**2017 BI3204 Bioinformatics Course**

**Instructor**: Dr. Jiankui He (hejk@sustc.edu.cn)

**Time**: Thursday. The lecture will be at 2-4pm, and lab will be at 8-10am and 4-6pm.

**Location**: The lecture will be at 505, The First Teaching Building. The lab will be at #4, The Second Teaching Building.

**Course materials:**

* The course materials including slides will be available in the server.

**Goal:**

Understanding the basic principles of next generation sequencing technology. This includes basic biological applications, basics in data processing, statistical and informatics theories in data analysis, advantages, limitations, and assumptions of different methodologies, and biological interpretation of the results.

**Expected outcomes:**

* Understand the basics of biological questions
* Implement different analysis tools (such as alignment, alternative splicing, variant call, ...)
* Fully understand the standard data structure; capable of retrieve relevant information through basic programming
* Understand statistical/computational models to solve different biological questions

**Lecture schedule**

Session 1 Overview of next generation sequencing technology

Session 2 Data Processing

* Analysis workflow
* Sequence quality evaluation
* Data formats
* Data visualization

Session 3 DNA sequencing

* NGS and diseases
* Sequencing mappability
* Refined alignment
* Variants identification

Session 4 RNA sequencing

* Biological theories on RNA-seq experiments
* Experimental considerations
* Major scientific advance using RNA-seq
* Alignment
* Gene expression analysis
* Differential expression analysis
* Alternative splicing
* Allele-specific expression
* RNA editing

Session 5 Student presentation

**Schedule for Lab**

DNA sequencing

Step 1: QC

* 1. Per base Q value
  2. Per sequence Q value
  3. GC content distribution

Step 2: Mapping

2.1 Read position

2.1 generate SAM file

Step 3: Variants calling

2.1 call all SNPs and generate VCF file

RNA sequencing

Step 1: Get gene position annotation

Step 2: mapping

Step 3: calculate FPKM