

# Contents

*Preface*

*page ix*

## **Part I Introduction**

1

### **1 Basics**

3

- 1.1 Overview 3
- 1.2 Undirected and directed graphs 3
- 1.3 Trees 7
- 1.4 Rooted DAGs 8
- 1.5 Traversals of trees and DAGs 9
- 1.6 Taxa, clusters, clades and splits 11

### **2 Sequence alignment**

13

- 2.1 Overview 13
- 2.2 Pairwise sequence alignment 13
- 2.3 Multiple sequence alignment 20

### **3 Phylogenetic trees**

23

- 3.1 Overview 23
- 3.2 Phylogenetic trees 24
- 3.3 The number of phylogenetic trees 27
- 3.4 Models of DNA evolution 29
- 3.5 The phylogenetic tree reconstruction problem 32
- 3.6 Sequence-based methods 33
- 3.7 Maximum parsimony 33
- 3.8 Branch-swapping methods 37
- 3.9 Maximum likelihood estimation 40
- 3.10 Bootstrap analysis 43

|                |                                                         |            |
|----------------|---------------------------------------------------------|------------|
| 3.11           | Bayesian methods                                        | 45         |
| 3.12           | Distance-based methods                                  | 50         |
| 3.13           | UPGMA                                                   | 52         |
| 3.14           | Neighbor-joining                                        | 54         |
| 3.15           | Balanced minimum evolution                              | 56         |
| 3.16           | Comparing trees                                         | 60         |
| 3.17           | Consensus trees                                         | 63         |
| 3.18           | The Newick format                                       | 66         |
| <b>4</b>       | <b>Introduction to phylogenetic networks</b>            | <b>68</b>  |
| 4.1            | Overview                                                | 69         |
| 4.2            | What is a phylogenetic network?                         | 69         |
| 4.3            | Unrooted phylogenetic networks                          | 71         |
| 4.4            | Rooted phylogenetic networks                            | 76         |
| 4.5            | The extended Newick format                              | 81         |
| 4.6            | Which types of networks are currently used in practice? | 83         |
| <b>Part II</b> | <b>Theory</b>                                           | <b>85</b>  |
| <b>5</b>       | <b>Splits and unrooted phylogenetic networks</b>        | <b>87</b>  |
| 5.1            | Overview                                                | 87         |
| 5.2            | Splits                                                  | 88         |
| 5.3            | Compatibility and incompatibility                       | 90         |
| 5.4            | Splits and clusters                                     | 91         |
| 5.5            | Split networks                                          | 93         |
| 5.6            | The canonical split network                             | 97         |
| 5.7            | Circular splits and planar split networks               | 102        |
| 5.8            | Weak compatibility                                      | 105        |
| 5.9            | The split decomposition                                 | 107        |
| 5.10           | Representing trees in a split network                   | 121        |
| 5.11           | Comparing split networks                                | 122        |
| 5.12           | T-theory                                                | 122        |
| <b>6</b>       | <b>Clusters and rooted phylogenetic networks</b>        | <b>127</b> |
| 6.1            | Overview                                                | 127        |
| 6.2            | Clusters, compatibility and incompatibility             | 128        |
| 6.3            | Hasse diagrams                                          | 132        |
| 6.4            | Cluster networks                                        | 133        |
| 6.5            | Rooted phylogenetic networks                            | 138        |
| 6.6            | The lowest stable ancestor                              | 140        |

|                 |                                             |            |
|-----------------|---------------------------------------------|------------|
| 6.7             | Representing trees in rooted networks       | 144        |
| 6.8             | Hardwired and softwired clusters            | 146        |
| 6.9             | Minimum rooted phylogenetic networks        | 149        |
| 6.10            | Decomposability                             | 150        |
| 6.11            | Topological constraints on rooted networks  | 156        |
| 6.12            | Cluster containment in rooted networks      | 168        |
| 6.13            | Tree containment                            | 171        |
| 6.14            | Comparing rooted networks                   | 171        |
| <b>Part III</b> | <b>Algorithms and applications</b>          | <b>185</b> |
| <b>7</b>        | <b>Phylogenetic networks from splits</b>    | <b>187</b> |
| 7.1             | The convex hull algorithm                   | 187        |
| 7.2             | The circular network algorithm              | 190        |
| <b>8</b>        | <b>Phylogenetic networks from clusters</b>  | <b>193</b> |
| 8.1             | Cluster networks                            | 193        |
| 8.2             | Divide-and-conquer using decomposition      | 194        |
| 8.3             | Galled trees                                | 198        |
| 8.4             | Galled networks                             | 201        |
| 8.5             | Level- $k$ networks                         | 210        |
| <b>9</b>        | <b>Phylogenetic networks from sequences</b> | <b>216</b> |
| 9.1             | Condensed alignments                        | 216        |
| 9.2             | Binary sequences and splits                 | 216        |
| 9.3             | Parsimony splits                            | 218        |
| 9.4             | Median networks                             | 219        |
| 9.5             | Quasi-median networks                       | 223        |
| 9.6             | Median-joining                              | 227        |
| 9.7             | Pruned quasi-median networks                | 232        |
| 9.8             | Recombination networks                      | 233        |
| 9.9             | Galled trees                                | 240        |
| <b>10</b>       | <b>Phylogenetic networks from distances</b> | <b>250</b> |
| 10.1            | Distances and splits                        | 250        |
| 10.2            | Minimum spanning networks                   | 251        |
| 10.3            | Split decomposition                         | 251        |
| 10.4            | Neighbor-net                                | 254        |
| 10.5            | T-Rex                                       | 261        |

|                                                                  |     |
|------------------------------------------------------------------|-----|
| <b>11 Phylogenetic networks from trees</b>                       | 265 |
| 11.1 Consensus split networks                                    | 265 |
| 11.2 Consensus super split networks for unrooted trees           | 268 |
| 11.3 Distortion-filtered super split networks for unrooted trees | 273 |
| 11.4 Consensus cluster networks for rooted trees                 | 274 |
| 11.5 Minimum hybridization networks                              | 275 |
| 11.6 Minimum hybridization networks and galled trees             | 285 |
| 11.7 Networks from multi-labeled trees                           | 287 |
| 11.8 DLT reconciliation of gene and species trees                | 289 |
| <b>12 Phylogenetic networks from triples or quartets</b>         | 300 |
| 12.1 Trees from rooted triples                                   | 300 |
| 12.2 Level- $k$ networks from rooted triples                     | 302 |
| 12.3 The quartet-net method                                      | 308 |
| <b>13 Drawing phylogenetic networks</b>                          | 312 |
| 13.1 Overview                                                    | 312 |
| 13.2 Cladograms for rooted phylogenetic trees                    | 312 |
| 13.3 Cladograms for rooted phylogenetic networks                 | 316 |
| 13.4 Phylograms for rooted phylogenetic trees                    | 323 |
| 13.5 Phylograms for rooted phylogenetic networks                 | 324 |
| 13.6 Drawing rooted phylogenetic networks with transfer edges    | 327 |
| 13.7 Radial diagrams for unrooted trees                          | 328 |
| 13.8 Radial diagrams for split networks                          | 329 |
| <b>14 Software</b>                                               | 332 |
| 14.1 SplitsTree                                                  | 332 |
| 14.2 Network                                                     | 333 |
| 14.3 TCS                                                         | 334 |
| 14.4 Dendroscope                                                 | 334 |
| 14.5 Other programs                                              | 335 |
| <i>Glossary</i>                                                  | 338 |
| <i>References</i>                                                | 343 |
| <i>Index</i>                                                     | 358 |