# Introduction to Transcriptomics Analysis

## Class 05 - Manipulations of Sequence Files in Linux II.

**INSTRUCTOR:**
Aureliano Bombarely
Department of Bioscience
Universita degli Studi di Milano
aureliano.bombarely@unimi.it

# Outline of Topics

- Exercise 1: Removing adapters and trimming low quality.
- Exercise 2: Evaluating quality

- Exercise 1: Removing adapters and trimming low quality.

Goal: Using fastq-mcf command remove the adapters and the low quality (-q 30) nucleotides for the FASTQ files. If the read has a length < 50 bp, remove the whole pair (-l 50). The output should be gzipped too.

Input: All the Artha_*.fastq.gz files

Recommended command: **fastq-mcf -q 30 -l 50 -o filename_q30_l50_R1.fq.gz -o filename_q30_l50_R2.fq.gz Adapters.fasta filename_R1.fastq.gz filename_R2.fastq.gz**

- Exercise 2: Evaluating quality

Goal: Using fastqc command generate a report for each of the raw and processed FASTQ files. Transfer the reports to your computer with FileZilla and open them with a web browser (e.g. Chrome).

Input: All the Artha_*.fastq.gz files

Recommended command: **fastqc filename**.