03: A comprehensive analysis of microorganisms and viruses from ancient samples using minion sequencing techniques

Lead supervisor:

Prof. Dr. Manja Marz
RNA Bioinformatics and High-Throughput Analysis
Faculty of Mathematics and Computer Science
Friedrich Schiller University Jena
07743 Jena, Germany
Email: manja@uni-jena.de
http://www.rna.uni-jena.de/en/

Co-supervisors:

Prof. Dr. Johannes Krause
Max Planck Institute for the Science of Human History
Department of Archaeogenetics
Kahlaische Str. 10
07745 Jena, Germany
Email: krause@shh.mpg.de
http://www.shh.mpg.de/2890/johanneskrause

Dr. Denise Kühnert
Max Planck Institute for the Science of Human History
Independent Research Group Leader: tide
Kahlaische Str. 10
07745 Jena, Germany
Email: kuehnert@shh.mpg.de
http://www.shh.mpg.de/tide

Involved subjects: Bioinformatics, Archaeogenetics

Number of positions requested: 1

Abstract:

In current metagenomic samples we are able to assign only up to 10% of the microorganisms, most of them being bacteria. This number drops significantly when investigating additionally viruses in especially ancient samples. One reason for not being able to identify the entire bacteria community is the lack of knowledge on genomic sequences of unculturable microbes in databases to compare with. Viruses have been just recently addressed to be of great importance and therefore many methods for specific extraction methods of small particles and the downstream analysis, e.g. assembly of viral genomes after sequencing, have not been established yet.

The aim of the project is the development of novel tools and usage of new techniques to comprehensively analyze microorganisms including viruses. This will lead to the possibility to study the spread and diversification of human-associated microbes and infectious diseases.

The approach will consist of three novel parts:

(1) We recently established sequencing with Minion (ONT) in Jena. This technique will lead to sequences of up to 500.000 nt length (expected mean about 80.000 nt) of bacteria and viruses. The challenge is to optimize the existing protocols to a minimum of input material due to the limited DNA extracted from ancient bones (in collaboration with co-supervisor J. Krause).

(2) Gained sequences will be assigned to known bacteria with a novel bioinformatical tool (developed in this project) based on artificial intelligence using information of k-mer and DNA composition instead of homology to existing sequences. With such an generic approach we expect to be able to classify also unknown bacteria to at least genus level (as preliminary data of our group show).

(3) We finally aim to identify viral sequences by assembling minion reads (as previously done by our group for cell culture samples). The challenge here is that no assembly tool for viral reads from Minion exists so far. Viruses usually come as quasispecies which makes it difficult to discriminate mutations from sequencing errors.

The PhD student should have a master debgree in molecular life sciences, bioinformatics or computer scientist. The student should have experience to work in wet lab (cell culture, PCR design, qRT-PCR, DNA/RNA extraction, DNA recombination, transfection, genome editing, EMSA, etc.) but at the same time also experience on the bioinformatical side (Linux, programming skills in at least shell/bash and/or a script language, genome assembly, genome annotation, genome comparison, pangenome analysis, phylogenetic reconstruction, etc.).

Experience with Minion Sequencing would be preferable.