



Nordic Intensive Master Course by NOVA SoilSoc network

- Course name: NIMC Bioinformatics for environmental scientists (3 ects)
- Course time: NOVA winter window 7-11.1. 2008
- Place: UH-AF Viikki Campus, lecture room and computer class
- Organiser: Kristina Lindström and research group
- Target students: M Sc students at all SoilSoc partner universities. In addition, PhD students
- Student numbers: 25 M Sc students; more can be accommodated at the lectures if required and during the whole course in case we find a bigger computer class

The need for this course

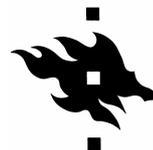
The SoilSoc network encompasses microbiologists, soil and environmental chemists, soil physicists, pedologists, hydrologists etc. A common theme for all these separate disciplines is experimental planning and the handling of large data sets. In biology, the handling of genomic data has recently become everyday practice in research and this discipline is often called bioinformatics.

We here use the term bioinformatics in a broader sense and include all kinds of data obtained from measurements in the environment. There is an urgent need for this kind of teaching in all SoilSoc member institutions, and our partners have enthusiastically welcomed our initiative to arrange this course.

In our microbiology curriculum at my department we introduced more mathematics and its use for data handling during the Bologna process. This course fills a big gap in the new curriculum at the master's level not only in Helsinki but also in the other Nordic countries.

The course program

The course program is planned to start from single microorganisms and methods for their classification "from scratch", in order for the students to understand the mathematics, logics and reasoning behind more sophisticated approaches (day 1). We then move to the study of single genes and computational methods which can be applied to these on a routine basis (day



2). Day 3 is devoted to different ways of analysing complex environmental samples by fingerprinting with the emphasis on different kinds of data and the different ways available to handle it. The themes of the last two days focus on whole genome sequence comparisons (day 4) and metagenomics – the analysis of complex gene samples from the environment (day 5).

The aim is to make the students acquainted with (i) the theoretical background behind the methods by using common sense; (ii) how to use them by hands on exercises; (iii) new perspectives in this rapidly moving field.

In addition to five full working days on site, the students will be given literature to read in advance and tasks which prepare themselves for the intensive learning process. After the course they will be requested to write an essay about what they learnt (the exact task will be specified later).

Teachers

We have managed to engage top teachers for all items to be taught.

Theme I: Identifying and differentiating pure cultures

Dr Hannu Rita is the AF faculty expert on experimental design, philosophy of data acquisition and statistics.

Dr Leena Räsänen from our group works as a research scientist and focuses at the moment on bioinformatics related to strain phylogenies and grouping. She is an expert on bacterial diversity and biogeography.

Docent Kristina Lindström is an expert on taxonomy and identification, is internationally known in the field and teaches the methods on a regular basis.

Theme II: Studying single genes

Head of laboratory Lars Paulin from the Institute of Biotechnology, UH, is in charge of the sequencing facility and an expert on all methods taught on the course.

Dr German Jurgens from our group is one of the world experts on the phylogeny of *Archaea* and teaches both domestically and abroad.

Theme III: Fingerprinting environmental samples



Docent Maarit Niemi from the Finnish Environment Institute (SYKE) developed a patented system for measuring soil enzyme activities and associated data handling.

M Sc (to be 2007) Anu Mikkonen is a graduate student from our group using the Bionumerics software for environmental applications and the best expert on this in our network.

Theme IVa: Data mining from large sequence pools – genomes

Theme IVb: Data mining from large sequence pools - metagenomes

Professor Peter Young from the University of York, UK, is the external invited teacher for the course. He is a long-standing collaborator of my group. He has a long experience from teaching at international courses (among them Helsinki 1993) and is well known for his genome sequencing efforts and comparative genomics. We are very lucky that he has accepted the invitation to participate.

The teachers responsible for the various themes will before the course discuss the items to be taught in detail and make up the final schemes for each lesson, which will involve collaboration of teachers as well.