAIST), Takayama-cho 8916–5, Ikoma, Nara 630–0101, Japan

M. ARITA

Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, 5–1–5 Kashiwanoha, Kashiwa, 277–8561 Japan, e-mail: arita@k.u-tokyo.ac.jp

H. ASAHI

Department of Bioinformatics and Genomics, Graduate School of Information Science, Nara Institute of Science and Technology (NAIST), Takayama-cho 8916–5, Ikoma, Nara 630–0101, Japan

M.H. BEALE

The National Centre for Plant and Microbial Metabolomics, Rothamsted Research, West Common, Harpenden, Herts. AL5 2JQ, UK, e-mail: mike.beale@bbsrc.ac.uk

R.J. BINO

Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands

C. BÖTTCHER

Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle/Saale, Germany

P.M. BRAMLEY

School of Biological Sciences, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK

F. CARRARI

Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476, Golm, Germany

Y.H. CHOI

Division of Pharmacognosy, Section Metabolomics, Institute of Biology, Leiden University, P.O. Box 9502, 2300RA, Leiden, The Netherlands

S. CLEMENS

Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle/Saale, Germany, e-mail: sclemens@ipb-halle.de

R.B. CROTEAU

Institute of Biological Chemistry, Washington State University, Pullman, WA 99164, USA

T. DASKALCHUK

Phenomenome Discoveries Inc., 204–407 Downey Road, Saskatoon, Saskatchewan, Canada S7N 4L8, e-mail: info@phenomenome.com

B.E. DEAVOURS

Plant Biology Division, Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA

R.A. DIXON

Plant Biology Division, Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA, e-mail: radixon@noble.org

A.R. FERNIE

Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476, Golm, Germany, e-mail: fernie@mpimp-golm.mpg.de

H. FOERSTER

Carnegie Institution, Department of Plant Biology, 260 Panama Street, Stanford, CA 94305, USA

M. FRANZ

Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle/Saale, Germany

P.D. FRASER

School of Biological Sciences, Royal Holloway, University of London, Egham, Surrey, TW20 0EX, UK, e-mail: p.bramley@rhul.ac.uk

Y. FUJIWARA

Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, 277-8561 Japan

E. FUKUSAKI

Department of Biotechnology, Graduate School of Engineering, Osaka Univ, 2-1 Yamadaoka, Suita, 565-0871, Japan, e-mail: fukusaki@bio.eng.osaka-u.ac.jp

I.A. GRAHAM

CNAP, Department of Biology (Area 7), University of York, PO Box 373, York YO10 5YW, UK

J. GULLBERG

Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, SE-901 87 Umeå, Sweden

R.D. HALL

Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands, e-mail: robert.hall@wur.nl

D. HEATH

Phenomenome Discoveries Inc., 204-407 Downey Road, Saskatoon, Saskatchewan, Canada S7N 4L8

M.Y. HIRAI

RIKEN Plant Science Center, Suehiro-cho 1-7-22, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

T. HIRAYAMA

International Graduate School of Arts and Sciences, Yokohama City University, 1-7-29 Suehiro, Tsurumi-ku, Yokohama, 230-0045 Japan

T. IKEGAMI

Department of Polymer Science and Engineering, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto, 606-8585, Japan

A.I. JOHANSSON

Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, SE-901 87 Umeå, Sweden

P. JONSSON

Research Group for Chemometrics; Organic Chemistry, Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden

K.-M. OKSMAN-CALDENTY

VTT Biotechnology, P.O. Box 1500, 02044 VTT, Finland,  
e-mail: Kirsi-Marja.Oksman@vtt.fi

S. KANAYA

Department of Bioinformatics and Genomics, Graduate School of Information Science, Nara Institute of Science and Technology (NAIST), Takayama-cho 8916-5, Ikoma, Nara 630-0101, Japan,  
e-mail: skanaya@gtc.naist.jp

R.E.B. KETCHUM

Institute of Biological Chemistry, Washington State University, Pullman, WA 99164, USA, e-mail: rketchum@wsu.edu

J. KIKUCHI

RIKEN Plant Science Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, 230-0045 Japan, e-mail: kikuchi@psc.riken.jp

H.K. KIM

Division of Pharmacognosy, Section Metabolomics, Institute of Biology, Leiden University, P.O. Box 9502, 2300RA, Leiden, The Netherlands,  
e-mail: verpoort@chem.leidenuniv.nl

J. KOPKA

Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany, e-mail: Kopka@mpimp-golm.mpg.de

K. KUROKAWA

Department of Bioinformatics and Genomics, Graduate School of Information Science, Nara Institute of Science and Technology (NAIST), Takayama-cho 8916-5, Ikoma, Nara 630-0101, Japan

T.R. LARSON

CNAP, Department of Biology (Area 7), University of York, PO Box 373, York YO10 5YW, UK, e-mail: trl1@york.ac.uk

B. MEHROTRA

Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Washington St., MC 0477, Blacksburg, Virginia 24061, USA

P. MENDES

Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Washington St., MC 0477, Blacksburg, Virginia 24061, USA, e-mail: mendes@vt.edu

T. MORITZ

Umeå Plant Science Centre, Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, SE-901 87 Umeå, Sweden, e-mail: thomas.moritz@genfys.slu.se

Y. NAKAMURA

Department of Bioinformatics and Genomics, Graduate School of Information Science, Nara Institute of Science and Technology (NAIST), Takayama-cho 8916-5, Ikoma, Nara 630-0101, Japan

Y. NAKANISHI

Intec Web and Genome Informatics Corporation

M. NAOUMKINA

Plant Biology Division, Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA

D. OHTA

Department of Plant Genes and Physiology, Graduate School of Agriculture and Biological Sciences, Osaka Prefecture University, Gakuen-cho 1-1, Sakai, Osaka 599-8531, Japan

M. OREŠIĆ

VTT Biotechnology, P.O. Box 1500, 02044 VTT, Finland

S.Y. RHEE

Carnegie Institution, Department of Plant Biology, 260 Panama Street, Stanford, CA 94305, USA, e-mail: rhee@acoma.stanford.edu

C.H. RIC DE VOS

Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands

H. RISCHER

VTT Biotechnology, P.O. Box 1500, 02044 VTT, Finland



K. SAITO

Chiba University, Graduate School of Pharmaceutical Sciences, Yayoi-cho  
1-33, Chiba 263-8522, Japan  
RIKEN Plant Science Center, Suehiro-cho 1-7-22, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan, e-mail: ksaito@faculty.chiba-u.jp

N. SAKURAI

Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba  
292-0818, Japan

N. SCHAUER

Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1,  
14476, Golm, Germany

D. SCHEEL

Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle/Saale,  
Germany

D. SHIBATA

Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba  
292-0818, Japan, e-mail: shibata@kazusa.or.jp

Y. SHINBO

New Energy and Industrial Technology Development Organization, Toshima,  
Tokyo 170-6028, Japan

L.W. SUMNER

The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore,  
OK 73401, USA, e-mail: lwsumner@noble.org

H. SUZUKI

Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba  
292-0818, Japan

N. TANAKA

Department of Polymer Science and Engineering, Kyoto Institute of  
Technology, Matsugasaki, Sakyo-ku, Kyoto, 606-8585, Japan

C. TISSIER

Carnegie Institution, Department of Plant Biology, 260 Panama Street,  
Stanford, CA 94305, USA

T. TOHGE

RIKEN Plant Science Center, Suehiro-cho 1-7-22, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan

T. TOKIMATSU

Kazusa DNA Research Institute, 2-6-7 Kazusa-kamatari, Kisarazu, Chiba  
292-0818, Japan

R.N. TRETHERWEY

metanomics GmbH and metanomics Health GmbH, Tegeler Weg 33, 10589  
Berlin, Germany, e-mail: richard.trethewey@metanomics.de

J. TRYGG

Research Group for Chemometrics; Organic Chemistry, Department of  
Chemistry, Umeå University, SE-901 87 Umeå, Sweden

E. V. ROEPENACK-LAHAYE

Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle/Saale,  
Germany

H.A. VERHOEVEN

Plant Research International, P.O. Box 16, 6700 AA Wageningen, The  
Netherlands

R. VERPOORTE

Division of Pharmacognosy, Section Metabolomics, Institute of Biology,  
Leiden University, P.O. Box 9502, 2300RA, Leiden, The Netherlands

J.L. WARD

The National Centre for Plant and Microbial Metabolomics, Rothamsted  
Research, West Common, Harpenden, Herts. AL5 2JQ, UK,  
e-mail: Jane.ward@bbsrc.ac.uk

L. WILLMITZER

Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1,  
14476, Golm, Germany

E. WILLSCHER

Leibniz Institute of Plant Biochemistry, Weinberg 3, 06120 Halle/Saale,  
Germany

Y. YAMAZAKI

Phenomenome Discoveries Inc., 204-407 Downey Road, Saskatoon,  
Saskatchewan, Canada S7N 4L8

P. ZHANG

Carnegie Institution, Department of Plant Biology, 260 Panama Street,  
Stanford, CA 94305, USA

## **Section I Analytical Technology**

# I.1 Gas Chromatography Mass Spectrometry

J. KOPKA<sup>1</sup>

## 1 Introduction

GC-MS technology has been used for decades in studies which aim at the exact quantification of metabolite pool size and metabolite flux. Exact quantification has traditionally been focused on a single or small set of predefined target metabolites. Today GC-MS is one of the most widely applied technology platforms in modern metabolomic studies. Since early applications in unravelling the mode of action of herbicides (Sauter et al. 1988) it has experienced a renaissance (Fig. 1) in post-genomic, high-throughput fingerprinting and metabolite profiling of genetically modified (e. g. Roessner et al. 2001a,b, 2002; Fernie et al. 2004) or experimentally challenged plant samples (e. g. Cook et al. 2004; Kaplan et al. 2004; Urbanczyk-Wochniak and Fernie 2005). Metabolic phenotyping and analysis of respective phenocopies by metabolite profiling has become an integral part of plant functional genomics (Fiehn et al. 2000b; Roessner et al. 2002; Fernie et al. 2004). The essence of metabolite profiling, namely the non-biased screening of biological samples for changes of metabolite levels relative to control samples, has been thoroughly discussed earlier and is clearly distinguished from fingerprinting approaches and the concept of exact quantification (Fiehn et al. 2000b; Sumner et al. 2003; Birkemeyer et al. 2005).

GC-MS-based metabolome profiling analysis is on the verge of becoming a routine technology. This fact substantially contributes to the development of metabolomics as a fourth integral part of the Rosetta stone for functional genomics and molecular physiology (Trethewey et al. 1999; Fiehn et al. 2000b; Trethewey 2004). Nevertheless, GC-MS technology is already challenged again by new bottlenecks and demands for improved data sets which are optimised for the mathematical modelling tools currently developed in the fields of bioinformatics and biological systems analysis.

The challenges of modern, multi-parallel, GC-MS based metabolite analysis are manifold: (i) automation of sample preparation, wet chemistry and data processing after acquisition for increased throughput and reproducibility, (ii) extension of the analytical scope of metabolomics studies, for example by combined analysis of single samples using multiple analytical technology platforms, and combined analysis with the proteome and transcriptome

---

<sup>1</sup> Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany, e-mail: Kopka@mpimp-golm.mpg.de

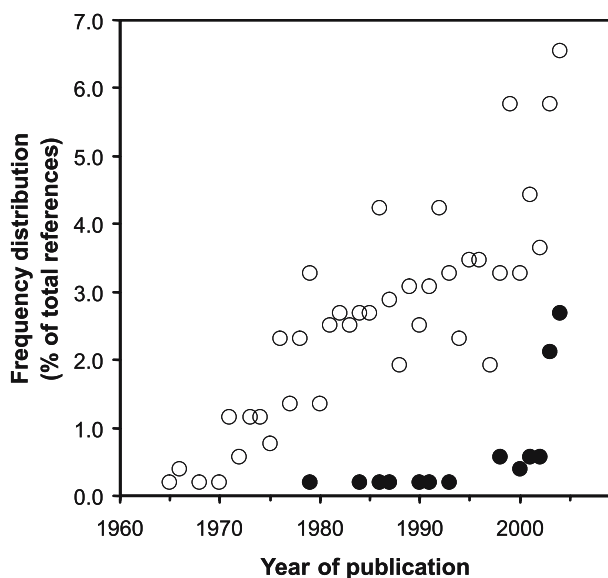


Fig. 1. Literature survey of publications which associate the concepts, “metabolite”, “profiling”, and “gas chromatography” performed on 1/2005. A total of ~500 citations without conference proceedings, abstracts and book chapters were found. The frequency of publications in all biological sciences (*open circles*) is compared to the contribution by plant metabolomics community (*closed circle*)

(Weckwerth et al. 2004b), (iii) profiling of trace compounds, or signalling molecules in the presence of bulk metabolites (Mueller et al. 2002; Birkemeyer et al. 2003; Schmelz et al. 2003, 2004), (iv) increasing accuracy in multi-parallel metabolite quantification (Birkemeyer et al. 2005), (v) combining profiling and flux analyses (Roessner-Tunali et al. 2004), (vi) establishment of quantitative repeatability, unambiguous nomenclature and comparability between analyses performed in different laboratories or using different analytical technology platforms (Schauer et al. 2005), and (vii) finally – perhaps the most important challenge of all metabolomic investigations – the identification of the unidentified majority of metabolic components from metabolite profiling experiments (Fiehn et al. 2000a; Schauer et al. 2005).

In agreement with the focus of this chapter the above challenges have predominantly analytic or technical motivation. The breakthrough of metabolomic investigations, however, will depend on the access to hitherto unavailable fundamental insights into metabolic and systems interactions. Increasingly integrative studies which consider the metabolome, proteome, transcriptome, and genome evolution of an organism have been initiated and are to be expected. Promising steps have been made – using GC-MS technology – towards network analysis (Fiehn 2003; Weckwerth et al. 2004a) and correlation studies between or within metabolome and transcriptome

constituents (Urbanczyk-Wochniak et al. 2003; Steinhäuser et al. 2004; Kopka et al. 2005). A detailed discussion of these general aspects including GC-MS studies and beyond can be found in the applications section of this book.

## 2 GC-MS Profiling Technology in a Nutshell

Metabolite profiling with GC-MS involves six general steps:

1. *Extraction* of metabolites from the biological sample, which should be as comprehensive as possible, and at the same time avoid degradation or modification of metabolites (e. g. Kopka et al. 2004).
2. *Derivatisation* of metabolites making them amenable to gas chromatography. Metabolites which are not volatile per se require chemical modification prior to GC analysis.
3. *Separation* by GC. High resolution GC can also be highly reproducible as it involves automated sample injection robotics, highly standardised conditions of gas-flow, temperature programming, and standardised capillary column material.
4. *Ionisation* of compounds as they are eluted from the GC. Electron impact (EI) ionisation is most widely used, as it is the technology which is least susceptible to suppression effects and produces reproducible fragmentation patterns.
5. Time resolved *detection* of molecular and fragment ions. Mass separation and detection can be achieved with different mass-detection devices, including sector field detectors, quadrupole detectors (QUAD), ion trap technology, and time-of-flight detectors (TOF). The choice of detectors depends on the targeted analytical niche. GC-MS systems with QUAD detection are most widely spread for routine analysis. Ion trap technology allows MS $\times$ MS (two-dimensional MS) analysis for structural elucidation and targeted quantification of trace compounds (e. g. Mueller et al. 2002). TOF detection can either be tuned to fast scanning rates (van Deursen et al. 2000) or to high mass precision comparable to sector field systems. Fast scanning GC-TOF-MS enables the, today, most advanced technology in the GC-MS field, namely two dimensional GC $\times$ GC-TOF-MS (two-dimensional GC-TOF-MS) (Ryan et al. 2004; Sinha et al. 2004a–c).
6. *Acquisition and evaluation* of GC-MS data files. All GC-MS system manufacturers provide software which is tuned for targeted, quantitative metabolite analysis. The targeted approach involves unequivocal identification of predefined metabolites by expected chromatographic retention times and mass-spectral fragmentation patterns and quantitative calibration by authentic standard concentrations. Recent software developments support the non-targeted analysis of GC-MS patterns, and the full evaluation of all resolved compounds. This feature of GC-MS allows discovery of novel hitherto

unknown metabolites. As we are far from knowing all possible metabolites of a given organism, non-biased, truly comprehensive data evaluation is the most essential requirement of metabolite profiling.

## 2.1 Chemical Derivatisation and Chromatography

The principles of fast metabolic sample inactivation and nondestructive extraction are common to all metabolome analyses. In contrast to all other technologies GC-MS is inherently restricted to volatile and temperature-stable compounds. The scope of GC-MS for metabolite analysis is limited by the typical temperature range of commercial capillary columns, for example up to 320–350 °C. The lower temperature range is determined by ambient temperature, but cold trapping devices and isothermal GC allow analysis of low molecular weight gases and highly volatile metabolites. GC received a considerable extension of applications through the development of a highly versatile tool box of derivatisation reagents, which chemically transform non-volatile metabolites into volatile analytes for GC-MS analysis (e. g. Knapp 1979; Blau and Halket 1993; Toyo'oka 1999). To date, GC-MS profiling of metabolites in plants has largely been confined to compounds, recovered in the methanol-water phase after methanol-water/chloroform extraction of tissues (Fiehn et al. 2000a; Roessner et al. 2000; Duran et al. 2003; Barsch et al. 2004; Gullberg et al. 2004; Strelkov et al. 2004; Broeckling et al. 2005). Although not all hydrophilic compounds can be volatilised by derivatisation, the following classes of compounds are detected routinely: amino-, organic-, and aromatic-acids, amines, sugars up to trisaccharides, alcohols and polyols, and some monophosphorylated metabolites.

The current limitations of metabolite preparation and derivatisation strategy, namely methoxyamination with subsequent direct trimethylsilylation of predominantly polar metabolites, call for extension. Application of other technology platforms is an obvious route and will be discussed in the following chapters. Here a short appraisal of the potential of chemical derivatisation is attempted. Four main types of reaction schemes will be discussed.

1. *Alkoxyamination* by reagents, such as methoxyamine  $\text{CH}_3\text{-O-NH}_2$ , stabilises carbonyl moieties in native metabolite structures, but forms E- and Z-isomers of the  $\text{-N=C<}$  double-bond substituents. Keto-enol tautomerism is suppressed, as is the decarboxylation of unstable  $\beta$ -carbonyl-carboxylic acids. In addition, the formation of acetal- or ketal-structures in aqueous solution is inhibited. These equilibrium reactions generate multiple intramolecular and water adducts, for example the typical  $\alpha$ - and  $\beta$ -conformers of reducing sugars. Ether- and ester-conjugates are mostly stable when exposed to methoxyamine reagent and maintain conformation. So far other alkoxy-reagents – for example hydroxylamine, ethoxyamine, or benzyloxyamine – have not been exploited for systematic discovery of metabolites with carbonyl moieties:

2. *Silylation* reagents classify into those which introduce either a trimethylsilyl (TMS) moiety,  $-\text{Si}(\text{CH}_3)_3$ , or a dimethyl-*(tert-butyl)*-silyl (TBS) moiety,  $-\text{Si}(\text{CH}_3)_2-\text{C}(\text{CH}_3)_3$ . TMS reagents have been well investigated and are known to have the widest derivatisation spectrum (Little 1999; Halket et al. 2005). TMS has the potential to substitute all exchangeable, “acidic” protons of a metabolite. Steric hindrance of TMS substitution is rare but common with the bulkier TBS reagent. The benefit of the TBS reagent is higher tolerance for the presence of water and clear mass spectral fragmentation. However, vicinal diols, which typically occur in sugars, are only partially derivatised.
3. *Alkylation* reactions, mostly methylation, are widely used to derivatise carboxylic acids and alcohols. The enormous reactivity of available reagents – some allow for flash derivatisation during hot GC injection – leads to transalkylation of ester-bonds and consequently breaks down complex metabolites, such as glycerol- and phospholipids. Alkylation of sugars leads to derivatives which are more volatile than the TMS derivatives and therefore allow analysis of higher sugar oligomers.
4. *Acylation* reactions, mostly acetylation or trifluoro-acetylation, are less reactive than transalkylation. Reagents usually form stable ester and amide bonds and break down only activated metabolic intermediates, e. g. thioesters.

In conclusion further developments of alternate GC-MS profiling techniques need to employ more selective combinations of metabolite fractionation and derivatisation schemes. Solid phase extraction can be explored to partition and concentrate metabolites amenable to alternate subsequent derivatisation. On the other hand, vapour phase extraction (VPE) for the separation and concentration of volatile derivatisation products prior to GC injection may prove promising (Schmelz et al. 2003, 2004). VPE has the potential to be a robust technique and was shown to operate with a range of commonly used reagents.

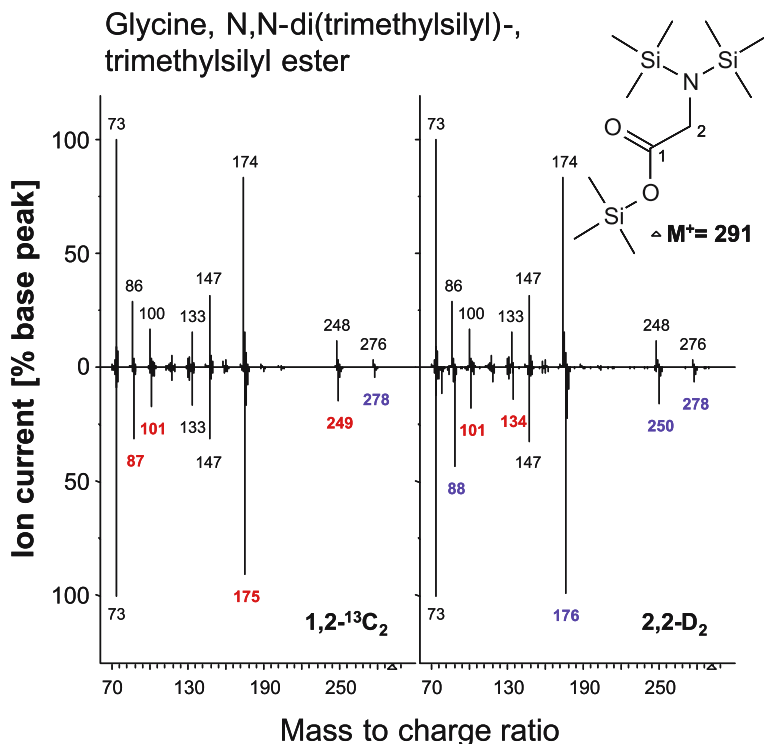
## 2.2 Mass Detection and Quantitative Calibration Techniques

One of the major criticisms and pitfalls of metabolome analyses is best explained by so-called matrix effects. This well-known effect describes unexpected losses or increased recovery of metabolites in complex extracts compared to pure authentic preparations. Matrix effects on one hand are caused by the presence of compounds which either specifically inhibit extraction or chemical analysis of metabolites. Positive matrix effects can stabilise otherwise labile compounds in the presence of suitable chemicals. Typical examples are suppression effects of soft ionization techniques, for example electrospray ionization (ESI) or matrix assisted laser desorption ionization (MALDI). Electron-impact ionization (EI) typically used in GC-MS profiling is not susceptible to suppression. Instead GC injection is the crucial step which may



cause discriminations, especially in view of the complex and rather crude extracts which are typically injected.

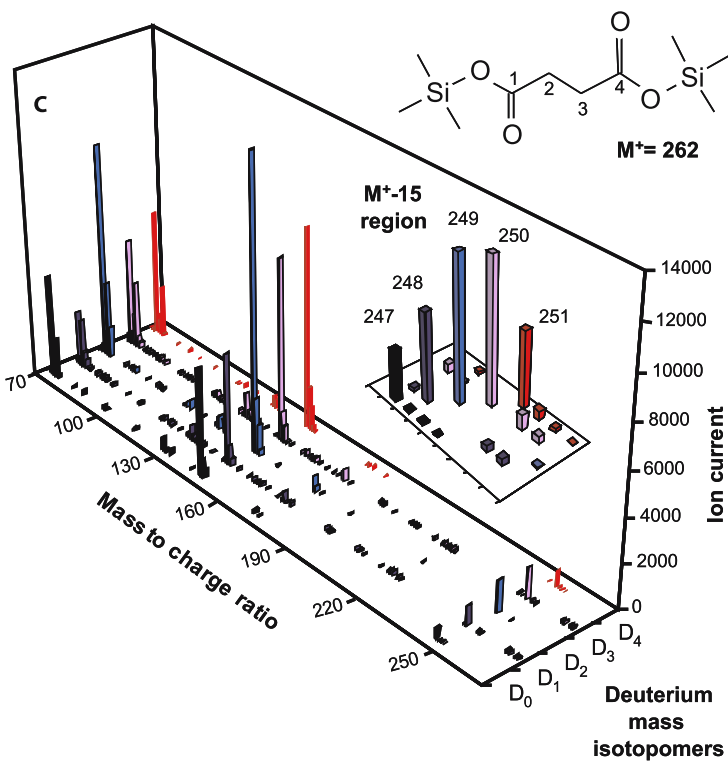
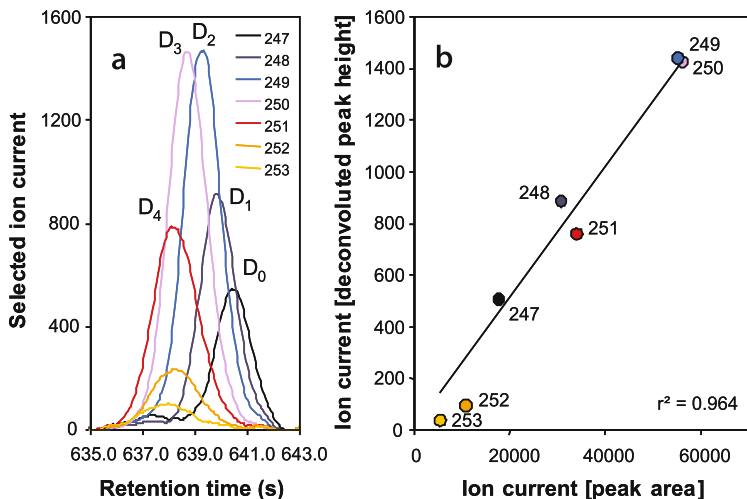
So far, only exemplary – albeit time demanding – thorough tests for unexpected matrix effects have been performed with selections of chemically



**Fig. 2.** Mass spectra of deuterated and  $^{13}\text{C}$  labeled MSTs help structural elucidation and recovery analysis of metabolites. Labeled and non-labeled MSTs of Glycine *N,N*-di(trimethylsilyl)-, trimethylsilyl ester are shown. *Oryza sativa* L. cv. Nipponbare was labelled in vivo using deuterated water or  $^{13}\text{CO}_2$ . MSTs representing the fully labeled mass isotopomers demonstrate presence of two carbon atoms (*left panel*) and two non-exchangeable hydrogen atoms (*right panel*). Mass fragments which exhibited a mass shift of 1 amu (*red*) or 2 amu (*blue*) are indicated

► **Fig. 3a–c.** Mass spectral deconvolution of deuterated mass isotopomers. Succinic acid di(trimethylsilyl) ester was partially labelled in vivo by exposing *Oryza sativa* L. cv. Nipponbare to deuterated water. Metabolite profiles were performed on a Pegasus II GC-TOF-MS system (LECO, St. Joseph, MI, USA) with 20 scans  $\text{s}^{-1}$ . Mass spectra were deconvoluted using ChromaTOF software version 1.00, with baseline offset just above noise, smoothing and peak width set to 10 and 2 scans, respectively: **a** selected ion traces of non-deuterated ( $\text{D}_0$ ,  $m/z = 247$ ) and deuterated ( $\text{D}_{1-4}$ ,  $m/z = 248 - 251$ )  $\text{M}^+ - 15$  mass fragments. Mass fragments at 252 and 253 amu are carbon mass isotopomers of  $\text{D}_4$ ; **b** peak area compared to deconvoluted peak height. Peak area integration does not allow differentiation of contributions by carbon mass isotopomers; **c** deconvoluted mass spectra of  $\text{D}_{0-4}$ . *Inset* shows partial deconvolution of  $\text{D}_{0-4}$  carbon mass isotopomers and missing carbon mass isotopomers of  $\text{D}_{1-3}$

diverse, representative metabolites (e. g. Roessner-Tunali et al. 2003; Gullberg et al. 2004). Therefore technologies are required to improve quantitative standardisation for the comparison of increasingly diverse biological samples and experimental conditions.



For this purpose, full saturating  $^{13}\text{C}$  in vivo labelling was developed using yeast which is one of the most important organisms in systems biology (e. g. Stephanopoulos et al. 2004). Metabolites of yeast were demonstrated to be fully labelled when provided with an exclusive carbon source, such as  $\text{U-}^{13}\text{C}$ -glucose (Mashego et al. 2004; Birkemeyer et al. 2005). Refer to Birkemeyer et al. (2005) for detailed discussion of potential applications for  $^{13}\text{C}$ -labelled metabolomes. Similar approaches are possible in plants (Figs. 2 and 3).

In short, standardised in vivo labelled extracts of yeast or other microorganisms can substitute the rather small number of chemically synthesised mass isotopomers used in earlier studies (Fiehn et al. 2000a; Gullberg et al. 2004). Typically a standardised labelled reference sample is combined in equal amounts with non-labelled experimentally challenged samples. The advantages of this approach are (i) the presence of a mass isotopomer for all identified but also all hitherto non-identified metabolites, (ii) the concentration of each mass isotopomer is inherently adjusted to the endogenous metabolite concentration, (iii) metabolic components can easily be distinguished from laboratory contaminations, and (iv) recovery of all metabolic components can be determined with the appropriate mass isotopomer.

Thus metabolite profiling will achieve the same level of transcriptome and proteome experiments, which utilize differential fluorescent probes or differential isotope coded tagging, respectively. In conclusion, comprehensive in vivo isotope labelling will help to establish quantitative between laboratory comparability of GC-MS based metabolome experiments. More importantly, we expect metabolome experiments with full mass isotopomer standardisation to be also independent of the mass spectrometric platform, e. g. CE-MS, LC-MS, or possibly even MALDI-TOF-MS.

### 3 Short Excursion into Nomenclature and Definitions

Concise and unambiguous description of GC-MS metabolite profiling results requires clear definitions. The definitions suggested within this section are biased towards the specifics of GC-MS technology but may also be applied to other technology platforms. This section is intended as a contribution to the ongoing process of unifying data formats and concepts within the field of plant metabolomics (e. g. Fiehn 2002; Bino et al. 2004; Jenkins et al. 2004).

#### 3.1 Metabolite and Analyte

Routine GC-MS profiling analysis (Fiehn et al. 2000b; Roessner et al. 2000) has an upper size exclusion limit which is roughly equivalent to a persilylated trisaccharide derivative (MW:1296), hexatriacontane (MW:506), or hentriacontanoic acid trimethylsilylester (MW:523). Even though it may appear tempting, metabolite and analyte are best not defined by molecular weight.

A *metabolite* may be described as a compound which is internalised, chemically converted or secreted by an organism, but is not synthesised by DNA replication, transcription, or translation. Post-processing events of DNA, RNA and proteins, such as DNA methylation, RNA splicing, sequence specific protease cleavage or post-translational modification are not attributed to the metabolome. The origin of a metabolite is not exclusively dependent on the biosynthetic capacity of an organism or delimited by the genomic inventory. Metabolites may readily be exchanged between organisms, for example in plant microbe interactions, and – like drugs or pesticides – can today be of anthropogenic/xenobiotic origin.

In contrast to LC- or CE-MS, GC-MS analysis requires clear distinction between metabolite and analyte, because – depending on choice of chemical derivatisation – metabolites may be chemically transformed before quantification. The term *analyte* may be used to address the chemical structure and compound which is submitted to GC-MS and finally detected and quantified. An analyte can be identical with the metabolite, if the metabolite is not chemically derivatised. Single metabolites may have more than one analyte, if the chosen derivatisation reaction generates more than one derivative, for example methoxyamination (see above). In these cases *preferred* and *alternate* analytes exist for quantification. Analytes of one metabolite may differ in abundance, i. e. a *major* and one, even multiple, *minor* analytes may exist. Standardisation by stable mass isotopomers corrects the quantification errors which may arise from unforeseen matrix effects on analyte ratios during chemical derivatisation of GC injection.

Different metabolites may be chemically transformed into the same analyte structure. In addition a single analyte may arise from inadequate chromatographic separation of isomers. For example, the biochemically distinct stereoisomeric structures of DL-amino acids are only separated by specialised chiroselective chromatography. These analytes have *composite* properties in contrast to absolutely *specific* analytes.

These concepts are not unique to GC-MS technology. Analyte sensitivity, accuracy, and potentially composite analyte properties need to be thoroughly considered in MS-MS applications, non-chiroselective capillary electrophoresis or liquid chromatography, and in cases of adduct-formation or multiply charged ions.

### 3.2 Mass Spectral Tag (MST) and Mass Fragment

GC-MS metabolite profiles resolve hundreds of analytes, which represent metabolites, but also internal standard substances and laboratory contaminations. Typical GC-MS profiles may contain approximately 100 identified analytes of metabolites. The chemical structure of the majority of GC-MS analytes, however, is still unknown. Each new biological object or experimental condition still gives rise to new, hitherto unidentified, chemical components.

Because in non-biased analysis of GC-MS profiles identified and unidentified components are equally important, we created the term *mass spectral tag* (MST), i. e. a mass spectrum which is characterised by a specific chromatographic retention and by repeated occurrence in a single or multiple types of biological samples (Colebatch et al. 2004; Desbrosses et al. 2005). MSTs represent analytes. MSTs can be identified, in other words, unequivocally linked to a chemical structure. The use of MSTs allows uncoupling of metabolite profiling experiments from the time consuming process of chemical identification. MSTs can be used to track analytes in different experiments or laboratories (Schauer et al. 2005). Thus MST identification can be performed even years after the first discovery.

MSTs of GC-EI-MS profiles are composed of multiple characteristic *mass fragments* in constant relative abundances. In most cases residual, non-fragmented molecular ions are rare or even absent. In consequence GC-MS allows selection of multiple mass fragments which all represent the same MST and exhibit the same quantitative changes. Typically one quantifying mass fragment (QM) and a set of specific, supporting qualifying mass fragments are selected in GC-MS analysis (Halket et al. 2005). The criteria for the proper choice of QMs are equal to the choice of a preferred analyte. QMs need to be selective, i. e. not composite, in the context of the complexity of co-eluting MSTs. Therefore, the best QM is the most abundant among the available selective mass fragments.

### 3.3 Response and Relative Quantification of Metabolite Pools

GC-MS metabolite profiling studies monitor relative changes in metabolite pool sizes and but also allow insight into flux, i. e. the dynamic turnover of metabolite pools or metabolite substructures (e. g. Fischer and Sauer 2003; Sauer 2004; Roessner-Tunali et al. 2004). Flux experiments are easily distinguished from above mentioned saturating *in vivo* labelling experiments. Flux experiments monitor the initial kinetics of labelling and thus stable isotopes are only partially incorporated into metabolite pools. In contrast, saturating *in vivo* labelling reaches the endpoint of a completely stable isotope labelled metabolome.

MSTs are quantified by ion currents of QMs which are recorded after analyte ionization, fragmentation and mass separation. Ion currents in GC-MS are monitored either by peak area or peak height. Both measurements need to be baseline corrected for electronic and chemical noise. The resulting corrected values are defined to be what we call *responses*, i. e.  $X_{QM}$  of fragment QM (Colebatch et al. 2004; Desbrosses et al. 2005). The fragment response is routinely normalised to the amount of the sample, for example fresh or dry weight. In addition each response is corrected for recovery effects, which may occur at any step of the analytical process between metabolic inactivation of the sample and final recording of ion currents. Different levels of recovery correction exist: (i) correction by extract and sample volume, (ii) correction

by addition of a constant amount of a representative internal standard compound (IS), and (iii) normalization by chemically identical, but stable-isotope labelled mass isotopomers of each metabolite. The *normalised response* ( $N_{QM}$ ) is, consequently,  $N_{QM} = X_{QM} \times X_{IS}^{-1} \times \text{sample weight}^{-1}$ , where  $X_{IS}$  ideally represents a mass isotopomer response of QM. In a further step, the normalised response of a fragment,  $N_{QM}$ , is divided by the average relative response of QM as determined in a set of reference samples,  $\text{avg}N_{QM(\text{ref})}$ . The resulting quotient,  $R_i = N_{QM} \times \text{avg}N_{QM(\text{ref})}^{-1}$ , is called *response ratio*  $R_i$ .  $R_i$  describes the x-fold changes in metabolite pools sizes relative to the reference samples. Typical reference samples are taken at the start of a time series experiment or are mock-treated biological controls.

In GC-MS profiling analyses the standard deviation of normalised responses is dependent on the chemical nature of metabolite and analyte. Average relative standard deviations (RSD) of 10% (Weckwerth et al. 2004b) or 13.8% (5.5–33.4%; Gullberg et al. 2004) were reported for replicate GC-MS analyses. These analyses included extraction as well as derivatisation and were performed using representative analytes. Use of isotope labelled standardisation was reported to reduce RSD further to approximately 6.9–9.7% residual experimental variance (Gullberg et al. 2004).

## 4 Present Challenges of GC-MS Profiling

### 4.1 Standardisation of GC-MS Systems

GC-MS profiles, with the exception of GC×GC-TOF-MS data, are in essence three-dimensional and comprise a chromatographic time-resolved axis, a second coordinate axis which represents the mass to charge ratio ( $m/z$ ,  $z = 1$  in GC-MS with only rare exceptions), and an intensity axis which monitors the ion current (IC) and thus the abundance of molecules or mass fragments. A substantial breakthrough for GC-MS analyses was the early establishment of generally accepted calibration substances and procedures, so-called tuning routines, which allowed comparison of mass spectra from GC-MS systems of virtually all manufacturers and from different hyphenated mass detection technologies. In addition the widely used electron-impact ionisation technique (EI) ensured stable fragmentation ratios, which are in first approximation independent of analyte concentration. However, comparability was only achieved by restriction to 1 amu precision.

The chromatography axis is less standardised, not least because of multiple types of available capillary GC-columns which have different chromatographic properties and thus serve different separation problems. In addition slight changes in temperature program, pressure and flow settings of both carrier gas and injection technique, as well as slight production differences of column manufacturers cause minor but perceptible changes in retention