

유전자 정보시스템 설계 및 구현 Implementation of Gene Information System

최 낙 중, 최 한 석*, 김 동 옥**
국립식량과학원, 목포대학교 컴퓨터공학*, 올디텍**

Nak-Joong Choi , Han Suk Choi*, Dong-Wook Kim**
National Institute of Crop Science,
Department of Computer Engineering Mokpo
National University*, AIIDTec. co. Ltd**

ABSTRACT

We have developed a web server for the high throughput annotation of gene. This system processes entire data sets with an automated pipeline of 13 analytic services, then deposits the data into the MySQL database and transforms it into three kinds of reports: preprocessing, assembling and annotation.

I. Introduction

As Genes are used in various areas of biological research, the generation of Genes has grown exponentially, and the need to handle large Gene data sets has also grown. Thus, high-throughput Gene sequencing technologies have steadily advanced to manipulate huge Gene data sets more quickly and accurately.

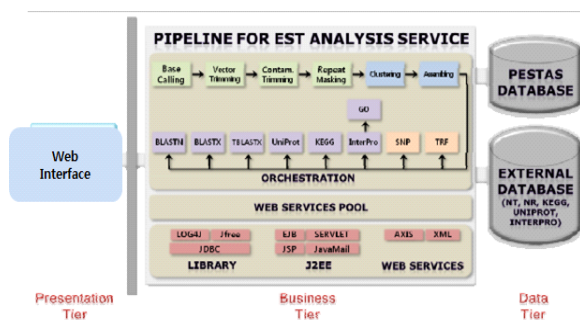
in the pipeline is then deposited into the data tier.

2. Key features

Web service technologies were used to convert thirteen analytical tools into web services. Each analytical web service can be executed by the user-controlled pipeline or individually. And since the web services are component-based and loosely coupled, This system is an extensible and flexible system.

II. IMPLEMENTATION

1. Architecture



▶▶ Fig.1. system Architecture

This System has a three-tier architecture. A presentation tier helps the scientist interact with this system through a web browser. In the business tier, a user-controlled pipeline is created with user-specified options and analytic services. All information generated



▶▶ Fig.2. System Interface

This system provides three kinds of reports: pre-processing, assembling and annotation (see Supplementary material). The preprocessing report presents information generated during the cleansing step (base calling, vector trimming, contamination trimming and repeat masking) of the pipeline module.

The main function of the curation service is to give

more intuitive and effective information to scientists. In general, automated highthroughput Gene analysis pipelines annotate sequences by using BLAST algorithms to search for similar sequences in various relevant databases[1][2].

III. Conclusion

This system is a valuable web server for analyzing gene sequences through an intuitive web interface. It provides various analytical services, detailed report services, a novel user-specifiable curation option and a comprehensive gene-chip analysis tool. This system has a user-friendly interface and well-organized contents. Response and uploading times are fast due to the use of Ajax. Every interaction between the scientist and This system is performed through a web interface. Compared to other methods, This system provides a more convenient research environment for extracting biologically relevant information from gene Data sets.

■ References ■

- [1] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. *J Mol Biol* 1990, 215:403-410.
- [2] Aurrecoechea C, Brestelli J, Brunk BP, Dommer J, Fischer S, Gajria B, Gao X, Gingle A, Grant G, Harb OS, et al: PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Res* 2009, 37:D539-543.