
Table of Contents

Justification, Objectives and Contributions	xv
---	----

Part I: Introduction

1. Introduction.....	1
1.1 The Omics Era and Bioinformatics.....	3
1.2 Gene expression microarray data	5
1.3 Time Course Microarray Data	8
2. Tools for Microarray Data.....	11
2.1. Introduction.....	13
2.2 Experimental design	14
2.3 Data pre-processing	16
2.3.1 The logarithm transformation	17
2.3.2 Normalization	17
2.4 Exploratory analysis.....	22
2.4.1 Hierarchical cluster.....	24
2.4.2 Non hierarchical cluster	26
2.4.3 Self Organizing Maps (SOM)	26
2.4.4 Self Organizing Tree Algorithm (SOTA)	28
2.4.5 Principal Components Analysis (PCA).....	28
2.5 Inferential analysis	29
2.5.1 Classical statistical tests.....	30
2.5.2 Relevant aspects in data analysis of microarray experiments.....	32
2.5.3. Specific statistical methods for microarray data analysis	35
2.6 Other applications of interest	37
2.7 Method comparison applied to Time Course data	39
2.7.1 Toxicogenomics experiment	39
2.7.2 Data visualization.....	40
2.7.3 Inferential analysis.....	49
2.8 Discussion	52

3. Review of Statistical Analysis of Time Course Microarray Data	55
3.1 Introduction.....	57
3.2 Clustering.....	59
3.3 Gene regulatory networks	62
3.4 Identifying differentially expressed genes	64
3.5 Discussion	67

Part II: Developed methodologies

4. maSigPro: a Method to Identify Significantly Differential Expression Profiles in Time-Course Microarray Experiments.....	71
4.1 Introduction.....	73
4.2 Methods	74
4.2.1 Definition of the model	74
4.2.2 First regression model: gene selection	75
4.2.3 Second regression step: variable selection	76
4.2.4 Visualization.....	78
4. 3 Results.....	78
4.3.1 Case 1: Toxicogenomics dataset.	78
4.3.2 Case 2: Simulated data.	82
4. 4 Discussion	83
5. ASCA-genes: Discovering gene expression patterns in Time Course Microarray Experiments by ANOVA-SCA.....	87
5.1 Introduction.....	89
5.2 Methods	91
5.2.1. Model definition	91
5.2.2 Gene-wise analysis.....	95
5.2.3 Comparative analysis	97
5.3 Results.....	98
5.3.1 Case 1: Simulated data.	100
5.3.2 Case 2: Toxicogenomics dataset.	105
5.4 Discussion	111

6. Using ASCA as data pre-processing technique in TCM	115
6.1 Introduction.....	117
6.2 Material and methods	118
6.2.1 The proposed filter	118
6.2.2 Datasets	121
6.2.3 The evaluation approach.....	125
6.3 Results.....	128
6.3.1. Simulation study 1: independent time-course data.....	128
6.3.2. Simulation study 2: longitudinal data.	133
6.3.3 Toxicogenomics dataset.....	134
6.3.4 NSF potato stress dataset	136
6.4 Discussion	137
7. Functional assessment of time course microarray data	141
7.1 Introduction.....	143
7.2 Methods	145
7.2.1 maSigFun	145
7.2.2 PCA-maSigFun.....	145
7.2.3 ASCA-Functional	146
7.2.4 Datasets	147
7.3 Results.....	150
7.3.1 Simulation studies.....	150
7.3.2 Experimental datasets	153
7.4 Discussion	160
<hr/>	
Part III: Conclusions, Annexes and References	
<hr/>	
8. General conclusion	165
Annexes	173
Annex 1: maSigPro package scheme	175
Annex 2: Supplementary figures of Chapter 5	177
Annex 3: ASCA-genes on simulated datasets with high number of genes	181

Annex 4: Supplementary tables of maSigPro applied to the simulation study 1 of Chapter 6	185
Annex 5: Details of the evolution of specificity and sensitivity with maSigPro on simulated study 1 of Chapter 6.....	191
Annex 6: Enrichment functional analysis of experimental data used in Chapter 6.....	195
Annex 7: Details of simulated datasets results of Chapter 7	197
Annex 8: Gene Ontology (GO)	203
References	207