

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Eukaryotic transcriptional regulation	4
1.2.1	Transcription factors	7
1.2.2	Cis-regulatory modules	13
1.2.3	Complexation is important in development	21
1.3	Outline and Goal	23
2	Related work	25
2.1	Protein complex prediction from networks	25
2.1.1	ClusterONE and cohesiveness	27
2.1.2	Quality measures for complex prediction	30
2.2	Protein complexes beyond plain networks	34
2.2.1	Domain-domain interaction model	37
2.3	Prediction of cooperative TFs	40
2.3.1	Expression coherence scoring	40
3	Materials and Methods	45
3.1	Domain-aware cohesiveness optimization	45
3.1.1	Why seeding from pairs is beneficial	51
3.2	Data sources, their retrieval and preprocessing	53
3.2.1	General yeast protein data	53
3.2.2	Weighted protein-protein interaction data	55
3.2.3	Domain-association and domain-interactions	56
3.2.4	Transcription factors and their binding sites	60
3.2.5	Expression data	64
3.2.6	Reference data	65
3.3	Workflow and implementation	66
3.3.1	Building the domain-domain interaction network	68
3.3.2	Domain-aware cohesiveness optimization	69

4	Results and Discussion	89
4.1	Impact of algorithm engineering on runtime	90
4.2	Common protein complex prediction benchmarks	91
4.2.1	Comparison to reference complexes	94
4.2.2	Assessment of biological relevance	97
4.2.3	Evaluation of postprocessing and thresholds	98
4.3	Analysis in the transcription factor context	100
4.3.1	Estimation of target genes	101
4.3.2	Estimating the modes of action	103
4.3.3	Significance in yeast cell cycle	106
5	Conclusion and Outlook	111
5.1	Conclusion	111
5.2	Outlook	111
A	Additional tables	113
A.1	Supplement to complex prediction benchmarks	113
A.2	Supplement to target analysis	119
	Bibliography	121

<http://www.springer.com/978-3-658-08268-0>

Predicting Transcription Factor Complexes
A Novel Approach to Data Integration in Systems
Biology

Will, T.

2015, XIX, 142 p. 29 illus., Softcover

ISBN: 978-3-658-08268-0