

Contents

Preface XIX

List of Contributors XXI

I	Transcriptome Analysis	1
A	Whole Genome Expression Analysis	1
1	Single Cell Expression Profiling: Transcript and Protein Analyses in Isolated Higher Plant Gametes and Zygotes	3
	<i>Stefan Scholten and Erhard Kranz</i>	
1.1	Introduction	3
1.2	Microdissection, Cell Isolation	6
1.3	<i>In Vitro</i> Fertilization	6
1.4	Techniques for Molecular Analyses of Single Cell Types	7
1.4.1	Sampling of Single, Living Cells	7
1.4.2	Analyses of Gene Expression	8
1.4.2.1	Single Cell Gene-by-Gene Analysis	8
1.4.2.2	Amplification of Whole cDNA Populations	9
1.4.2.3	Quantification of Transcript Levels	11
1.4.2.4	Library Construction and EST Sequencing	12
1.4.2.5	Targeted Approaches Using cDNA Subtraction	12
1.4.2.6	Microarray Analyses	13
1.5	Analyses of Protein Expression	14
1.6	Prospects	15
	References	16
2	AFLP-Based RNA Fingerprinting: Novel Variants and Applications	21
	<i>Christian W.B. Bachem, Wim H. Vriezen, Craita E. Bitá, and Asun Fernandez del Carmen</i>	
2.1	Introduction	21
2.2	Methods and Protocols	23

2.2.1	Theoretical Considerations	23
2.2.2	State-of-the-Art cDNA-AFLP Protocol	24
2.2.2.1	Isolation of cDNA Fragments	24
2.2.2.2	Non-Selective Pre-Amplification	25
2.2.2.3	Selective Amplification-Reaction Using ^{33}P -Labeled Primer and Gel Analysis	26
2.2.2.4	Downstream Analysis	29
2.3	Applications of the Technology	31
2.3.1	Fruit Development	31
2.3.2	Tuber Development	32
2.3.3	Transcript BSA	32
2.3.4	Domain Profiling	33
2.3.5	VIDISCA	33
2.4	Perspectives	34
	References	34

3 SuperSAGE: The Most Advanced Transcriptome Technology for Functional Genomics 37

Ryohei Terauchi, Hideo Matsumura, Detlev H. Krüger, and Günter Kahl

3.1	Introduction	37
3.2	Methods and Protocols	40
3.2.1	Linker Preparation	41
3.2.2	RNA Sample	42
3.2.3	cDNA Synthesis	43
3.2.4	Tag Extraction from cDNA	43
3.2.5	Purification of Linker-Tag Fragment	44
3.2.6	Ditag Formation and Amplification	44
3.2.7	Tag Extraction from Sequence Data	45
3.3	Applications of the Technology	47
3.3.1	Interaction Transcriptome	47
3.3.2	Application of SuperSAGE to Non-Model Organisms	47
3.3.3	SuperSAGE-Array	48
3.3.4	GMAT	49
3.4	Perspectives	49
	References	51

4 From CAGE to DeepCAGE: High-Throughput Transcription Start Site and Promoter Identification for Gene Network Analysis 55

Matthias Harbers, Thomas Werner, and Piero Carninci

4.1	From Genomes to Transcriptomes	55
4.2	Addressing the Complexity of Transcriptomes	56
4.3	The Shift From CAGE to DeepCAGE	57
4.4	Applications of CAGE and DeepCAGE Libraries	58
4.5	Preparation of a DeepCAGE Library	59
4.6	CAGE Data Analysis and Genome Mapping Approaches	62

4.7	Expression Profiling: Putting CAGE Tags into a Biological Context	67
4.8	Perspectives	68
	References	70
5	Gene Identification Signature-Paired End diTagging (GIS-PET): A Technology for Transcriptome Characterization	77
	<i>Patrick Ng, Yen-Ling Lee, Chia-Lin Wei, and Yijun Ruan</i>	
5.1	Introduction	77
5.1.1	Microarray Analysis	79
5.1.2	cDNA Sequencing, Including EST- and flcDNA-Sequencing	79
5.1.3	DNA-Tagging Methods	80
5.1.4	Advanced DNA Sequencing Technologies	81
5.2	Protocol	82
5.2.1	Construction of a GIS-PET flcDNA Library	84
5.2.1.1	Reverse-Transcription of mRNA (polyA-RNA) Sample	84
5.2.1.2	Oxidation	86
5.2.1.3	Biotinylation of RNA Ends	86
5.2.1.4	RNaseONE Selection for Full-Length (–) cDNA/RNA Heteroduplex	86
5.2.1.5	Binding Biotinylated (–) cDNA/RNA Heteroduplex to Streptavidin Beads	87
5.2.1.6	Alkaline Hydrolysis to Release (–) Strand flcDNA	88
5.2.1.7	Double-Stranded cDNA (ds cDNA) Synthesis	88
5.2.1.8	Further Processing of ds flcDNA	89
5.2.1.9	cDNA Size Fractionation	90
5.2.1.10	Cloning of flcDNA in pGIS4a Vector	90
5.2.1.11	Perform QC on flcDNA Library	92
5.2.2	Construction of a Single-PET Library	92
5.2.2.1	Plasmid DNA Preparation	92
5.2.2.2	Tagging by Mme I Digestion	93
5.2.2.3	Intramolecular Circularization to Create Single-PET Plasmids	93
5.2.2.4	Transform Cells	94
5.2.2.5	Perform QC on GIS Single-PET Library	94
5.2.3	Construction of a GIS-PET Sequencing Library for Sanger Sequencing of Ditags	95
5.2.3.1	Single-PET Plasmid DNA Preparation	95
5.2.3.2	Bam HI-Digestion to Release Single-PETs	95
5.2.3.3	PAGE-Purification of 50-bp BamHI-Cohesive Single PETs	96
5.2.3.4	PET Concatenation	97
5.2.3.5	Purification of Concatenated PETs	97
5.2.3.6	Cloning Concatenated PETs in pZErO-1 Vector	98
5.2.3.7	Transform Cells	99
5.2.3.8	Carry out QC on GIS-PET Sequencing Library	99

5.2.4	Construction of diPETs for 454-Sequencing	100
5.2.4.1	Single-PET Plasmid DNA Preparation	100
5.2.4.2	Bse RI Linearization of Single-PET Plasmid DNA	100
5.2.4.3	BamHI Digestion to Release Asymmetric PETs	101
5.2.4.4	Recovery and Quantitation of Purified Asymmetric PETs	101
5.2.4.5	Formation of diPETs	102
5.3	Data Analysis	103
5.4	Discussion	103
5.5	Perspectives	105
	References	109
6	High-Throughput Functional Screening of Genes <i>In Planta</i>	113
	<i>Thomas Berberich, Yoshihiro Takahashi, Hiromasa Saitoh, and Ryohei Terauchi</i>	
6.1	Introduction	113
6.2	Methods and Protocols	115
6.2.1	Extraction of Total RNA	116
6.2.2	Purification of Poly(A) ⁺ -mRNA	119
6.2.3	Synthesis of cDNA and Ligation to Binary, PVX-Based Expression Vectors	121
6.2.4	Amplification of the cDNA Library in <i>E. coli</i>	124
6.2.5	Transformation of cDNA Library into <i>Agrobacterium tumefaciens</i> Cells	126
6.2.5.1	Preparation of Competent <i>Agrobacterium tumefaciens</i> Cells	126
6.2.6	Toothpick Inoculation of Leaves	128
6.2.7	Agroinfiltration	129
6.2.8	Recovery of the cDNA Fragments	130
6.3	Application of the Technology	131
6.4	Perspectives	133
	References	134
7	Microarrays as Tools to Decipher Transcriptomes in Symbiotic Interactions	137
	<i>Helge Küster and Anke Becker</i>	
7.1	Introduction	137
7.2	Methods and Protocols	143
7.2.1	Spotting and Storage of Mt16kOLI1/Mt16kOLI1Plus 70mer Oligonucleotide Microarrays	143
7.2.2	Synthesis of Targets by Indirect Reverse Transcription Cy-Labeling	144
7.2.2.1	Components Stored at −20 °C	144
7.2.2.2	Components Stored at 4 °C (−20 °C after Aliquoting in 1/10 Volumes)	145
7.2.2.3	Components Stored at Room Temperature	145

7.2.2.4	Reverse Transcription of Total RNA to obtain Aminoallyl-Labeled First-Strand cDNA	145
7.2.2.5	Hydrolysis of RNA	146
7.2.2.6	Clean-Up of Aminoallyl-Labeled First-Strand cDNA	146
7.2.2.7	Coupling of Fluorescent Dyes to Aminoallyl-Labeled First-Strand cDNA	147
7.2.2.8	Quenching of all Remaining NHS Esters	147
7.2.2.9	Clean-up of Fluorescently Labeled Targets	147
7.2.2.10	Quality Control of Fluorescently Labeled Targets	148
7.2.3	Pre-Processing, Hybridization and Scanning of Mt16kOLI1/Mt16kOLI1Plus Microarrays	148
7.2.4	Handling and Evaluation of Microarray Data	150
7.3	Applications of the Technology	151
7.3.1	Microarray-Based Identification of <i>Medicago truncatula</i> Genes Induced during Different Arbuscular Mycorrhizal Interactions	151
7.3.2	Microarray-Based Identification of <i>Medicago truncatula</i> Genes Activated during Nodulation and Mycorrhization	154
7.4	Perspectives	156
	References	157

B Gene-by-Gene Analysis 163

8	Genome-Wide Analysis of mRNA Expression by Fluorescent Differential Display	165
	<i>Suping Zhou, Jonathan D. Meade, Samuel Nahashon, Blake R. Shester, Jamie C. Walden, Zhen Guo, Julia Z. Liang, Joshua G. Liang, and Peng Liang</i>	
8.1	Introduction	165
8.2	Methods and Protocols	169
8.2.1	Materials	169
8.2.1.1	Total RNA Isolation and Removal of Genomic DNA from Total RNA	169
8.2.1.2	Single-Strand cDNA Synthesis by Reverse Transcription	170
8.2.1.3	Fluorescent Differential Display-PCR (FDD-PCR)	170
8.2.1.4	Gel Electrophoresis	170
8.2.1.5	Reamplification of Selected Differentially Expressed Bands	171
8.2.1.6	Cloning of Reamplified PCR Products	171
8.2.1.7	Verification of Cloned PCR Products	171
8.2.1.8	Confirmation of Differential Gene Expression by Northern Blot	171
8.2.2	Methods	172
8.2.2.1	Total RNA Isolation and Removal of Genomic DNA	172
8.2.2.2	Gel Preparation	174
8.2.2.3	RNA Loading Sample Preparation	174
8.2.2.4	Single-Strand cDNA Synthesis by Reverse Transcription	174
8.2.2.5	Fluorescent Differential Display-PCR	175

8.2.2.6	Gel Electrophoresis	176
8.2.2.7	Reamplification of Selected Differentially Expressed cDNA Bands	177
8.2.2.8	Cloning of Reamplified PCR Products	179
8.2.2.9	Verification of the Cloned Inserts	180
8.2.2.10	Confirmation of Differential Gene Expression by Northern Blot	181
8.3	Applications of the Technology	183
8.4	Perspectives	184
	References	185

9 Real-Time Quantitation of MicroRNAs by TaqMan MicroRNA Assays 187

Toni L. Ceccardi, Marianna M. Goldrick, Peifeng Ren, Rick C. Conrad, and Caifu Chen

9.1	Introduction	187
9.1.1	What are microRNAs?	187
9.1.2	Why are Researchers Interested?	188
9.1.3	Current Technologies for miRNA Quantitation	189
9.2	Methods and Protocols	190
9.2.1	Bioinformatic Tools for miRNA Discovery	190
9.2.2	MicroRNA Isolation from Plants	191
9.2.2.1	Extraction from Plant Tissue	192
9.2.2.2	PCR Directly from Cells in Culture	193
9.2.3	Description of TaqMan MicroRNA Assays	194
9.2.3.1	Principle of TaqMan MicroRNA Assays	194
9.2.3.2	Performing the TaqMan MicroRNA Assay	195
9.2.4	Data Normalization	196
9.3	Applications of the Technology	197
9.3.1	Quantitation of miRNAs	197
9.3.2	Absolute Quantitation of miRNAs	198
9.3.3	Expression Profiling of miRNAs	198
9.3.4	Verification of Predicted Novel miRNAs	199
9.3.5	MicroRNAs in Plant Growth and Development	199
9.3.6	Discovery of miRNA Biomarkers	201
9.3.7	Discovery and Validation of Plant miRNA Targets	201
9.4	Perspectives	202
	References	203

II Gene Silencing, Mutation Analysis and Functional Genomics 207

10 RNA Interference 209

Chris A. Brosnan, Emily J. McCallum, José R. Botella, and Bernard J. Carroll

10.1	Introduction	209
10.2	Methods and Protocols	213

10.3	Applications of the Technology	216
10.3.1	Targeting Transgenes in <i>Arabidopsis</i> using RNAi	216
10.3.2	Tissue-Specific RNA Silencing of an Endogenous Gene in Tobacco	218
10.3.3	Advantages of Using RNAi in Plant Functional Genomics	219
10.4	Perspectives	219
	References	220
11	Extending Functional Genomics: VIGS for Model and Crop Plants	227
	<i>Steven Bernacki, John Richard Tuttle, Nooduan Muangsan, and Dominique Robertson</i>	
11.1	Introduction	227
11.2	Methods and Protocols	232
11.2.1	Constructing Geminivirus VIGS Vectors	232
11.2.1.1	Structure of Geminivirus Plasmids	233
11.2.1.2	Construction of an AR1 Replacement Vector	234
11.2.1.3	Construction of an Insertion Vector	235
11.2.2	Silencing an Endogenous Gene	236
11.2.2.1	Visible Markers for Testing and Optimizing VIGS	236
11.2.2.2	Cloning a Target Gene Fragment(s) into the CaLCuVA:ChlI Vector	237
11.2.2.3	Plant Preparation	237
11.2.2.4	Microprojectile Bombardment	238
11.2.2.5	Assessment of VIGS	239
11.3	Applications of the Technology	241
11.4	Perspectives	242
	References	243
12	TILLING: A Reverse Genetics and a Functional Genomics Tool in Soybean	251
	<i>Khalid Meksem, Shiming Liu, Xiao Hong Liu, Aziz Jamaï, Melissa Goellner Mitchum, Abdelhafid Bendahmane, and Tarik El-Mellouki</i>	
12.1	Introduction	251
12.2	Methods and Protocols	252
12.2.1	Production of Suitable Mutant Population for TILLING	252
12.2.2	DNA Extraction and DNA Construction of Pools	253
12.2.3	TILLING Screening for Mutations	254
12.2.3.1	Gene-Specific Primers for TILLING	254
12.2.3.2	PCR Amplification and Heteroduplex Formation	255
12.2.3.3	M13-Tailed PCR Amplification for TILLING	255
12.2.3.4	Endonuclease Digestion and Purification of the Amplified DNA	256
12.2.3.5	Gel Electrophoresis and Image Analysis	257
12.3	Applications of TILLING to Soybean	258

12.3.1	Mutation Discovery, Density and Distribution in two Mutagenized Soybean Populations	258
12.3.2	Confirmation and Segregation Patterns of TILLING Mutations in Soybean	260
12.4	Discussion and Perspectives	262
	References	264
13	Transposon Tagging in Cereal Crops	267
	<i>Liza J. Conrad, Kazuhiro Kikuchi, and Thomas P. Brutnell</i>	
13.1	Insertional Mutagenesis in Plants	268
13.2	Transposon Tagging in Maize	270
13.2.1	Mutator Insertional Mutagenesis	270
13.2.2	Activator/Dissociation Mutagenesis	274
13.2.2.1	Forward Genetics	274
13.2.2.2	Reverse Genetics in Maize Using Ds	275
13.3	Large-Scale Reverse Genetics in Rice	275
13.3.1	<i>Tos17</i> in Rice	276
13.3.2	The Maize <i>Ac/Ds</i> Transposons in Rice	277
13.3.3	<i>En/Spm</i>	279
13.4	<i>Ac/Ds</i> Transposon Tagging in Barley	280
13.5	Future Direction of Tagging in Cereals	281
13.5.1	Potential for an Endogenous Candystripe1 Tagging System in Sorghum	281
13.5.2	Transposon-Mediated Deletions in Maize	281
13.5.3	Future Tagging Resources in Rice	282
13.5.4	Saturation Mutagenesis	282
	References	282
14	Fast Neutron Mutagenesis for Functional Genomics	291
	<i>Christian Rogers and Giles Oldroyd</i>	
14.1	Introduction	291
14.1.1	Advantages of Fast Neutron Mutagenesis	291
14.1.2	Features of Fast Neutron Mutagenesis	292
14.1.3	Fast Neutron Mutagenesis for Reverse Genetics	294
14.2	Methods and Protocols	294
14.2.1	Screening Strategies	294
14.2.2	Automation	297
14.2.3	Establishing the Populations	298
14.2.4	Pooling Strategies	299
14.2.5	Characterization of the Populations	301
14.3	Applications of the Technology	301
14.3.1	Targeting Small Genes	302
14.3.2	Deletions Can Span Multiple Genes	303
14.3.3	Fast Neutron Reverse Genetics for Crop Improvement	303

14.4	Perspectives	304
	References	304
III	Computational Analysis	307
15	Bioinformatics Tools to Discover Co-Expressed Genes in Plants	309
	<i>Yoshiyuki Ogata, Nozomu Sakurai, Nicholas J. Provart, Dirk Steinhauser, and Leonard Krall</i>	
15.1	Introduction	309
15.2	The Expression Angler Tool of the Botany Array Resource	311
15.2.1	Methods and Protocols	311
15.2.2	Applications of the Technology	317
15.2.3	Perspectives	319
15.3	The CSB.DB Tool	320
15.3.1	Introduction	320
15.3.2	Methods and Protocols	321
15.3.2.1	Simple Protocol for the Use of CSB.DB	322
15.4	The KaPPA-View 2: Co-Expression Analysis on the Plant Metabolic Pathway Maps	324
15.4.1	Introduction	324
15.4.2	Application of the Technology	324
15.4.3	Methods and Protocols	326
15.4.4	Perspectives	328
15.5	The KAGIANA Tool for Co-Expression Network Analysis of <i>Arabidopsis</i> Genes	328
15.5.1	Introduction	328
15.5.2	Methods and Protocols	329
15.5.2.1	Initial Setting	329
15.5.2.2	Retrieval of Co-Expressed Genes	329
15.5.2.3	The Other Tools of KAGIANA	330
15.5.3	Perspective	331
	References	331
16	AthaMap, a Database for the Identification and Analysis of Transcription Factor Binding Sites in the <i>Arabidopsis thaliana</i> Genome	337
	<i>Reinhard Hehl</i>	
16.1	Introduction	337
16.2	Methods and Applications	339
16.2.1	Using the Web Interface at http://www.athamap.de/	339
16.2.1.1	The Search Function	339
16.2.1.2	Co-Localization Analysis	342
16.2.1.3	Gene Analysis	343
16.2.1.4	External Links	344
	References	345

17	Structural Phylogenomic Inference of Plant Gene Function	347
	<i>Nandini Krishnamurthy, Jim Leebens-Mack, and Kimmen Sjölander</i>	
17.1	Introduction	347
17.2	Challenges in Protein Function Prediction	349
17.2.1	Gene Duplication	350
17.2.2	Domain Shuffling	350
17.2.3	Speciation	352
17.2.4	Propagation of Existing Annotation Errors	352
17.3	The Nomenclature of Homology	352
17.4	Structural Phylogenomic Inference of Function	354
17.5	Recommended Protocols for a Structural Phylogenomic Pipeline	356
17.5.1	Step 1: Homolog Selection	357
17.5.2	Step 2: Constructing and Analyzing a Multiple Sequence Alignment	358
17.5.3	Step 3: Constructing and Analyzing a Phylogenetic Tree	359
17.5.4	Step 4: Predicting Function using a Phylogenetic Tree	361
17.6	Web Servers and Databases useful in Phylogenomic Inference	362
17.7	Discussion	363
	References	364
18	Structural, Functional, and Comparative Annotation of Plant Genomes	373
	<i>Françoise Thibaud-Nissen, Jennifer Wortman, C. Robin Buell, and Wei Zhu</i>	
18.1	Introduction	373
18.2	Methods, Protocols, and Applications	374
18.2.1	Structural Annotation	374
18.2.1.1	Cognate Transcript Sequences are the Most Reliable Data Available	378
18.2.1.2	<i>Ab initio</i> Gene Finders are Good ORF Finders	379
18.2.1.3	Integrated Approaches are Ideal Solutions for the Automated Gene Prediction	379
18.2.1.4	Gene Prediction is an Iterative Process	380
18.2.1.5	Manual Curation is still an Indispensable Process in Gene Prediction	380
18.2.1.6	Other Considerations for Gene Prediction	380
18.2.2	Functional Annotation	381
18.2.2.1	Sequence Similarity	381
18.2.2.2	Domain Searches	382
18.2.2.3	Phylogenomics	382
18.2.2.4	Expression Data	383
18.2.3	Comparative Annotation	383
18.2.3.1	Comparative Annotation Using Transcripts	383
18.2.3.2	Comparative Genomics Using Genome Sequences	385
18.2.3.3	Algorithms for Comparative Genomics	386
18.3	Perspectives	387
	References	389

19 Large-Scale Genomic Sequence Comparison and Gene Identification with ClustDB 397

Jürgen Kleffe

- 19.1 Introduction 397
- 19.2 Methods and Protocols 401
 - 19.2.1 Reading Sequences 401
 - 19.2.2 Substring Clusters 402
 - 19.2.3 Maximally Extended Pairs of Common Substrings 405
 - 19.2.4 Match Extension with Errors 406
 - 19.2.5 Complete Matches 407
 - 19.2.6 Reference Query Problems 407
 - 19.2.7 Complementary Sequences 408
 - 19.2.8 Handling Ambiguity Letter Codes 408
 - 19.2.9 Sequence Clusters 409
 - 19.2.10 Memory Analysis 410
- 19.3 Applications 410
 - 19.3.1 Deriving Clusters of Identical Plant ESTs 411
 - 19.3.2 Deriving Substring Clusters for All Plant ESTs 412
 - 19.3.3 Checking the TIGR Medicago BAC Assembly 412
- 19.4 Perspectives 413
- References 415

IV Functional Genomics and Emerging Technologies 417

20 Nanotechnologies and Fluorescent Proteins for *in planta* Functional Genomics 419

C. Neal Stewart Jr.

- 20.1 Introduction 419
- 20.2 Green Fluorescent Protein 420
- 20.3 Protocol: Seeing GFP in Transgenic Plants 424
- 20.4 Nanotechnology for Monitoring Gene Expression 425
 - 20.4.1 Aptamers and Quantum Dots 425
 - 20.4.2 Molecular Beacons 425
 - 20.4.3 Split GFP Tagging and Detection 426
- 20.5 Barriers to Implementation 427
- 20.6 Conclusions 427
- References 428

21 New Frontiers in Plant Functional Genomics Using Next Generation Sequencing Technologies 431

Robert C. Nutter

- 21.1 Introduction 431
 - 21.1.1 Advent of Massively Parallel Sequencing Systems 431
 - 21.1.2 Overview of the Sequencing by Synthesis System 432

21.1.3	Overview of Single Base Extension System	433
21.1.4	Overview of the (SOLiD) System	433
21.2	Library Generation	434
21.3	Emulsion PCR	436
21.3.1	Bead Purification	436
21.3.2	Bead Deposition	437
21.4	Sequencing by Ligation	437
21.5	Base Calling	439
21.6	Potential Applications	440
21.6.1	Resequencing	440
21.6.2	<i>De novo</i> Sequencing	440
21.6.3	Gene Expression via Sequence Tags	441
21.6.4	Other Tag-Based Applications	443
21.7	Conclusions	444
	References	444

22 **454 Sequencing: The Next Generation Tool for Functional Genomics** 447

Lei Du, Jan Frederik Simons, Maithreyan Srinivasan, Thomas Jarvie, Bruce Taillon, and Michael Egholm

22.1	Introduction	448
22.2	Methods and Protocols	450
22.2.1	DNA Library Preparation	450
22.2.1.1	Option A: Nebulized Library Procedure	451
22.2.1.2	Option B: Amplicon Library Procedure	454
22.2.2	Emulsion PCR	456
22.2.3	Loading of PTP and Instrument Run	460
22.2.4	Data Analysis	467
22.2.4.1	Whole Genome Assembly	467
22.2.4.2	Resequencing and Mutation Detection	469
22.2.4.3	Ultra-deep Sequencing	469
22.3	Applications of the Technology	470
	References	472

Glossary 477

Index 537