Heinrich Heine Universität Düsseldorf

Heinrich-Heine-Universität Düsseldorf 40204 Düsseldorf
Dekanat der Mathematisch-Naturwissenschaftlichen Fakultät

An alle
hauptamtlichen Professoren/innen
und Privatdozenten/innen
des Faches Informatik der
Mathematisch-Naturwissenschaftlichen Fakultät

Mathematisch-Naturwissenschaftliche Fakultät

Dekanat

Promotionsangelegenheiten
Universitätsstraße 1
40225 Düsseldorf
Telefon: +49 (0)211 81 15092
E-Mail: promotionmnf@hhu.de

01.10.2024

Promotionsverfahren von **Herrn M.Sc. Sebastian Alexander Fuchs**
Auslage der Dissertation und Gutachten sowie Termin der mündlichen Prüfung
Anlage: Einseitige Zusammenfassung der Dissertation

Sehr geehrte Damen und Herren,

in dem oben genannten Promotionsverfahren wird die Annahme der Dissertation

Short- and long-read based resolution of complete bacterial genomes with applications in outbreak analysis and tracking of resistance genes

von den Berichterstattenden Prof. Dr. A. Dilthey und Prof. Dr. G. Klau beantragt. Sie kann zusammen mit den Gutachten in der Zeit

vom 20.10.2024 bis 31.10.2024

eingesehen werden. Bitte wenden Sie sich zur Einsicht an das Promotionsbüro (promotionmnf@hhu.de).

Einsprüche gegen diese Dissertation können nur zwei Tage nach der vorgenannten Frist geltend gemacht werden. Erfolgt kein Einspruch, so gilt die Dissertation als angenommen (§ 7 Ziffer (5) PO).

Sofern die Dissertation angenommen wird, findet die mündliche Prüfung am

06.11.2024 um 10.00 Uhr

im **Seminarraum A in Geb. 22.21, OG 3** statt. Als Prüferinnen bzw. Prüfer sind vorgesehen: Prof. Dr. K. Pfeffer, Prof. Dr. B. Henrich und Prof. Dr. L. Rose.

Die Öffentlichkeit ist bei der Befragung zugelassen.

Mit freundlichen Grüßen
im Auftrag

Daniela Schleiffer

Abstract – Sebastian Alexander Fuchs

Research in bacterial pathogen genomics has witnessed significant advancements in sequencing technologies. In the realm of bacterial genomics, our work addresses both bioinformatic challenges, as well as cost and labour constraints, associated with the use of novel long-read sequencing technologies within three different projects.

First, introducing Ultraplexing, we present a method that substantially reduces per-sample sequencing costs and hands-on time in Nanopore sequencing for hybrid assembly. Ultraplexing eliminates the need for molecular barcoding bs bioinformatically determining which specific sequenced isolate a long-read belongs to; this is done by comparing each long-read to the k-mer spectrum of the sequenced isolates, measured using Illumina data. This method holds promise for large-scale bacterial genome projects that utilize hybrid assembly strategies, enabling considerable savings without compromising assembly quality. These advantages are enabled by the possibility to multiplex at least 100 isolates together, representing roughly fourfold increase of isolates possible at the time of publication, thus also reducing hands-on time in the lab by a factor of four.

Second, shifting focus to the hospital associated pathogen *Acinetobacter baumannii*, we investigate genome plasticity and horizontal gene transfer mechanisms in the context of transmission of colistin resistance elements. Through short- and long-read sequencing and creation of hybrid assemblies, we identify two probable recombination events in the *pmrCAB* operon, which confers colistin resistance. Our findings highlight the role of homologous recombination and shed light on the possible contribution of mobile genetic elements to this phenomenon in *A. baumannii*. This study contributes to the understanding of antibiotic resistance dynamics in clinical isolates of *A. baumannii*, specifically those belonging to International Clone 7.

Third, expanding the scope to genomic pathogen surveillance in healthcare facilities, we introduce NanoCore, a user-friendly method developed for Nanopore-based outbreak surveillance and investigation. NanoCore enables the determination and visualization of cgMLST-like sample distances directly from raw Nanopore reads by mapping Nanopore data to a core genome reference, variant-calling and calculating distances from the results, thus offering a fast and flexible solution. Validated on methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococcus faecium* (VRE) datasets, NanoCore demonstrates high accuracy, producing results quasi-identical to those of current gold-standard tools with an average difference of 0.75 alleles for MRSA and 0.81 alleles for VRE in Nanopore-only-mode and 3.44 and 1.95 alleles respectively in hybrid-mode (measured in closely related isolates). The computational efficiency, open-source availability, and user-friendly installation via bioconda make NanoCore a valuable tool for effective bacterial pathogen surveillance in healthcare settings.

In conclusion, the work presented in this thesis spans the development of methods for hybrid genome assembly, long-read-based genomic surveillance and the investigation of the transmission of antibiotic resistance elements. The presented work demonstrates the potential of combining data generated by different sequencing technologies for bacterial genomics, as well as the potential of bioinformatics methods development for emerging sequencing technologies.