**Expression profiling by high throughput sequencing: RNA-seq of longissimus dorsi muscles in Japanese black (Wagyu) and Chinese Red Steppes cattle**

**Important to note**

You are provided RNAseq samples from two cattle breeds. The FASTQs were sampled from larger files using seqtk and the reference genome contains five chromosomes. This was done so that the computations can be run on a laptop.

**Project description**

Comparative transcriptome profile of genes differentially expressed in longissimus dorsi muscles between Japanese black (Wagyu) and Chinese Red Steppes cattle by RNA-seq Overall design: RNA-seq analysis were performed in longissimus dorsi muscles samples between Wagyu and Chinese Red Steppes cattle using sequencing platform of the Illumina HiSeq2000, respectively

Total RNA was isolated from longissimus dorsi muscles using Trizol Reagent (Invitrogen, USA) according to the manufacturer's instructions. Total RNA was treated with DNase I (NEB, Beijing, China), extracted with phenol-chloroform, and precipitated with ethanol. The quality and quantity of total RNA were determined using an Agilent 2100 Bioanalyzer (Agilent technologies, Palo Alto, CA). The mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. 1.5 µg of mRNA from each sample was used to construct six cDNA libraries for sequencing. The mRNA was treated with a Thermomixer (Eppendorf AG, Hamburg, Germany) to generate fragments with an average size of 200 bp (200 ± 25 bp) for the paired-end libraries. The fragmented mRNA was then used as templates for synthesizing the first-strand cDNA. The double-stranded cDNAs were purified and ligated to adaptors for Illumina paired-end sequencing. Library concentration was quantified by qPCR and a Qubit® 2.0 Flurometer (Life Technologies, CA, USA), and the insert size was checked on an Agilent Bioanalyzer 2100 system (Agilent Technologies, Palo Alto, CA).

| **Sample** | **Breed** |
| --- | --- |
| SRR13107018 | Japanese black (Wagyu) cattle |
| SRR13107019 | Japanese black (Wagyu) cattle |
| SRR13107020 | Japanese black (Wagyu) cattle |
| SRR13107021 | Chinese Red Steppes cattle |
| SRR13107022 | Chinese Red Steppes cattle |
| SRR13107023 | Chinese Red Steppes cattle |

Design two pipelines that utilise the data provided

1. Variant calling from RNAseq using GATK best-practices
2. Differential expression analysis using any combination of tools discussed in the previous class