

# Contents

|   |             |
|---|-------------|
| <b>Declaration</b>  | <b>i</b>    |
| <b>Abstract</b>   | <b>iii</b>  |
| <b>Resum</b>  | <b>v</b>    |
| <b>Resumen</b>  | <b>vii</b>  |
| <b>List of Figures</b>  | <b>xiii</b> |
| <b>List of Tables</b>   | <b>xv</b>   |
| <b>1 Introduction</b>   | <b>1</b>    |
| 1.1 Systems Biology: aim and definition . . . . .                                       | 3           |
| 1.2 Human protein interactome . . . . .   | 5           |
| 1.2.1 Definition . . . . .  | 5           |
| 1.2.2 PPIs identification . . . . .   | 5           |
| 1.2.3 Functional modules . . . . .  | 6           |
| 1.2.4 The interactome as a network . . . . .  | 8           |
| 1.2.5 Functional and evolutionary implications of the interactome<br>topology . . . . . | 8           |
| 1.2.6 Structural details . . . . .  | 10          |
| 1.3 Human genomics . . . . .  | 13          |
| 1.3.1 Sequencing the human genome . . . . .   | 13          |
| 1.3.2 Genotype to phenotype: the challenge of heterogeneity . . . . .                   | 14          |
| 1.3.3 Deleterious variability in healthy populations . . . . .                          | 15          |

---

|          |   |           |
|----------|---|-----------|
| 1.3.4    | Genetic heterogeneity among cancer patients . . . . .   | 16        |
| 1.4      | Thesis outline . . . . .  | 21        |
| <b>2</b> | <b>A methodology for functional profiling using protein-protein interactions data</b>             | <b>25</b> |
| 2.1      | Overview and objectives . . . . .   | 26        |
| 2.2      | Materials and Methods . . . . .   | 28        |
| 2.2.1    | Construction of protein interactomes . . . . .  | 28        |
| 2.2.2    | Modelling the interactome as a network . . . . .  | 29        |
| 2.2.3    | Identification of candidate functional subnetworks . . . . .                                      | 30        |
| 2.2.4    | Algorithm for subnetwork enrichment analysis in a ranked list                                     | 32        |
| 2.2.5    | GWAS analysis in bipolar disorder . . . . .   | 34        |
| 2.3      | Results and Discussion . . . . .  | 35        |
| 2.3.1    | Collection and curation of protein interactomes . . . . .   | 35        |
| 2.3.2    | Study of network parameters characteristic of real functional subnetworks . . . . .               | 38        |
| 2.3.3    | NetworkMiner: a tool for subnetwork enrichment analysis in a ranked list . . . . .                | 43        |
| 2.3.4    | Applications . . . . .  | 44        |
| <b>3</b> | <b>Role of the interactome in the maintenance of deleterious variability in human populations</b> | <b>47</b> |
| 3.1      | Overview and objectives . . . . .   | 48        |
| 3.2      | Materials and Methods . . . . .   | 50        |
| 3.2.1    | Human healthy individuals exome sequencing data . . . . .   | 50        |
| 3.2.2    | Cancer donors . . . . .   | 51        |
| 3.2.3    | Analysis of exome sequencing data . . . . .   | 51        |
| 3.2.4    | Selection of variants with functional impact . . . . .  | 52        |
| 3.2.5    | Construction of the human PPI network . . . . .   | 54        |
| 3.2.6    | Deciphering the effect of the deleterious variants on the interactome . . . . .                   | 55        |
| 3.3      | Results and Discussion . . . . .  | 58        |
| 3.3.1    | Variability in human healthy populations . . . . .  | 58        |
| 3.3.2    | Topological role of proteins carrying deleterious variants . . . . .                              | 65        |

|          |   |           |
|----------|---|-----------|
| 3.3.3    | Robustness of the interactome structure to homozygous deleterious variants of healthy individuals . . . . . | 68        |
| 3.3.4    | Robustness validation in tissue-specific interactomes . . . . .   | 73        |
| 3.3.5    | Distribution of deleterious variants among the interactome modules . . . . .                                | 76        |
| 3.3.6    | Comparison of germinal to somatic cancer-specific mutations . . . . .                                       | 79        |
| <b>4</b> | <b>Mutational oncogenic signatures on structurally resolved protein interacting interfaces</b>              | <b>83</b> |
| 4.1      | Overview and objectives . . . . .   | 84        |
| 4.2      | Materials and Methods . . . . .   | 86        |
| 4.2.1    | TCGA and ICGC cancer datasets . . . . .   | 86        |
| 4.2.2    | Analysis of exome sequencing data . . . . .   | 87        |
| 4.2.3    | Construction of the structurally resolved human protein interactome . . . . .                               | 88        |
| 4.2.4    | Statistical analysis of the mutations distribution among protein region types . . . . .                     | 89        |
| 4.2.5    | Identification of significantly mutated protein interacting interfaces . . . . .                            | 89        |
| 4.2.6    | Survival analysis . . . . .   | 90        |
| 4.2.7    | Identification of the Minimal Connected 3D Network mutated in cancer . . . . .                              | 90        |
| 4.3      | Results and discussion . . . . .  | 92        |
| 4.3.1    | Cancer donors . . . . .   | 92        |
| 4.3.2    | Construction of a three dimensional structurally resolved protein interactome . . . . .                     | 93        |
| 4.3.3    | Relevance of the protein interacting interfaces in cancer . . . . .   | 93        |
| 4.3.4    | Identifying significantly mutated protein interacting interfaces . . . . .                                  | 97        |
| 4.3.5    | Topological properties of enriched interfaces . . . . .   | 103       |
| 4.3.6    | The 3D cancer interactome: new insights into the cancer hallmarks . . . . .                                 | 104       |
| 4.3.7    | Clinical relevance of mutations affecting interacting interfaces: survival implications . . . . .           | 110       |

---

|          |  |            |
|----------|--|------------|
| <b>5</b> | <b>Summarizing discussion</b>                                    | <b>113</b> |
| 5.1      | Assembling the human protein interactome . . . . .               | 114        |
| 5.2      | NetworkMiner . . . . .   | 115        |
| 5.3      | The interactome as a buffer of deleterious variability . . . . . | 117        |
| 5.4      | Oncogenic signatures on PPI interfaces . . . . .                 | 119        |
| 5.5      | Future directions . . . . .                                      | 121        |
| 5.5.1    | The interactome model . . . . .                                  | 121        |
| 5.5.2    | Modelling mutation consequences . . . . .                        | 122        |
| <b>6</b> | <b>General conclusions</b>                                       | <b>125</b> |
|          | <b>Publications</b>  | <b>130</b> |
|          | <b>Bibliography</b>  | <b>131</b> |