

A U S H A N G

FREIE UNIVERSITÄT BERLIN Fachbereich Mathematik und Informatik

Promotionsbüro, Arnimallee 14, 14195 Berlin

DISPUTATION

Dienstag, 22. November 2016, 10.30 Uhr

**Ort: Seminar Room I (ground floor, MPIMG)
Max Planck Institute for Molecular Genetics
Innestrasse 73, 14195 Berlin**

Disputation über die Doktorarbeit von

Frau Verena Heinrich

**Thema der Dissertation:
Aspects of Quality Control for Next Generation Sequencing Data
in Medical Genetics**

**Thema der Disputation:
Linkage Analysis and its Application
on Whole Genome Sequencing Data**

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

Abstract: Linkage analysis is a well-established approach for studying the connection between the pattern of the occurrence of a Mendelian or complex trait and the inheritance pattern of certain genes in a given pedigree. During the last years the focus shifted to genome-wide association studies (GWAS) which are more powerful in the analysis of common variants.

However, as high-throughput sequencing techniques now allow a comprehensive detection of rare variants throughout the exome or whole genome, linkage analysis is again emerging as a powerful tool to detect associations between rare variants and complex diseases. In my presentation I will give an overview about the basic concepts of association and linkage analyses, whereby the major emphasis is on two computational models to handle linked markers, namely the Elston-Steward algorithm and the Lander-Green algorithm. Both approaches are in general use for parametric linkage analysis across a set of markers but are based on different strategies. The major difference between the two algorithms is the complexity. The Lander-Green algorithm uses a hidden markov model (HMM) where the recursion takes place over single loci so that the computing effort increases linearly with the number of markers but exponentially with the size of the family. On the other hand the Elston-Steward algorithm considers genotypes simultaneously which causes the algorithm to increase exponentially in the number of loci but scales linearly in the number of pedigree members. However with respect to missing data the Lander-Green approach is preferable and I will provide a detailed explanation of this algorithm.

The second part of the presentation will summarize the last chapters of my thesis entitled "Aspects of Quality Control for Next Generation Sequencing Data in Medical Genetics".

The identification of disease causing mutations is often based on the analysis of several family members and sample mix-ups can lead to erroneous conclusions when filtering for potentially pathogenic variants.

I will present an approach to reconstruct entire family structures solely based on whole exome sequencing (WES) datasets including rare variants.

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