

Contents

Preface	vii
Acknowledgements	xi
1 Introduction	1
Introduction.....	1
Bioinformatics.....	2
What Is Bioinformatics?	2
References.....	2
Trends in Bioinformatics	3
What Is the State of the Art?.....	3
The Role of Bioinformatics in Studying Health and Disease	3
Bioinformatics Within Life Sciences Research Projects	4
Types of Analysis.....	4
Ontologies	5
Genomics and Proteomics.....	5
StructuralG enomics.....	6
Systems Biology	7
What Is Systems Biology?	7
References.....	7
Trends in Systems Biology	8
What Is the State of the Art?.....	8
Why Do Systems Biology?.....	9
Approaches to Modelling.....	9
Systems Biology Within Life Sciences Research Projects	10
Research Areas for Analysis.....	13
Dimensions of Life Sciences Research.....	13
Choice of Area	13
“Right-Size” Science	14
Collaborative Research	14
The Nature of Collaborative Research.....	14
Advantages of Collaborative Research	15
Access to Resources.....	15

Research Paradigms	16
European Commission Research Projects	17
Framework Programmes	17
Projects.....	17
A Practical Guide to Project Websites, Participants, Coordinators, Publications.....	18
Websites, Participants and Coordinators.....	18
Project Publications	18
Resources and Infrastructures	20
Accessing Knowledge and Projects	20
Participating in Projects	21
Key Sources	21

Part I Fundamental Collaborative Research and Computational Biology

2 Bioinformatics	25
Introduction.....	25
Genome Sequences	25
Annotation.....	25
European Contributions	26
Key Areas.....	26
Genome Annotation	26
A European Virtual Institute for Annotation	26
Distributed Annotation.....	27
An Integrated Approach.....	27
Annotation Deliverables	27
Genome Browser and Distributed Annotation Viewer	28
Experimental-Computational Collaboration	31
Thematic Collaborations.....	31
Critical Mass of Resources	31
Bioinformatics Tools For Annotation	32
Integrated Tool Development.....	32
Integrated Layer for Genomic and Proteomic Data.....	32
Protein Structure	33
Expression Data	34
Protein-Protein Interactions	36
Protein Sequence and Function Database.....	36
Protein Sequence Grouping	36
Gene Definition/Alternative Transcripts and Splicing.....	37
Gene Definition.....	37
Gene Definition and Alternative Splicing Methods.....	37
Alternative Transcription Goals.....	38
Alternative Transcription Methods	38
Future Research	39

Gene Regulation and Expression	39
Gene Regulation and Expression Processes.....	39
DNA Microarray Data	40
Expression Research Goals.....	40
Gene Regulation Research Goals.....	41
Systems Biology of Transcription and Regulation	42
Functional Annotation of Proteins	42
Protein Sequence, Structure and Function Integration.....	42
Sequence to Structure to Function Results	43
Functional Sites Results.....	43
Small-Ligand Binding.....	44
Future Plans	44
Post-translation Modification, Membrane and Localisation	
Prediction	44
Membrane Proteins and Results.....	44
Post-translation Modification and Localisation and Results.....	45
Future Post-translation Modification and Localisation.....	47
Protein Complexes, Networks and Pathways.....	47
Protein-Protein Complexes.....	47
Network Prediction	47
Metabolic Pathway Net.....	48
Future Pathway Work.....	48
Encyclopaedia of DNA Elements	49
Functional Elements in the Human Genome	49
International Collaboration	49
Functional Identification Methods	49
Genome Analysis Future.....	50
Major Result: Most DNA Is Transcribed to RNA	51
3 Systems Biology	53
Introduction.....	53
Systems Biology Projects	53
Networks and Dynamics	54
Cell Cycle.....	54
Cell Cycle Regulation	54
Cell Cycle Ontology and Knowledge Warehouse.....	55
Cell Cycle Model Organisms.....	55
Cell Cycle Research Objectives.....	55
Systems Biology Data Generation	56
Cell Cycle and Health	56
Standard Cell Synchronisation.....	56
Periodically Regulated Genes	57
Functional Modules in the Cell Cycle	57
Data Visualisation and Analysis	58

P53	58
The p53–Mdm2 Regulatory Network	58
Collaborative Approaches to Data Handling	58
Negative Feedback: p53 and Mdm2 Experiments	59
Negative Feedback: p53 and Mdm2 Modelling	59
Publications.....	60
Spindle Formation and Imaging.....	60
Light Microscopy	60
Microtubule Formation	61
Regulatory Feedback in Microtubule Formation	61
Signalling and Control	62
Central Role of Cell Signalling.....	62
Cell Growth, Differentiation and Survival Pathways.....	62
The Ras/Raf/Mek/ERK Pathway	63
Data and Modelling Interaction	65
Quantifying Signal Transduction.....	65
Mitogen-Activated Protein Kinase Pathways	65
AMP-Activated Protein Kinase Signalling Pathway	66
Health Applications of AMPK Modelling.....	66
Metabolic Regulation.....	67
<i>Bacillus subtilis</i> as a Model Organism.....	67
Regulation of Transcription in Bacteria.....	67
Technologies and Modelling.....	68
Systems Biology Approach to Transcriptional Modelling	68
RNA Metabolism	69
Applications of RNA Metabolic Analysis	70
Circadian Clock	70
Nature of the Circadian Clock	70
Entrainment.....	70
Multiple Pathway Integration.....	71
Resources for Systems Biology	71
Determining Protein Function from Sequence	71
Functional Sites via Structural Recognition	71
Exploitation of Features.....	72
Regulation, Transcription and Signalling	73
Use of Gene Expression Data	73
Systems Modelling.....	73
International Collaborations on Systems Biology Tool Development	74
Cellular Systems Biology	74
Intergovernmental Collaboration on Bacterial Systems Biology.....	74
National Programmes on Cellular Systems Biology	75
Local and Worldwide Scientific Collaboration.....	75
National Programmes.....	75
International Survey of Systems Biology	76

European Intergovernmental Programmes.....	76
USA Glue Grants	76
US Integrative Cancer Biology Programme	77
Informal but Structured International Collaboration	77
A Major FP7 Initiative in Systems Biology.....	78
Projects in Systems Biology Research.....	78
4 Developmental Biology and Ageing.....	85
Introduction.....	85
The Importance of Developmental Biology and Ageing	85
Plant Development.....	86
Leaf Biology	86
Plant Research.....	86
<i>Arabidopsis thaliana</i> as a Model Organism.....	86
Growth Factors.....	87
Integrated Approach to Leaf Development.....	87
Developmental Biology Technologies	88
Results.....	88
Impact	89
Stem Cells and Development.....	90
Stem Cell Databases and Bioinformatics Tools.....	90
Stem Cell Genomic Data	90
Mitochondria and Ageing	91
Evolutionarily Conserved Mechanisms	91
Databases for Model Organisms and Developmental Biology.....	92
Mitochondria Signalling	92
Metabolic Processes.....	92
Nuclear Receptors and Ageing	93
Nuclear Receptors.....	93
Nuclear Receptor Networks	93
Databases and Bioinformatics Tools.....	94
Implementation in the Seventh Framework Programme.....	94
Stem Cells	94
 Part II Resources Supporting Bioinformatics and Systems Biology Research	
5 Databases, Computational Tools and Services	99
Introduction.....	99
Computational Resources	100
Major Gateway to Collaborative Research	100
A Central Resource Gateway.....	100
General Search Tool.....	102

Tools and Services	102
Interlinking of Tools and Databases.....	103
Distributed Development of Resources	103
Centres and Partners	104
Collaborative Bioinformatics Resource and Infrastructure Projects.....	104
Database Infrastructure	104
Fungal Database.....	104
Distributed Annotation System.....	105
Distributed Annotation.....	105
Annotation Server Information Service	105
Database Grid Integration	106
A Bioinformatics Grid for Europe	106
Grid Connectivity and Worldwide Access.....	106
Objectives of a Grid Infrastructure	107
Expected Grid Results.....	108
Linked Databases	108
Linked Tools	109
Open Software Linked Tools	109
Automated Gene Finding.....	110
Manual Versus Automated Gene Finding.....	111
Sequence Motif Analysis	111
Integrating Promoter Motif Analysis and Gene Expression	112
Protein Family Analyses	112
Ontologies	113
Gene Ontology	113
Gene Ontology Annotations	113
Open Biomedical Ontologies.....	113
Medical Ontologies.....	114
Text Mining.....	114
Reference Searching Plus Full Data Overview	114
Text Mining on Texts	114
Systems Biology Toolboxes.....	115
Systems Biology Toolboxes for Experimentalists	115
Systems Biology Toolbox Applications Procedures.....	116
Yeast Systems Biology.....	117
Expression Tools.....	118
Power-Law Analysis Tools	118
Multigenic Disease Modelling Platforms	119
Modelling Thousands of Reactions	119
Platform for Editing, Analysing and Varying Biochemical Models	119
Object Classes for Biological Entities from Gene to Cell	121
Stochastic Cell Simulation.....	121

6 Supporting Infrastructures	123
Introduction.....	123
Wet-Laboratory Infrastructure and Data	123
Education, Human Resources and Publications.....	124
Wet-Laboratory Research Infrastructures	124
European Collaborative Projects.....	124
A European Systems Biology Institute	124
Workshop on Research Infrastructures	125
Model Organism Biobanks	125
Protein Structure Facilities.....	125
Genomics	127
Proteomics.....	127
Imaging Living Systems	128
Three-Dimensional Electron Microscopy.....	129
Research Infrastructures: Future Priorities	129
European Strategy Forum on Research Infrastructures	129
Future Infrastructures for Bioinformatics	130
Bioinformatics Infrastructures for Systems Biology	130
Medical Information Systems	132
Medical Informatics	132
Knowledge Management Systems	132
Choices of Systems Complexity	132
Education	134
School of Bioinformatics	134
Training Workshops and Courses	135
Outreach.....	135

Part III Disease-Related Collaborative Research and Computational Biology

7 Infectious and Major Diseases	139
Introduction.....	139
Approach to Disease Research.....	139
Computational Biology and Disease.....	139
Viral and Bacterial Pathogens	140
Cardiovascular–Pulmonary Disease	140
Diabetes.....	141
Neurological Diseases.....	141
Immunology.....	141
Drug Development References	142
Viral Pathogens	142
HIV/AIDS and Hepatitis C Virus/Hepatitis C Coordinated Studies and Databases.....	142

Bioinformatics Prediction of HIV Antiretroviral Resistance	
Trajectory	142
HIV Genotypes	143
HCV Sequence Alignment.....	144
Herpes	144
Bacterial Pathogens.....	144
Pathogens <i>Bacillus anthracis</i> and <i>Staphylococcus aureus</i> ,	
Using <i>Bacillus subtilis</i>	144
<i>Bacillus</i> Cell Factory.....	145
<i>Streptomyces lividans</i>	146
Protein Secretion in <i>Streptomyces</i>	146
Cardiovascular–Pulmonary Diseases and Diabetes	147
Congestive Heart Failure, Chronic Obstructive Pulmonary	
Disease and Type-2 Diabetes	147
Diabetes Screening.....	148
Proteomics and Modelling Approaches	149
Neurological Diseases.....	150
Systems Biology of the Neuron	150
Macular Degeneration.....	150
Modelling Molecular Interaction Networks.....	151
Proteomics Support	152
Immunology	153
Immunology Grid Models.....	153
Drug Development.....	153
Biosimulation in Drug Development	153
Mass-Spectrometric Resolution of Protein Phosphorylation	
in Hormonal Signalling.....	154
Metabolic Fates of Pharmaceuticals in Living Cells	154
Microcompartments Associated with Microtubular Networks	155
Regulation of Pancreatic α - and β -Cells	155
Neuronal and Systemic Models of Mental Diseases	
and Sleep Regulation	156
Synchronisation of Nephron Pressure and Flow Regulation	157
Models of Full-Scale Cardiac Arrhythmias	158
Spatio-temporal Organisation of Intracellular	
and Intercellular Ca^{2+} Dynamics	159
Modelling of Molecular Regulatory Mechanisms	
of Circadian Rhythms	159
Deep Brain Stimulation and Medication.....	160
Biological Networks, Data Analysis and Pharmacokinetic Models	161
Modelling Human Metabolism, Body Weight Regulation	
and the Treatment of Diabetes	161
Live Cell Imaging by Use of Interference Microscopy	161
Application of Methods from Non-linear Dynamics	
to Describe Complex Cellular Phenomena	162

Implementation in the Seventh Framework Programme.....	162
Innovative Medicines Initiative.....	162
Seventh Framework Programme Research Projects in the Systems Biology of Disease.....	163
8 Cancer	165
Introduction.....	165
Cancer Research Programmes	165
The Nature of Cancer.....	166
Challenges of Cancer Research	167
Relevance of Collaborative Research Projects.....	168
Nature of Cancer and Biology and Genetics of Cells and Organisms.....	168
Systems Biology of Cancer.....	168
Experimental and Clinical Data and Theoretical Models	169
Tumour Viruses.....	169
Role of Chronic Infections.....	169
Bioinformatics and Technology Platform.....	170
Cellular Oncogenes.....	170
Oncogenes Mutation Databases.....	170
Alternative Transcripts as Cancer Markers.....	172
Growth Factors and Their Receptors	173
Cell Growth Modelling	173
Cytoplasmic Signalling Circuitry	173
Ras/Raf/MEK/ERK and JAK/STAT Signalling.....	173
MAPK Signalling.....	173
Wnt and ERK Pathways.....	174
Regulatory Single-Nucleotide Polymorphisms.....	174
Cell Cycle.....	174
Cell Cycle Functional Modules	174
Tumour Suppressor Genes	175
pRb Tumour Suppressor	175
P53 and Apoptosis	176
Role of p53.....	176
p53 and Mdm2 Feedback Loops.....	176
p53 Mutations	176
p53 Database.....	177
Apoptosis Modelling.....	177
p53, p63 and p73 Comparisons.....	178
Cell Immortalisation, Tumourigenesis and Senescence.....	179
Irreversible Growth, Apoptosis and Premature Senescence	179
Predictive Dynamic Model	180
Multistep Tumourigenesis.....	180
Virtual Tumour Progression Features	180
Succession of Genetic Mutations in Colon Cancer.....	180

Genomic Integrity and the Development of Cancer	181
DNA Repair	181
Angiogenesis and Lymphangiogenesis	181
Angiogenesis.....	181
Lymphangiogenesis	182
Tumour Microenvironment Interactions	182
Metastasis.....	183
Metastasis of Breast Cancer.....	183
Tumour Immunology and Immunotherapy.....	184
Cancer Immunotherapy.....	184
Implementation in the Seventh Framework Programme.....	184
Seventh Framework Programme Research Project in the Systems Biology of Cancer.....	184
Implications of the New Project.....	185
9 Genetic Variation and Diseases.....	187
Introduction.....	187
Background.....	187
Call to Action.....	188
Genetic Variation Workshop	188
Genetic Variation Workshop Organisation and Goals	188
Editorial Encouragement	188
Status of Genetic Variation Research.....	189
Value of Cross-Linking Data	189
Linked Databases	189
Genetic Variation Data Sources	190
New Elements Facilitating Data Integration.....	190
Genotype to Phenotype.....	191
Biological and Medical Research: Genotype to Phenotype.....	191
Data and Analysis Challenges	192
Standards and Ontologies	192
Data Submission	193
Display and Analysis Tools.....	193
Analysis Challenges.....	194
Linking to Systems Biology Analysis.....	194
Databases	195
Databases to Be Linked	195
Approaches to Linking Databases in the Public Domain	197
Data Access.....	198
European-Level Support and International Collaboration	198
National-Level Programmes	198
Role of Model Organism Databases	198
Data Generation	199

Related Genetics Research and Infrastructures:	
Biobanks and Testing	199
Control Populations	200
Copy Number Variation	200
Patient as Data Source	200
Data-Taking Procedures	200
Genetic Etiology	201
The Way Forward	201
Obstacles	201
Linking as a Solution	202
Hub Software	202
Customised Entry Points	202
Principal Conclusions on Linking Genetic Variation	
Databases	202
Research in the Fifth Framework Programme	
and the Sixth Framework Programme	203
Population Genetics	203
Down's Syndrome and Relevant Genetic Regions	204
Experimental Tools for Studying Genetic Variation	205
Genetic Variation Mapping Data	206
QTL Data	206
Software Tools	207
Next Steps	207
A Major Seventh Framework Programme Initiative	
in Genetic Variation	207
Projects in Genetic Variation Research	207
Implications of New Projects	211

Part IV Science Management, Perspectives and Conclusions

10 Science Management	215
Introduction	215
Motivation	215
Framework Programmes for Research	216
Bioinformatics and Systems Biology Research	216
Methods for Executing the Specific Programme	217
Funding Methods	218
Topics in Calls for Proposals	218
Proposal Evaluation	218
Nature of Contracts and Grant Agreements	219
Project Management	219
How To Participate and Develop Proposals	220
How To Participate in Research Projects	220
European and International Participation and Collaboration	220

Calls for Proposals	221
Mechanics of Proposal Preparation	222
Support Services for the Mechanics of Proposal Preparation	222
Finding Partners for Proposals.....	223
Documents Relevant to the Proposal Process	223
Characteristics of Proposals.....	224
Evaluation of Proposals	228
Evaluation Process	228
Evaluation Criteria	229
Announcement of Evaluation Results.....	230
Project Negotiation	231
FP7 Negotiating Guidelines.....	231
Financial and Legal Negotiations	231
Negotiation Relevant Documents	231
Negotiation Key Points	233
Operating Project Attributes and its Management.....	235
The Model Grant Agreement.....	235
Project Coordination and Management Structures	235
Role of the Commission Scientific Officer.....	237
Role of Commission Financial and Contract Officers	238
Contractual and Financial Obligations	238
Role of Contracts in Collaboration.....	238
Model Grant Agreements	238
Project Reporting	239
Information Dissemination	239
Publications and Websites	239
A Systems Biology Website	240
A Bioinformatics Website.....	240
IPR and Exploitation.....	240
IPR Rules	240
Approaches to IPR Policy.....	241
OpenSource.....	242
IPR and Commercial Exploitation Support Services.....	243
11 Perspectives.....	245
Introduction.....	245
Role of Perspectives.....	245
Bioinformatics	246
Bioinformatics Perspectives	246
Model Systems and Biobank Resources.....	246
Standards and Ontologies	246
Obtaining Data Within and Beyond Present “Omics”, Extending Databases.....	247
Systems Biology	248

Systems Biology Perspectives	248
Developing Systems Biology: Cellular and Subcellular Systems	248
Multiple Interacting Systems at the Cellular and Physiological Levels	249
Physiology	250
Biotechnology	250
Synthetic Biology	251
Disease	251
Applying Systems Biology to Disease, Medicines and Treatment	251
Systems Biology of Targeted Diseases	252
Therapeutic Applications of Computational Biology and Chemistry	252
Seventh Framework Programme	253
Health Research	253
Physiological Integration	253
Longer Term	254
Systems Biology Forward Look	254
Very Long Term	254
12 Outstanding Results and Conclusions	255
Introduction	255
Tests of the Value of Collaborative Research	255
The Impact of Collaborative Research	256
Outstanding Resources Created	256
Integrated Data Access Capabilities	256
Establishment of Standards and Repository for Gene Expression Data	256
A European Bioinformatics Grid and Linked Resources	256
European Virtual Institute of Genome Annotation	257
European School of Bioinformatics and Training and Outreach	257
Outstanding Scientific Results	258
The Majority of Human DNA Is Transcribed into RNA	258
Understanding the Dynamic Behaviour of the p53–Mdm2 Network	258
Understanding the Dynamics of Spindle Formation in Cells	259
Establishment of the Role of Alternative Transcription and Splicing in Cancer	259
Small-Ligand Binding	260
Multiple Drug Treatment for HIV/AIDS Drug-Resistant Mutation Pathways	260
Cancer as A Signalling Disease, Applied to Protein Kinase Pathways	261
Future Project Outcomes	261

OverallC onclusions.....	261
TheC hallenge.....	261
Successes of Collaborative Research in Bioinformatics and Systems Biology	262
References	263
Index	281