A U S H A N G

FREIE UNIVERSITÄT BERLIN

Fachbereich Mathematik und Informatik

Promotionsbüro, Arnimallee 14, 14195 Berlin

D I S P U T A T I O N

Freitag, 27. Februar 2015, 11.00 Uhr

Ort: Max-Planck-Institut für molekulare Genetik,
Ihnestraße 63-73, 14195 Berlin,
Seminarraum III, 0.3.06

Disputation über die Doktorarbeit von

Frau Alena van Bömmel

Thema der Dissertation:
Prediction of transcription factor co-occurrence using rank based statistics

Thema der Disputation:
Rank list aggregation methods applied to miRNA

Die Arbeit wurde unter der Betreuung von **Prof. Dr. M. Vingron** durchgeführt.

Abstract: Due to the rapid development of experimental techniques in molecular biology over the past few decades, large amounts of high-dimensional data are now available. One popular representation of the results from the analysis of such high-dimensional data is the ranked list representation. In particular, the ranked list representation has an advantage in the simplicity of combining and comparing results of several experiments. For this purpose, rank aggregation techniques are applied. In general, rank aggregation techniques estimate a central (consensus) ranked list that reflects the overall preference within the set of ranked lists. In the framework of an optimization problem, the consensus list is a list with the smallest rank based distance (e.g. Spearman's footrule or Kendall's tau) to all individual ranked lists. Although this idea is very simple, different rank aggregation techniques differ greatly both in their mathematical complexity and in the obtained results. Here, the existing methods for rank aggregation that are frequently applied to biological data are reviewed. These methods range from simple techniques involving rank averaging to more complex techniques that employ advanced statistical methods such as Cross-entropy Monte Carlo or Markov chain models. The different approaches are demonstrated by applying them to miRNA sequencing data for which ranked lists for two different sequencing platforms and three different library preparation methods are available. The performance of various rank aggregation methods is then evaluated using the rank based distance measures.

In the second part of the presentation, the summary of the thesis with the title "Prediction of transcription factor co-occurrence using rank based statistics" is presented.

Die Disputation besteht aus dem o. g. Vortrag, danach der Vorstellung der Dissertation einschließlich jeweils anschließenden Aussprachen.

Interessierte werden hiermit herzlich eingeladen

Der Vorsitzende der Promotionskommission
Prof. Dr. M. Vingron