Bioinformatics on Big Data: Computing on the Human Genome

**Date**
July 14-15, 2016 - Downtown Toronto, Ontario

**Faculty**
Michael Brudno, George Mihaiescu, Francis Ouellette and Mark Phillips

**Course Objectives**
Several big data genomics projects, including the ICGC, are deciding to host their data in the Cloud and to provide access to configurable virtual machines (VM) with which to compute on this data (thereby removing the need to purchase and maintain your own compute cluster). Similarly, many labs are moving to renting compute time from various cloud providers. Analysis of a single genome or a smaller selected subset differs from analysis of multiple genomes, particularly in the compute infrastructure required.

To navigate through working in this new compute space, the CBW has developed a 2-day course providing an introduction to security and privacy issues related to working on human genome data and the processes necessary to access such data. After reviewing cloud computing infrastructure, the workshop will also provide a hands-on introduction to launching and configuring your own virtual machine (VM), accessing cloud-based data sets, and how to scale up the number of VMs to meet your analysis needs. Customizing VMs with your own tools and cloud-computing best practices will also be discussed.

Participants will gain practical experience and skills to be able to:

- Launch their own virtual machine (VM)
- Configure a VM with prepackaged tools
- Pull in data sets from Cloud repositories
- Follow best practices in data and workflow management
- Customize a VM with their own tools
- Scale up their VM to meet their analysis needs

For information send an email to: course_info@bioinformatics.ca

For course requirements and to apply visit: bioinformatics.ca