

Bioinformatics analysis of lateral gene transfer
Sim Lim, University of Adelaide

Lateral gene transfer, sometimes known as horizontal gene transfer is a process by which an organism incorporates genetic material from another organism without being an offspring of that organism. In the past years, a number of studies observed lateral gene transfer process from eukaryote – eukaryote and prokaryote – prokaryote. Due to the fast development of bioinformatics field and availability of genome sequence, fragments of prokaryotic genes were detected in eukaryotic genomes. Specifically, Dunning Hotopp et.al showed strong evidence that gene transfer from bacteria to *Drosophila* is possible.

Recently, David Adelson (my supervisor) accidentally discovered that two glean models from the *Bos taurus* genome, contained DNA sequences which is similar to some bacteria gene. Therefore he wanted to carry out systemic analyses to identify potential bacterial to bovine gene transfer event. The aim of this summer scholarship is to analyze and identify of possible lateral gene transfer event from bacteria to mammals by using the bioinformatics analysis.

Before I started the summer scholarships, I didn't have any knowledge of the computing or programming. I completed a Bachelor of Science majoring in genetics and microbiology, but the courses that I took did not provide any computing skills training in bioinformatics. The training they provided so far was how to use BLAST and other simple search job for bioinformatics through the GUI style and NCBI webpage.

Therefore when I started to work under Dave Adelson, I didn't have any idea how to use the command line to run BLAST or other bioinformatics software under Linux. When I first started using the Linux programs, I was really uncomfortable because I was "spoiled" with the GUI program which used in the Microsoft Windows in desktop or laptop.

As a result of this summer scholarship I have learned how to use the Unix/Linux command line, write simple shell script using the emacs editor, create simple database to carry at my research. As a result, I have identified a large number of regions in the bovine genome assembly with significant similarity to bacterial DNA sequencing confirming and extending the initial hypothesis.