Abstract

A raw genomic sequence is of no value without its structural and functional elements being annotated. Having an essentially complete genome sequenced, Sri Lankan Human Genome, the objective of the research project was to annotate a sample sequence from Sri Lankan Human Genome and implement an automated annotation pipeline.

With increasing speed and accuracy, automated annotation increasingly has a place even though manual annotation is considered the gold standard. Tools that need to perform annotation are freely available as installable applications or web services. Using these

web services reduce the overhead of maintaining the necessary hardware and software. Though few catalogues do exist that list some of those free web services, most of them are outdated. Also published articles comparing the quality of these web services are almost non-existent.

Due to difficulty in measuring the suitability and quality of various web services that can be used to perform the annotation, services from well-known institutions and those that are used frequently by the community were used in this annotation pipeline.

Annotation was automated using Perl scripts which would bind several web services and resources to come up with annotations. The annotations were published using tracks added to the GBrowse genome browser which is used to visualise the Sri Lankan

Human Genome at http://www.srilankangenome.net/.