

Contents

1	About this Thesis	1
1.1	Motivation	1
1.2	Objectives	2
1.3	Organization	4
1.4	Context	5
2	Introduction	7
2.1	Introduction	7
2.2	Biological background	8
2.3	The omics technologies and bioinformatics	13
2.3.1	Transcriptomics	15
2.3.2	Epigenomics	18
2.3.3	Proteomics	19
2.3.4	Metabolomics	20
2.4	From <i>omics</i> to Systems Biology: towards a more complete picture of life at the molecular level	21
2.4.1	Data integration in Systems Biology	23

3 Managing multi-omics experiments: the STATegra Experiment Management System	31
3.1 Introduction	31
3.2 Methods	34
3.2.1 STATegra EMS architecture	34
3.2.2 User system	41
3.2.3 Data specification	43
3.3 Results	49
3.3.1 Availability and requirements	49
3.3.2 Discussion	50
4 Experiment annotation using the STATegra EMS: an example of its use in Systems Biology	53
4.1 Introduction	53
4.1.1 The STATegra project	53
4.1.2 Studied system and experimental design	54
4.2 Project annotation using the STATegra EMS	56
4.2.1 Installing the STATegra EMS	57
4.2.2 User registration	57
4.2.3 Study annotation	57
4.2.4 Sample annotation	58
4.2.5 Analysis annotation	63
4.3 Results	73
4.4 Discussion	76
5 Integrative visualization of multi-omics data: The PaintOmics 3 Platform	77
5.1 Introduction	77
5.2 Methods	80
5.2.1 The PaintOmics 3 architecture	80

5.2.2	The KEGG pathways database	86
5.2.3	User system in PaintOmics 3	90
5.2.4	Accepted input data	92
5.2.5	Pathway enrichment analysis	104
5.2.6	Exploring data with PaintOmics 3	106
5.2.7	Administrator tools for PaintOmics	115
5.3	Results	118
5.3.1	Availability and requirements	118
5.3.2	Discussion	119
6	Visualizing multilevel integration models using PaintOmics 3: a use case in Systems Biology	123
6.1	Introduction	123
6.2	Running PaintOmics 3.	124
6.2.1	Data preparation for PaintOmics 3	124
6.2.2	Data submission	125
6.3	Results	128
6.3.1	Identifiers and names conversion	128
6.3.2	Obtaining pathways	130
6.4	Discussion	131
6.4.1	Significant pathways.	131
6.4.2	Pathway network	133
6.4.3	Pathway analysis	140
6.5	Conclusions	146
7	General discussion and Conclusions	147
7.1	General discussion	147
7.2	Conclusions	151
7.3	Reach and relevance	153
7.4	Major scientific contributions	154

7.5 Major future perspectives	156
Appendices	159
A Selection of pathways for the PaintOmics 3 use case	161
A.1 Introduction.	161
A.2 Selection of pathways	162
B PySiQ, a Python Simple Queue system	169
B.1 Introduction.	169
B.2 Features	169
B.3 Design	170
B.4 Example of use	171
B.5 Advanced use.	173
B.6 Availability	175
Acronyms	177
Bibliography	181