

## 2<sup>nd</sup> International Conference on **Predictive, Preventive and Personalized Medicine & Molecular Diagnostics**

November 03-05, 2014 Embassy Suites Las Vegas, USA

## Organization and analysis of large scale mutation datasets produced by next generation sequencing technologies

## Alireza Hadj Khodabakhshi

Genomic Institute of Novartis Research Foundation, USA

The advent of Next Generation Sequencing (NGS) technologies has significantly improved our understanding of many diseases and enhanced our ability to treat them. As NGS technologies advance, researchers are overwhelmed by the massive volume of genomic and transcriptomic sequence data generated. Unfortunately the existing computational infrastructures cannot keep up with the ever increasing data generation capacity of NGS. This shortage will become a greater challenge when the NGS data begin to be utilized in clinical and healthcare applications, i.e. personalize medicine. In the near future large clinical labs need to employ computational infrastructures that are capable of organizing and processing hundreds of NGS samples on daily basis. In this research we present our efforts in creating such infrastructure that includes a consortium of databases for organizing various genomic mutation types and a high throughput mutation analysis pipeline. The high throughput computational pipeline is currently developed to process the SNP and Indel variation types. It is capable of processing the entire variation set in our database (over 4 billion variations across more than 4000 sample) in a timely fashion with the goal of identifying recurrent mutational targets in various cancer genomes. This pipeline can also utilize the resources on a computational cluster to statistically validate mutations by processing raw data in a parallel way. We have successfully applied this pipeline to perform a comprehensive study of aberrant somatic hypermutation (aSHM) targets in Diffuse Large B-cell Lymphoma (DLBCL). This study led to the discovery of 32 novel targets of aSHM in DLBCL.

## **Biography**

Alireza Hadj Khodabakhshi has completed his PhD at the age of 31 years from Simon Fraser University in Canada. He has completed two Postdoctoral studies in BC Cancer Research Centre and Canada's Michael Smith Genome Sciences Centre. He held a Scientist position at Canada's Michael Smith Genome Sciences Centre and is currently a Research Investigator in Genomic Institute of Novartis Research Foundation. He is the author or co-author of more than 20 articles of reputed journals.

alirezak@bcgsc.ca