Data science & computational analysis for metagenomics			
Dr. Edoardo Pasolli		(e-mail) edoardo.pasolli@unina.it (phone) 081 253 9397	
6 CFU	Classroom	Lectures	14 hours
		Laboratory	28 hours

Objectives: Metagenomics aims at studying genetic material recovered directly from environmental samples. Because of its ability to reveal previously hidden diversity of microbial life, it offers a powerful tool to advance our understanding of different microbiomes including those coming from food, human, and environmental sources. As the cost of sequencing continues to fall and thanks to advancements in processing methodologies, it allows microbiomes to be investigated at a much greater scale and detail than before. However, such opportunities come with the need to acquire expertise in running computational and bioinformatics pipelines and applying specific data analysis methodologies. The course will focus on recent advances in high-throughput sequencing technologies and will give an overview of the entire processing pipeline (from pre-processing to final data analysis) required to analyze shotgun metagenomics data.

Learning outcome: The course will provide both theoretical and practical basis for studying complex microbial communities through shotgun metagenomics. At the end of the course the student should be able to: - understand basic concepts and challenges in study microbial communities through shotgun metagenomics approaches; - run existing computational pipelines to process shotgun metagenomics datasets; - run basic scripts for data analysis; - integrate experimental and publicly available data; - evaluate results in a critical way.

Topics: The course will be composed by three main parts: - basis of microbial ecology and introduction to metagenomics (2 CFU); - execution of existing computational pipelines to analyze shotgun data. This will comprise raw-reads pre-processing, reference-based approaches for taxonomic and functional profiling, assembly-based approaches for de novo genome reconstruction and functional annotation, phylogenetic analysis, integration of publicly available data (2 CFU); - execution of scripts for data analysis and biological interpretation of the results (2 CFU).

Evaluation: Class attendance is mandatory (> 80%; exceptions are allowed but need to be discussed with the lecture in advance). The final evaluation will consist in a practical project (working on own data will be encouraged).

Recommended readings: Course slides and code provided by the lecturer.

Students should read the paper "Quince *et al.* (2017) Shotgun metagenomics, from sampling to analysis. Nature Biotechnology, 35(9), 833-844", which will be discussed during the first class.

Basis of bash are required.