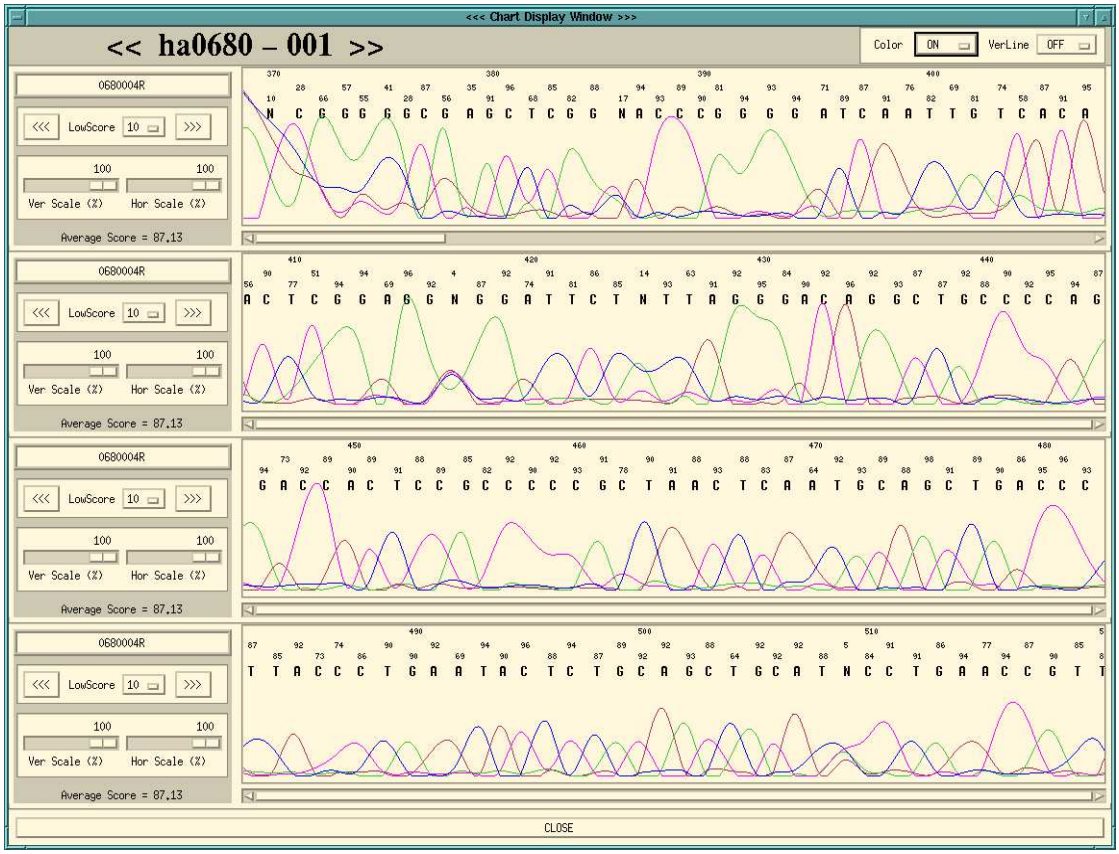

Figure 1: Example of a contig editing by MAEditor

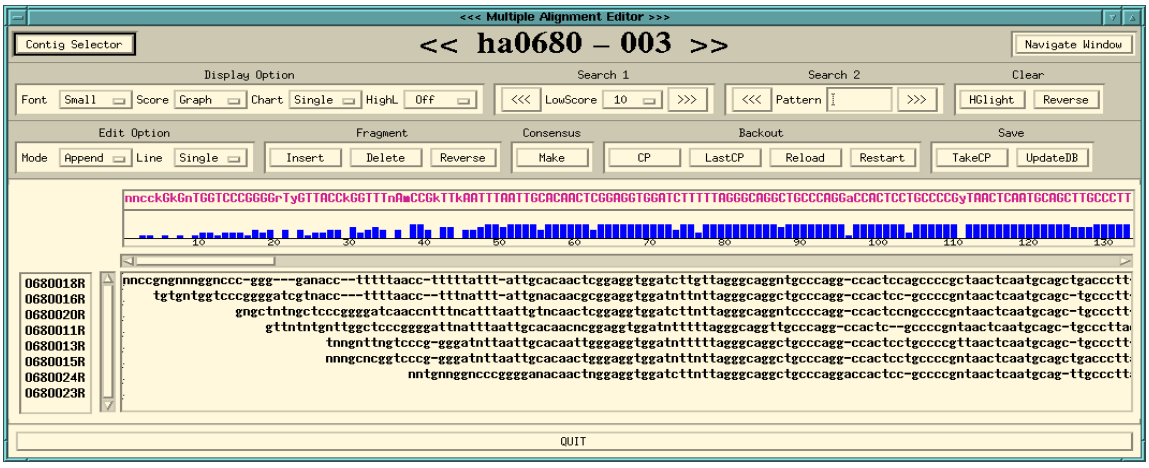

Figure 2: Example of a trace editing by MAEditor

An Integrated System for Large-Scale Sequencing: Dr.AGCT

Zhongqing Wang [1] zwang@kazusa.or.jp
Akira Ohyama [2] akr@hydra.mki.co.jp
Nobuyuki Miyajima [1] miyajima@kazusa.or.jp

[1] Department of Genome Informatics, Kazusa DNA Research Institute
1532-3 Uchino Yana Kisarazu, Chiba 292 Japan

[2] Mitsui Knowledge Industry Co. Ltd. 3-7-4
Kojimachi, Chiyoda-ku, Tokyo 102, Japan

Abstract

We have developed a set of tools for large scale sequence analysis, Dr.AGCT. MAEditor can incorporate the outputs from other assembly software and it can also be used to display a single trace from both ABI and ALF sequencers. Sequence quality control is the most important feature of Dr.AGCT. Other project monitoring systems are also used to reduce the human intervention in several steps of sequencing process.

1 System Description

Current methods of large scale sequence analysis involve the use of various set of software such as database homology searching and fragment assembly. We have developed a set of tools that can be useful for random sequencing and utilize the results of different programs such as GCG, XGAP or SHOTGUN. MAEditor has two mode. One is used to display and edit the contigs assembled by other software then create the contigs that can be reassembled by other software (Figure 1) and the other is used to manage a single trace data from ABI or ALF sequencers. The trace data can be seen in a way as one line or as several lines (Figure 2). The multiple line mode is useful to see the sequence as a whole. Sequence quality can be calculated by using Staden or other methods, and displayed graphically to check the accuracy of the base.

[1]王忠清, 宮嶋伸行：かずさ DNA 研究所，〒 292 千葉県木更津市矢那内野 1532-3
[2]大山彰：三井情報開発（株）〒 102 千代田区麹町 3-7-4