

Zusammenfassung

In dieser Arbeit verwenden wir für die Untersuchung genomweiter Transkriptionssysteme Methoden der integrierten Datenanalyse. Es werden einige solche Methoden entwickelt und auf genomweite heterogene Datensätze des Modellorganismus *Saccharomyces cerevisiae* angewandt. Unser Ziel ist es, ausgeprägte funktionell interpretierbare Signale und die damit verbundenen Mechanismen aus stark verrauschten genomweiten Transkriptionsdaten nachzuweisen.

Unsere Analyse beginnt mit Gengruppen, die durch gemeinsame Transkriptionsfaktoren reguliert werden, den sog. Transkriptionsmodulen. Diese werden mittels genomweiter Protein-DNA-Interaktionsdaten identifiziert und werden dann mit Genexpressions- und Annotationsdaten kombiniert um funktionelle Signale zu identifizieren. Durch Anwendung von Standardmethoden auf unterschiedliche *large-scale* Genexpressionsdatensätze wird gezeigt, dass die auf diese Weise identifizierten funktionellen Module sowohl konditionsinvariant als auch konditionsspezifisch sein können.

Im zweiten Schritt erweitern wir diese Modulanalyse, um eine Priorisierung der genetischen Regulationswechselwirkungen in diesen funktionellen Modulen zu erzielen. Mittels einer von uns entwickelten einfachen integrativen Methode für diese Priorisierung erzielen wir eine statistisch signifikante Erhöhung der Vorhersagegenauigkeit für zwei Arten von Referenzdatensätzen im Vergleich zu einer allein auf genomweiten Protein-DNA-Interaktionsdaten basierenden Analyse. Zusätzlich prognostiziert unsere Methode bei ähnlicher Vorhersagegenauigkeit regulatorische Interaktionen, die mit anderen Algorithmen nicht vorhergesagt wurden.

Zum Schluss wenden wir uns - im Anbetracht der Allgegenwärtigkeit kombinatorischer Regulation durch multiple Transkriptionsfaktoren - verschiedenen Gensätzen zu, welche unter verschiedenen Konditionen durch Paare von Transkriptionsfaktoren reguliert werden. Wir entwickeln eine Methode um konditionsspezifische Kofaktoren jener Regulatoren zu identifizieren, welche unter verschiedenen experimentellen Bedingungen den Satz ihrer Zielgene signifikant ändern. Wir wenden diese Methode auf genomweite Protein-DNA-Interaktionsdaten

aus verschiedenen experimentellen Bedingungen an. Unsere Vorhersage enthält sowohl neue als auch schon bekannte kooperative Regulatorpaare, die durch Genexpressions-, Protein-Protein-Interaktions-Daten und konservierten Sequenzmotiven unterstützt werden. In einer weiteren Analyse zeigen wir, dass solche konditionsspezifischen kombinatorischen Regulatoren öfter auftreten als man es rein zufällig erwarten würde.

Zusammenfassend können wir sagen, dass wir mittels unserer Analysen biologisch relevante Aussagen und konkrete Hypothesen aufgrund heterogener genomweiter Datensätze aus Hefeexperimenten formulieren können. Daher hoffen wir mit dieser Arbeit unsere experimentelle Kollegen zur Überprüfung dieser Hypothesen, sowie zur Untersuchung detaillierterer biologischer Mechanismen zu motivieren.