

### **Assignment: De Novo Genome Assembly**

- 1) Write a short literature review (~2 pages) on De novo sequence assembly algorithms indicating their main technique, improvements with respect to the then state-of-the-art in terms of their experimental results, and limitations.
- 2) Use an implementation of one of the algorithms you mentioned in 1) above to assemble the paired-end sequence reads provided in the files "mutant\_R1.fastq" and "mutant\_R2.fastq". Note that these files contain paired-end reads, where reads from one end of each sequence fragment is in one file and the other end is in the other file at the corresponding row. You can view [this](#) video for a better understanding of paired-end sequencing.
- 3) Evaluate the quality of your assembly at least in terms of the N50 length and genome coverage. Use the reference genome in "wildtype.fna" for evaluation purposes. Explain what these numbers mean. You may have to change the parameters of the algorithm to improve your assembly.