

Contents

| | |
|---|-------------|
| List of Figures | XI |
| List of Tables | XIII |
| 1 Introduction | 1 |
| 1.1 Biological Background | 3 |
| 1.1.1 Genotype and Phenotype | 3 |
| 1.1.2 Microbial Genotypes and Phenotypes | 6 |
| 1.1.3 Alternative Approaches for Genotype-Phenotype Relations | 7 |
| 1.2 Computational Background | 9 |
| 1.2.1 Basics of Machine Learning | 9 |
| 1.2.2 Association Rule Mining | 14 |
| 1.2.3 Support Vector Machines | 16 |
| 1.2.4 Databases for Orthologs | 22 |
| 2 Materials and Methods | 25 |
| 2.1 Genotype data | 25 |
| 2.2 Phenotype data | 27 |
| 2.3 Accuracy Measures | 27 |
| 2.4 KRONA | 29 |
| 2.5 NetCAR | 29 |
| 2.6 PICA | 32 |
| 2.7 GenTraitor | 38 |

| | |
|--|------------|
| 3 Results and Discussion | 41 |
| 3.1 Comparison and Evaluation of Phenotype Prediction Packages | 41 |
| 3.1.1 NetCAR | 41 |
| 3.1.2 PICA | 44 |
| 3.1.3 Complexity Estimation | 47 |
| 3.1.4 Prediction on Incomplete Genomes | 50 |
| 3.2 Extension to Novel Phenotypes | 54 |
| 3.2.1 Methane Oxidation | 54 |
| 3.2.2 Nitrification | 58 |
| 3.2.3 Intracellular lifestyle | 62 |
| 3.3 Phenotype Prediction in Metagenomics | 67 |
| 3.3.1 Coral metagenome | 69 |
| 3.3.2 Sponge metagenome | 70 |
| 3.3.3 Biogas-fermenter metagenome | 71 |
| 4 Conclusion and Outlook | 79 |
| Abbreviations | 83 |
| References | 85 |
| A Source code | 97 |
| B Tables | 105 |
| C Acknowledgments | 109 |



<http://www.springer.com/978-3-658-14318-3>

Machine Learning for Microbial Phenotype Prediction

Feldbauer, R.

2016, XIII, 110 p. 29 illus., Softcover

ISBN: 978-3-658-14318-3