APPLICATION INFORMATION

Genetic Analysis: CEQ Series

BACTERIAL ARTIFICIAL CHROMOSOME (BAC) END SEQUENCING ANALYSIS

Zhiming Jiang
Beckman Coulter, Inc.

Introduction

One of the goals of genetic analysis is to increase the understanding of gene function through the determination of DNA sequences that make up the genome of different species. In search of the best route to these ends, researchers have generated several different types of useful chromosomal maps. Eventually, the genome will be represented by DNA chromosome sequences with various levels of annotation. Bacterial artificial chromosomes (BAC), which typically contain 100 to 200 kb inserts of DNA, were designed as larger, more stable recombinant DNA clones that would represent the genome more uniformly than plasmid systems. Therefore, BACs are important tools for positional cloning, gene analysis, and physical mapping. During studies with BAC clones, it is often necessary to organize them into contigs. To finalize, join, and extend the contigs, both cloning and sequencing of the end of the inserts are required. However, BAC end sequencing has proven to be difficult due to the large molecular weight and secondary structure of the BACs. Here we present a new way of performing BAC end sequencing on the CEQ™ 8000 Genetic Analysis System from Beckman Coulter with high degrees of accuracy, automation, and throughput. The same system could be applied to perform end sequencing of other large DNA constructs such as PAC, YAC, or cosmid, etc. This new method allows scientists to rapidly and efficiently sequence BAC DNA on a large scale, producing sequence data with relatively high signal levels, good quality values, and long read lengths.

Methods

Growth of BAC Clone and DNA Isolation

A total of six human BAC clones (catalog number 96012) from Invitrogen were used for experiments. All clones were supplied as glycerol stocks and revived by transferring a small portion of the frozen sample onto an LB agar plate with Chloramphenicol (12.5 μg/mL) prior to incubation at 37°C for 12–16 hours. Single colonies were isolated from the freshly streaked plate and inoculated in a starter culture of 2-5 mL LB selective medium followed by an incubation of 12–16 hours at 37°C with vigorous shaking (~300 rpm). The starter culture was diluted into 500 mL selective LB medium (1/500 to 1/1000 dilution) and grew for 12–16 hours at 37°C with vigorous shaking. BAC DNA was then isolated using the QIAGEN Large-Construct Kit (catalog number 12462) with an expected yield of 20 to 50 μg BAC DNA from each 500 mL culture.

To maintain fresh stocks of BAC clones, the BAC cultures were streaked onto a selective LB plate every other week and single colonies were isolated into selective LB medium. Glycerol stocks were then made by adding sterile glycerol into the BAC LB culture to achieve a final glycerol concentration of 20%. The glycerol stocks were then mixed thoroughly and frozen at -70°C.

End Sequencing Reaction Using CEQ™ DTCS Kit (P/N 608000) or CEQ DTCS Quick-Start Kit (P/N 608120)

BAC DNA was quantitated by UV spectroscopy at 260 nm. To confirm the size of each BAC clone, 1 µg of BAC DNA was digested by *Hin*dIII at 37°C for one hour and then separated on 1% agarose gel. All clones were shown to have a size of over 100 kb. BAC sequencing reactions were performed using CEQ DTCS kit or CEQ DTCS Quick-Start Kit (Tables 1 and 2). The cycling conditions are listed in Table 3. Sequencing reactions were purified by ethanol precipitation (CEQ DTCS Kit insert). Samples were dried in a speed vacuum, resuspended in 40 µL SLS, and separated on the CEQ 8000 Genetic Analysis System using the run method "LFR-a" (Table 4).

Results

BAC Sequencing Analysis on the CEQ 8000 Genetic Analysis System

Raw data were collected and analyzed by the sequencing module of the CEQ 8000 software. All six clones were sequenced multiple times using the protocol described above. Figure 1 shows the typical sequencing results from BAC clones #1 and #6. Relatively high signal levels were achieved with high-quality value scores and long read lengths. Control sequences were obtained from GenBank and utilized as alignment templates. The 98% cutoff base number, a measurement of data accuracy, was calculated using the batch alignment function of the CEQ 8000 software. As shown in Table 5, using the recommended dye-terminator sequencing protocol and CEQ 8000 system, we were able to sequence all six BAC clones with high degrees of accuracy. The overall average 98% cutoff base number from all six clones was 668 bp, whereas the overall average read length was 796 bp. Our end-sequencing success rate is approximately 95%.

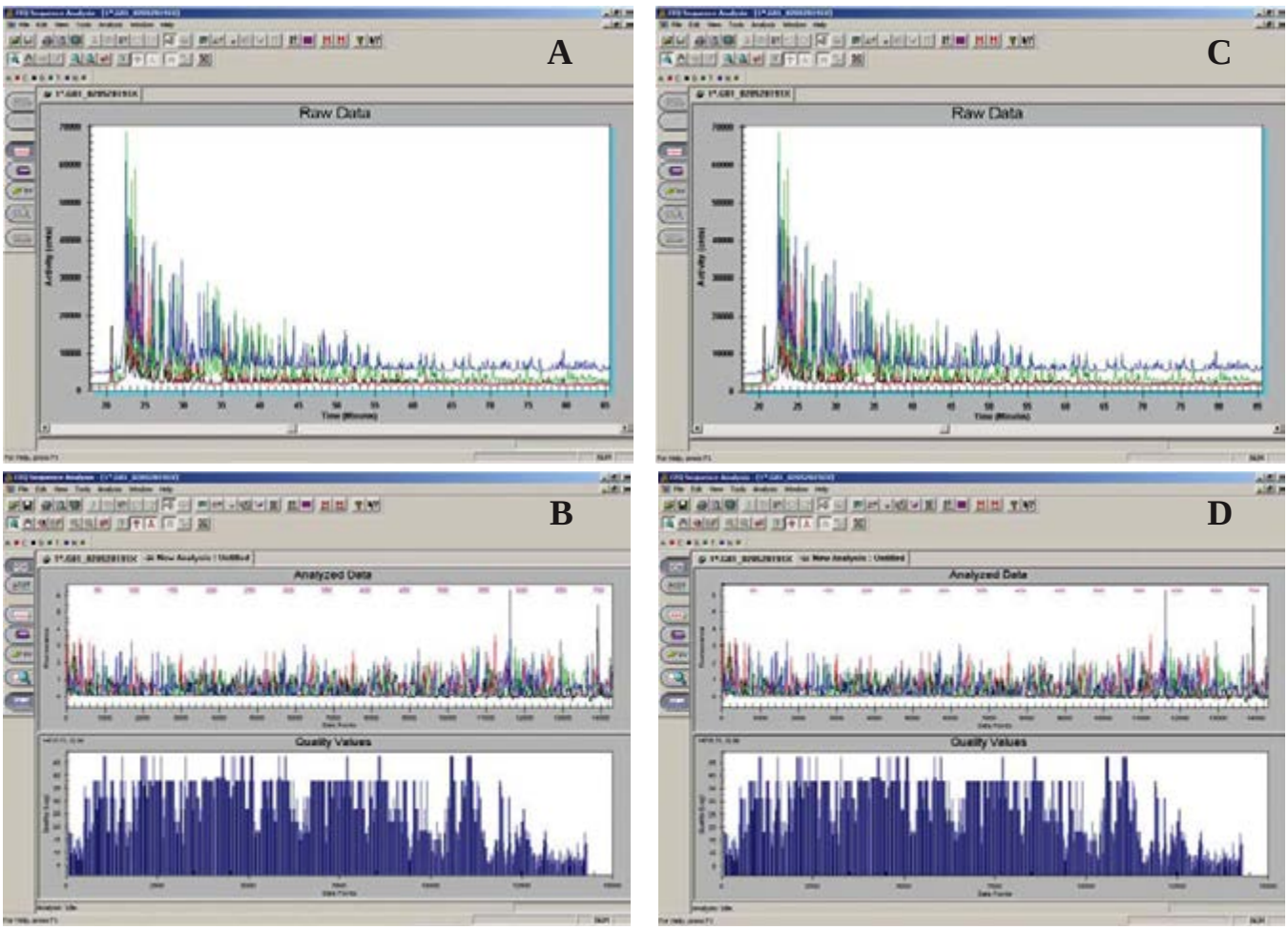

Figure 1. *Typical BAC sequencing results with relatively high signal level, good quality values, and long read length.* ***A:*** *Raw data—clone 1.* ***B:*** *Analyzed data—clone 1.* ***C:*** *Raw data—clone 6.* ***D:*** *Analyzed data—clone 6.*

Table 1. BAC Sequencing Protocol Using CEQ™ DTCS Kit	
Reagents	***Amount***
BAC DNA	2–3 µg
End Primer	60 pmoles
DTCS Premix	12.0 µL
Water	q.s.
Total	20.0 µL

Table 2. BAC Sequencing Protocol Using CEQ DTCS Quick-Start Kit	
Reagents	***Amount***
BAC DNA	2–3 µg
End Primer	60 pmoles
DTCS Premix	8.0 µL
Water	q.s.
Total	20.0 µL

Table 3. BAC Sequencing Cycling Conditions		
Temperature	***Time @ Temperature***	***Cycles***
95°C	5 min	1
95°C	30 sec	50
50-55°C	10 sec	
60°C	4 min	
Hold at 4°C	Forever	1

Table 4. BAC Sequencing Run Method "LFR-a"		
Capillary	Temperature	50 °C
	Wait for Temperature	Yes
Denature	Temperature	90 °C
	Duration	120 seconds
Pause	Duration	0 seconds
Injection	Voltage	2.0 kv
	Duration	15 seconds
Separate	Stage 1	
	Primary Voltage	4.0 kv
	Ramp Duration	5 minute
	Stage 2	
	Separation Voltage	4.0 kv
	Start Time	5 minute
	Ramp Duration	0 minute
	Total Separation Duration	110 minutes

Discussion

To achieve the best BAC sequencing results, several issues need to be taken into consideration:

1) Fresh Stocks of BAC Clones

Most of the BAC clones contain 100 to 200 kb inserts of DNA. It is easy for the bacterial cells to lose large vector constructs when stored in selective liquid culture or on agar plates at low temperature. The percentage of BAC-carrying cells decreases dramatically after refrigeration for extended periods of time, with eventual loss of all BAC-containing clones. To avoid the loss of BAC, we recommend streaking the BAC culture onto a selective LB plate every other week and inoculating single colonies into 2-5 mL selective LB medium prior to starting a large selective culture (500 mL). The solubility of chloramphenicol in water is much lower (2.5 mg/mL) than its solubility in ethanol (50 mg/mL). The recommended concentration for chloramphenicol in LB medium or agar is 12.5 µg/mL. However, the higher the concentration (up to 50 µg/mL) used, the more stringent growth conditions appeared to help BAC stability.

2) QIAGEN Large-Construct Kit

High-purity DNA is essential for achieving high-quality sequencing results. We suggested using the QIAGEN large-construct kit for BAC DNA isolation. This kit was designed for isolation of up to 50 µg ultrapure genomic DNA-free BAC, PAC, P1, or cosmid DNA. The recommended culture volume for BAC is 500 mL. Since large DNA constructs are generally present in low or very low copy numbers in cells, the actual yield varies depending on the construct insert, the vector, and the host strain. The instructions in the kit should be followed closely.

In addition, the following suggestions are also helpful for obtaining pure BAC DNA with a higher yield:

- Reduce the culture volume to avoid overloading columns.
- Reduce DNA shearing by minimizing the vortexing or pipetting up and down during lysis.
- If experiencing difficulty in dissolving DNA, check whether the DNA was over-dried or contaminated by isopropanol or salt.
- If the BAC DNA pellet is invisible, rinse the bottom of the centrifuge tube with 0.5 mL of water, transfer the solution to a 1.5 mL tube, and then dry the DNA using a speed vacuum. Avoid over-drying the pellet. Resuspend the BAC DNA in 10–20 µL of TE buffer (pH 8.0).
- Try to dissolve BAC DNA in as low a volume of buffer as possible. The concentration of BAC DNA should be kept at 0.3–1.0 µg/µL.

3) Sequencing Primer

Good primer design is critical for achieving good sequencing results. We recommend using a commercial software package to design sequencing primers. In this study, we used the Oligo 6 software package from Molecular Biology Insights, Inc. Dye-terminator reactions for plasmid vectors usually require 3.2 pmol of primer per reaction. For an equivalent amount of BAC DNA, there are 50–100 times fewer priming sites available. In addition, the BAC target contains increased numbers of imperfect binding sites which may titrate away some primer molecules. Both longer primers and higher primer concentrations seem to be helpful for sequencing quality improvement. Out of the series of primer concentrations we tested, the best working concentration for BAC sequencing is 60 pmol primer/20 µL sequencing reaction (Tables 1 and 2).

4) BAC Template

High-molecular-weight DNA may be hard to dissolve at high concentration. Over-drying of BAC DNA pellets may make this problem worse. Warming the solution slightly, and allowing more time for the pellet to dissolve can help overcome this problem. We performed sequencing reactions using different amounts of template DNA per reaction. The higher amount of BAC DNA in the reaction, the better the sequencing results. However, using too much BAC DNA may result in a sticky cluster in the sequencing reaction, incomplete denaturing of the template, or insufficient primer annealing, which therefore leads to lower signal levels. We recommend using 2.5–3.0 µg BAC DNA in a 20 µL sequencing reaction.

5) Cycling Conditions

Different combinations of denaturing, annealing, and extension temperature and time were tested. The optimal conditions are presented in Table 3. The initial 5-minute incubation step at 95°C is included to insure complete denaturation of the BAC templates prior to primer annealing. Preheating has proven to improve sequencing quality as well.

Table 5. Sequencing Data Summary on All Six BAC Clones

Clone Number	*ResGen Clone ID*	*Human Clone Number*	*Number of Reactions*	*Average Read Length (bp)*	*Average 98% Cutoff Base Number*
1	2015B8	Chr 1 clone RP11-122M14	11	804	689
2	2326G7	Chr 7 clone CTA-281B9	14	821	674
3	2581C3	Chr 15 clone CTD-2033D15	8	695	521
4	2049P16	Clone hRPK.60_A_1	9	780	732
5	2337M5	Chr 17 clone CTD1-2246P4	10	856	640
6	2253H2	Clone hRPK.60_A_1	11	817	756
				Overall Average 796	Overall Average 668

6) DNA Precipitation

Following the dye terminator sequencing reactions, excess free dyes were removed by ethanol precipitation. To achieve good DNA precipitation and avoid losing pellet, we recommend:

- Add stop solution and glycogen as a master mix prior to adding cold ethanol.
- Mix thoroughly before adding the cold 95% ethanol.
- Add cold 95% ethanol to the tube, mix thoroughly, then centrifuge immediately at 14,000 rpm, 4°C for 20 minutes.
- Rinse the DNA pellet twice with 500 µL (instead of 200 µL) cold 70% ethanol followed by immediate centrifugation.
- Avoid over-drying the DNA pellet.

7) Separation on the CEQ™ 8000

Good resuspension of the final pellet in SLS is also critical for good sequencing results from BAC DNA. Add 40 µL of SLS to each DNA pellet and incubate at room temperature for 5 minutes. Vortex for 10 seconds and then pipette up and down multiple times to insure complete resuspension in SLS. Reactions were loaded to the CEQ and separated by "LFR-a" method with a separation voltage of 4.0 kV for 110 minutes.

8) Data Analysis

Raw data were collected and analyzed using default sequencing analysis parameters. Call threshold was adjusted from 0.4 to 0.6 depending on the sequence quality. Control sequences for batch alignment were obtained by a BLAST search in GenBank at: http://www.ncbi.nlm.nih.gov/Genbank/GenbankSearch.html.

Conclusions

- The CEQ 8000 Genetic Analysis System and CEQ DTCS kit or DTCS Quick-Start Kit from Beckman Coulter enable scientists to perform BAC end sequencing rapidly and efficiently on a large scale.
- This new method produces high-quality BAC sequence results with relatively high signal levels, good quality values, and long read lengths.
- Both DNA preparation and sequencing reaction are amenable to future automation. In addition, no other equipment is required other than common lab centrifuges and vacuum manifolds.

For Research Use Only. Not for use in diagnostic procedures.

SCIEX

www.sciex.com/ce

View SCIEX products at www.sciex.com
Find your local office at www.sciex.com/offices

AB SCIEX Headquarters
500 Old Connecticut Path | Framingham, MA 01701 USA
Phone 508-383-7700
www.absciex.com