Making Sense of Individual Genome in a Hospital Environment

A dissertation submitted to
The University of Manchester
for the degree of
Master of Science in Advanced Computer Science
in the Faculty of
Engineering and Physical Sciences

2012

Nabanita Basu

School of Computer Science

Abstract

The project ‘Making Sense of individual genome in a hospital environment’ is co-sponsored by the National Health Services and the National Genetic Reference Laboratory, Manchester, United Kingdom. It involves extensive study of the tools that are used for predicting the effect of non-conservative missense mutations that are in general harmful to humans. With the successful completion of the ‘Human Genome Project’ in 2003 now obtaining the gene sequence of an individual has become cheap (only a few thousand dollars) and easy. With a gene sequencer tool in every genetic research laboratory and a wide array of gene SNP (Single Nucleotide Polymorphism) prediction tools, we are currently on a cusp of revolution in healthcare genomic medicine which focuses largely on patient diagnosis and treatment based on the information from a person’s DNA sequence or genome. Single Nucleotide Polymorphisms (SNPs) have been found to be involved in the etiology for many complex diseases in humans and hence are of particular interest in pharmacogenetics. Increased knowledge and better use of technology for deciphering genomic as well as genetic information will facilitate better/more accurate, early identification of diseases. It shall also contribute to the development of better therapies for treatment of diseases. New technology and genetic data forms the basis for reclassification of diseases. This piece of research is an attempt to find out a tool that is not only robust but also makes trustworthy, near accurate prediction of the effect of missense mutation in most cases. Polyphen-2 and SIFT toolkit have been marked out to be the best missense prediction tools, on the basis of rigorous research. Apart from providing good levels of prediction accuracy, these tools have also withstood the test of time. However, Polyphen-2 has been found to be more robust than SIFT. Hence Polyphen-2 has been test integrated on the Amazon EC2 cloud.