Next Generation Sequencing Analysis on Next Generation computing infrastructure of CESGA
 - Actualizado (04.11.2010)

Date: Thursday 11 November 2010
Time: 12:00 am
Place: Conference room CESGA
Speaker: Abhay Krishna Bioinformatician Eye Disease: from genetics and   epigenetics to therapy research group CABIMER Sevilla Andalucía Spain
New DNA mega-sequencing technologies pioneered and developed in form of various sequencing platforms including Roche 454 GS System®, Illumina Genome Analyzer®, Applied Biosystems SOLiD® System, Helicos Heliscope®, Complete Genomics®, and others render much of the traditional methodology (Sanger sequencing) impractical. With the advent of new technologies of sequence generation there are several computational challenges in assembly of genomes and analysis of the massive datasets that are generated by these mega-sequencers.
In the first place, next-generation sequencers have a much greater throughput at much reduced cost, allowing new possibilities and also posing new analysis and computational challenges. These challenges are mainly due to the amount of the data that these mega-sequencers generate and comparably smaller read lengths of these mega sequencers compared to Sanger method .
The need to adapt current methods to the possibilities of new technologies has been pointed out by many experts. Therefore, it is paramount to discover and apply new techniques in order to extract meaningful information and truly exploit the data mining opportunities presented by the huge amounts of data available.

The technique presented herein receives the name of Multi-Resolution Analysis (MRA) using Wavelet transform (WT), that has solved several signal filtering/noise removal problem in datasets of comparable size eg. finger prints data with unmatched efficiency . In this kind of data processing, the data is split into different resolution levels, so that different information scopes can be treated and analyzed independently. This modality of analysis is especially suitable for location and classification of patterns at all different resolutions levels. Supercomputing infrastructure like CESGA can be leveraged for such an integrative analysis.