



3<sup>rd</sup> Science for the Environment Conference  
Aarhus Denmark 1-2 October 2015

## SEQUENCING NEW GENOMES TO IMPROVE IDENTIFICATION OF eDNA – A NGS WORKFLOW FOR BIODIVERSITY ASSESSMENT IN SOIL

Frank Panitz, Toke T. Høye, Lars-Erik Holm, Peter Wind, Paul Henning Krogh, Liselotte Wesley Andersen

<sup>1</sup>Dept. Molecular Biology and Genetics, Aarhus University; <sup>2</sup>Aarhus Institute of Advanced Studies, Aarhus University; <sup>3</sup>Dept. Bioscience, Aarhus University

### ABSTRACT

A next generation deep sequencing approach was applied to evaluate its use for assessing or monitoring biodiversity using environmental DNA (eDNA). In a pilot study, soil samples from two coastal heath habitats in north-western Denmark were compared: one native, the other growing a Sitka spruce plantation for about forty five years. From 250 mg soil eDNA was extracted and randomly sequenced on an Illumina HiSeq2000. This shotgun approach is supposed to generate a more diverse spectrum of sequences than the conventional PCR-based targeting of organism groups (e.g. plant (trn), insects (COI), Collembola (18S RNA), bacteria (16S RNA)) followed by sequencing, omitting primer targeting and PCR amplification biases of the latter. To improve identification of springtails in the habitats, specimen from 15 species were individually collected, pooled and sequenced. Main components of the NGS-workflow comprise: [1] partial assembly of Collembola genomes into contigs (length dependent on sequencing coverage and repeat content); [2] identification and removal of bacterial/viral genomes (as complete animals including head and interior organs were used); [3] creating a local database of 15 sequenced Collembola species; [4] soil sample sequences are annotated by identifying hits against a) bacterial/viral genomes, b) Collembola genome contigs, c) non-redundant protein database. The resulting species lists will be compared between habitats and to the composition of the soil samples as determined by the metabarcoding PCR-marker based approach. Data from our deep genomic sequencing approach represents a larger spectrum of species including assessment of local springtail populations and allows an improved biodiversity monitoring using eDNA-barcodes. Besides monitoring plants and animals, the deep sequence coverage generated provides the opportunity to investigate the complex metagenome in greater detail, which might yield insights into the function of the soil's microbiome



AARHUS  
UNIVERSITY

DCE - DANISH CENTRE FOR ENVIRONMENT AND ENERGY