

DNA Microarrays

Current Applications

Edited by **Emanuele de Rinaldis**
and **Armin Lahm**

Contents

List of contributors	vii
Preface	xi
1 Applications for DNA Microarrays: An Introduction	1
Emanuele de Rinaldis	
Introduction 1	
DNA microarray technologies 2	
A typical microarray experiment 7	
Applications of microarrays 8	
– <i>Expression profiling</i> 8	
– <i>Analysis of genomic DNA</i> 12	
Challenges in using DNA microarrays 14	
– <i>Noise</i> 14	
– <i>Normalization</i> 15	
– <i>Sample preparation</i> 15	
– <i>Experimental design</i> 15	
– <i>Large number of genes</i> 16	
– <i>Biological factors</i> 16	
– <i>Array quality assessment</i> 16	
Microarray data analysis 16	
– <i>Unsupervised clustering methods</i> 18	
– <i>Similarity distances</i> 19	
– <i>Supervised clustering methods</i> 22	
Public resources for expression data 23	
– <i>Stanford Microarray Database (SMD)</i> 23	
– <i>Gene Expression Omnibus (GEO)</i> 23	
– <i>Microarray Gene Expression Data (MGED)</i> 23	
– <i>ArrayExpress</i> 23	
– <i>Expression Profiler</i> 24	
Book's content 24	

2	Gene Networks and Application to Drug Discovery	29
	A. Ambesi, M. Bansal, G. Della Gatta, and Diego di Bernardo	
	Introduction 29	
	Classifier-based algorithms 31	
	– <i>Target identification and validation</i> 32	
	– <i>Hit identification, lead identification, and optimization: mode of action (MOA)</i> 34	
	– <i>Hit identification, lead identification, and optimization: efficacy and toxicity</i> 37	
	Gene network reconstruction 39	
	– <i>Target validation, hit identification, lead identification, and optimization</i> 45	
	Discussion 49	
3	Pathway Analysis of Microarray Data	55
	Matteo Pellegrini and Shawn Cokus	
	Introduction 55	
	Term enrichment analysis 56	
	Gene set analysis 57	
	Pathway coherence 59	
	Reconstruction of networks using expression data 61	
	Integrated pathway analysis of expression data and transcription factor binding data 65	
	Conclusions 67	
	Future developments and trends 67	
4	Toxicogenomics: Applications of Genomics Technologies for the Study of Toxicity	71
	Uwe Koch	
	Introduction 71	
	Toxicogenomics in the preclinical testing funnel 73	
	Mechanistic toxicogenomics 73	
	Phenotypic anchoring 74	
	Examples 76	
	– <i>Proteomics</i> 76	
	– <i>Metabonomics</i> 77	
	Predictive toxicogenomics 78	
	– <i>Example</i> 78	
	Data analysis 79	
	<i>In vitro</i> methods 80	
	– <i>Example</i> 80	
	Absorption 80	
	Genotoxicity 81	
	Integration 81	
	– <i>Computational tools</i> 81	
	– <i>Data</i> 82	

John C. Castle, Chris J. Roberts, Chun Cheng, and Jason M. Johnson

Introduction 89

Primary uses 90

- *Where is my gene expressed?* 90
- *What genes are expressed in my tissue?* 91
- *What groups of gene are specific to my tissue?* 91
- *Biomarkers: which genes are expressed specifically in condition X?* 94

Additional uses 95

- *Tissue correlation comparisons* 95
- *Gene correlation comparisons* 95
- *Species correlation comparisons* 97
- *Transcription factors and targets* 97

Conclusion 99

6 Supervised Classification of Genes and Biological Samples 101

Adrian Tkacz, Leszek Rychlewski, Paolo Uva, and Dariusz Plewczynski

Introduction 101

Nearest shrunken centroids (NSCs) 103

k-Nearest neighbor 104

Support vector machine (SVM) 107

Supervised neural networks 109

Random forests (RFs) 111

Conclusions 112

Acknowledgements 112

Appendix A: Classification algorithms 112

- *Support vector machine (SVM)* 112
- *Nearest shrunken centroids (NSCs)* 113
- *k-Nearest neighbor (kNN)* 114

Appendix B: Analysis software systems 116

- *Gene Cluster 2.0* 116
- *PAM* 116
- *Genesis* 117
- *Stuttgart Neural Network Simulator (SNNS)* 117

7	A Case Study: The Mammary Carcinogenesis in HER2 Transgenic Mice	121
	Federica Cavallo, Guido Forni, Anna Grassi, PierLuigi Lollini, and Raffaele Calogero	
	Introduction 121	
	– <i>The HER2 oncogene</i> 122	
	– <i>The BALB-neuT breast carcinoma murine model</i> 122	
	– <i>Halting BALB-neuT carcinoma by HER2 vaccination</i> 123	
	Vaccination effects and transcriptional profiling 124	
	– <i>Triplex vaccination</i> 124	
	– <i>Prime-and-boost vaccination</i> 126	
	Deriving tumor-associated antigens (TAAs) from BALB-neuT breast carcinoma transcription profiles 127	
	Bioinformatic approaches for microarray data analysis 129	
	– <i>Oligonucleotide chips</i> 129	
	– <i>GeneChips data analysis</i> 130	
	Conclusions 137	
8	DNA Microarrays: Beyond mRNA	143
	Armin Lahm	
	Introduction 143	
	mRNA variants (splicing) 144	
	Genomic tiling; transcription outside annotated regions of the genome 146	
	Profiling of non-coding RNAs 148	
	DNA methylation 149	
	“ChIP-on-chip” for mapping protein–DNA interactions and epigenetic marks 152	
	Detection of genome variations 155	
	Integrative genomics 156	
	Microarrays focusing on viruses and bacteria 156	
	Conclusions 157	
	Index	169