Web-Based System for Sample Tracking in Rice Genome Research Program

Kazunori Waki[1] waki@staff.or.jp
Kimiko Yamamoto[1] yamamoto@staff.or.jp
Isamu Ohta[1] ohta@staff.or.jp
Masaki Kise[2] kise@mi.mss.co.jp
Katsumi Sakata[1] ksakata@abr.affrc.go.jp
Takuji Sasaki[1] tsasaki@abr.affrc.go.jp

[1] Rice Genome Research Program, NIAR/STAFF 446-1 Ippaizuka, Kamiyokoba, Tsukuba, Ibaraki, 305-0854, Japan
[2] Mitsubishi Space Software Co.,Ltd.
1-17-15 Sengen, Tsukuba, Ibaraki, 305-0047, Japan

Keywords: sample tracking, genome sequencing, rice, web-based system, graphic view

1 Introduction

Rice is one of the major cereal crops and is the principal source of food for about half of the world's population. In terms of genome analysis, it has an advantage among other cereals because it has the smallest genome size estimated at 430 Mb. At the Rice Genome Research Program (RGP), sequencing of the entire genome was launched in 1998. So far, more than 1.9 Mb of genome sequence has already been completed and made available to the public domain through the DNA Data Bank of Japan (DDBJ) and RGP home page (`http://www.dna.affrc.go.jp:82/`).

Currently, RGP is processing approximately 3000 samples per day in various stages of genome sequencing. This is expected to increase in the near future as we accelerate the sequencing process. In order to monitor the progress of sequencing and the quality of sequence data, a sample tracking system has been devised [1]. The system was built using a web-based software, SDTM (Sample Data Tracking Manager) , and the first operational version has been completed and evaluated. The major features of the system include: (1) comprehensive system covering various processes in the laboratory from sequencing reaction to annotation, (2) graphic view of data based on a web-interface, and (3) bar-code data management. The system provides a lot of useful information necessary for management of large scale sequencing projects.

2 Laboratory Process and Sample Tracking System

The different processes involved and data derived in each step could be systematically managed from the web-based system for sample tracking (Fig. 1). This strategy also provides an efficient method of tracking the sample or dataflow at different stages as well as the data produced and recorded. The system updates a work history file in a server computer by running a command at the end of each day to check the resulting data at each stage. The progress and quality of information are processed using the work history file.

3 Graphic View of Data

A graphical view of data is a main feature of the system that allows the user to access all relevant information. An example of graphic view displaying the progress of laboratory process, number of base assembled and quality of sequence data is shown in Fig. 2.

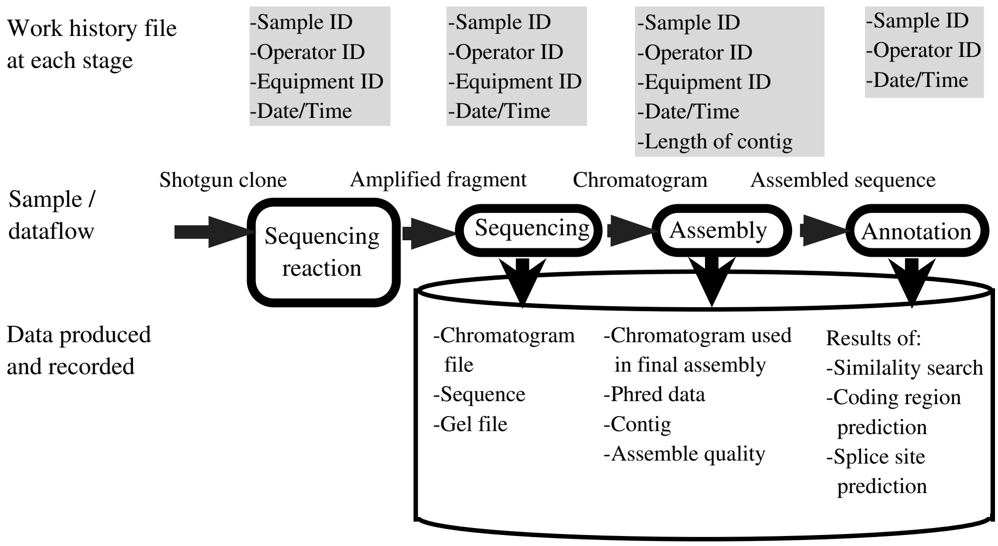

Figure 1: A schematic diagram of the different processes involved and data derived from the web-based system for sample tracking.

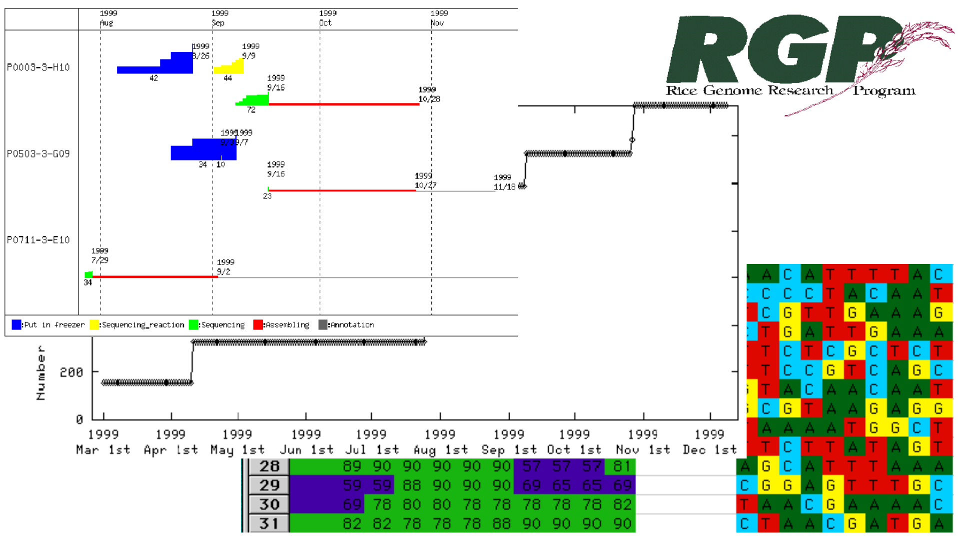

Figure 2: An example of a graphical view of data showing the progress in the different steps of sequencing as well as the quality of sequence data.

References

[1] Sakata, K., Waki, K., Sasaki, T., Shimomura, M., and Kise, M., A sample tracking tool for rice genome sequencing, *Genome Informatics 1998*, Universal Academy Press, 222–223, 1998.