

Web-based system for analysis and management of next-generation genomic sequence data

Byungwook LEE and Minkyung SUNG

Korean BioInformation Center (KOBIC), KRIBB, Aeung-dong, Yuseong-gu, Daejeon 305-817, Korea

Abstract

With the rapidly falling cost of high throughput sequencing technology, the bottleneck of effectively using genomic analysis in the laboratory is shifting to one of effectively managing, analyzing, and sharing these massive genomic data. Here, we present web based system, called NGSpass, for analyzing and managing NGS genomic sequence data. Our system accepts a FASTQ-formatted sequencing file as inputs, and it then executes back-end analysis pipelines already constructed by users. Users can simply build analysis pipelines by adding or deleting programs and adjusting parameters of each program. Users also easily monitor running states of each NGS project. In addition, our system contains a module that allows researchers to build a Sequence Read Archive (SRA) submission files with their NGS and related data. Final results can be easily downloaded using a web browser. Our system is very flexible to update and modify because it has been developed using Google Web Toolkit (GWT) based on JAVA and JavaScript language and MySQL database. It has user-friendly interface and can be installed at any platform such as Linux, Mac and Windows. We believe that our web-based system will be very useful for NGS genomic data analysis and processing.