Course Report: Next Generation Sequencing Bioinformatics

Wellcome Genome Campus, Hinxton, Cambridge, UK.
1-7 October 2017.

Ramasamy Kandasamy
D4 student.
Division of Neurogenetics

Next Generation Sequencing (NGS) Bioinformatics is an annual course offered by "Wellcome Genome Campus Advanced Courses and Scientific Conferences". This year the course was held at the Wellcome Genome Campus at Hinxton. This course comprised of several lectures and exercises on analyzing NGS data. Twenty four students from 16 different countries took part in this course. All the course lectures and practical sessions took place from day-1 to day-5. In these five days, starting with an introduction to various NGS technologies, this course took us through various steps followed in analyzing NGS data and ended with a short project. Here I will give a brief summary of the course.

Day-0:

The course started with short talks by the participating students about their current research. This session was very interesting to me because most of the students were working diverse topics such as parasitology, virology, cancer biology, human genetics etc. I was amazed to see how various NGS technologies were applied people working in different fields. We also had a networking event which helped us to get to know each other.

Day-1:

During the first day we had two sessions. In the morning session we were introduced to various NGS technologies and their application. I learnt about the advantages and disadvantages of various NGS technologies and which technology is suitable for what kind of applications. We also had a tour of the sequencing facility at the Wellcome Trust Sanger Institute (WTSI). WTSI played a very important role in the historic Human Genome Project and is currently taking part in various important genome sequencing projects. As a consequence a tour of their sequencing facilities was great experience for me. There were several sequencing systems of several kinds and this gave the feel of being in a factory. The afternoon session introduced us to UNIX/LINUX environment. We gained skills that are fundamental to all the future sessions.

Day-2:

Second day's lectures and exercises dealt with various file formats used to store NGS data, preprocessing raw NGS data and to examine their quality. Following this we learnt about aligning the reads to a reference genome. These steps are crucial for further analyses and are essential for various kinds of NGS data analysis.

Day-3 to Day-5:

During these days we had lectures and exercises focusing specifically on four different types of NGS experiments and various tools that used to analyze these data. On Day-3 we learnt how to identify nucleotide and structural variants from the sequencing data. This is useful for researcher trying to identify disease causing variants and cancer researcher looking at mutation and deletions in the genome of cancer tissues. On Day-4 we learnt how to analyze ChIP-seq and RNA-seq data. This part

of the course was the most interesting for me because in my project I am using RNA-seq. I learnt how to use two new tools (that I didn't know existed) to assemble reads and quantify the gene expression. I also came to know about genome viewers and how to use them to visualize the reads aligned to the reference genome. On the last day (Day-5), we learnt to assemble the reads de novo. This is essential when a reference genome is not available. In the afternoon I took part in a group task. In this we used a published RNA-seq data of a Plasmodium and analyzed them. Finally we, discussed our results and interpretations.

Seminars:

Besides the courses, we also had the opportunity to attend interesting seminars that made use of NGS technologies to address various biological questions from different fields of biology. A seminar by Prof. Steven Munger from Jackson Laboratories was the most interesting for me. In this talk Prof. Munger talked about using genome sequencing, RNA-seq and proteomics to correlate genetic variation to mRNA expression and protein expression. This study also revealed several protein-protein interaction networks. I think the knowledge from this research might be useful for subsequent analysis of my data.

Conclusion:

This course was very useful and I learnt a lot of things NGS data analysis in a short period of time. I gained skills and knowledge that would be very useful for my research. I would strongly recommend this course for anyone interested in learning NGS data analysis. The accommodation and food at Wellcome Genome Campus were excellent and comfortable.

Clockwise from top: Willow court, I was given accommodation here during the course. Tennis court suite, lectures and exercise took place here. I visited Buckingham palace on the last day.