

FREIE UNIVERSITÄT BERLIN
Fachbereich Mathematik und Informatik

Promotionsbüro, Arnimallee 14, 14195 Berlin

DISPUTATION

Donnerstag, 26. Februar 2015, 10.00 Uhr

**Ort: Max-Planck-Institut für molekulare Genetik,
Ihnestraße 63-73, 14195 Berlin,
Seminarraum SR 1**

Disputation über die Doktorarbeit von

Frau Rina Ahmed

**Thema der Dissertation:
Bioinformatic Analysis of microRNA Genes
in Free-Living and Parasitic Nematodes**

**Thema der Disputation:
Mapping Next-Generation Sequencing Data
with Burrows-Wheeler Transform and FM Index**

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

Abstract: Next-generation sequencing (NGS) machines produce millions of short sequences (reads) in a single run. A fundamental step in the analysis of NGS datasets is the alignment (mapping) of reads to a reference sequence. This read mapping problem can be essentially seen as an approximate string matching problem, i.e. find the locations on the reference string that are similar to the pattern string under some measure of similarity. However, mapping the enormous amount of NGS reads to a large reference sequence is a non-trivial computational task, and numerous read mapping approaches have been developed recently that target different speed-accuracy trade-offs.

In this talk I will provide an overview of different read mapping strategies with particular focus on prefix trie based mappers. These mappers utilize a fast and accurate read mapping approach by creating an efficient index of the reference sequence, which is based on the Burrows-Wheeler transform (BWT). BWT-based implementations typically use the full-text minute-space (FM) index data structure, which has been referred to as a compressed suffix array. Using backward search, BWT-based mappers essentially mimic a top-down traversal on the prefix trie of the reference sequence and can determine exact matches in time linear to the read length and independent of the size of the reference sequence. After elaborating on the concepts of prefix trie based mappers, I will conclude the talk with a performance comparison of prefix trie and hash table based mappers.

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