

Update on the Continuation of the DNA Sequencing Research Group 2000 Study: An evaluation of the methods used to sequence and isolate bacterial artificial chromosomes.

T. Thannhauser¹, L.S. Hall², J. Hawes³, T. Hunter⁴, D. Leviten⁷, K. Knudtson⁶, E. Jackson-Machelski⁵, M.A. Robertson⁸ and J. VanEe¹.

¹Cornell University, Ithaca NY; ²Albert Einstein College of Medicine, Bronx NY; ³Indiana University School of Medicine, Indianapolis, IN; ⁴University of Vermont, Burlington VT;

⁵Washington University School of Medicine, Saint Louis MO; ⁶Univ. of Iowa, IA; ⁷ICOS Corp., Bothell WA; ⁸Cimarron Inc., Houston TX.

Introduction:

At the 2000 meeting of the Association of Biomolecular Resource Facilities (ABRF) the DNA Sequencing Research Group (DSRG) presented the results of a study in which the protocols used to sequence bacterial artificial chromosomes (BACs) in member laboratories were surveyed in an attempt to correlate methodology with sequencing outcome. The results of this study suggest that there are two critical determinants of sequence quality when dealing with BAC templates:

- amount of template used
- number of cycles in the thermocycling program

The correlation between these factors and sequencing quality were not as strong as expected. It was concluded that the weak correlation was due to the accumulation of uncertainty arising from systematic differences in the experimental parameters used in the participating laboratories. To eliminate this uncertainty the DSRG conducted an internal study on the effects of these factors under rigorously controlled conditions. The results of this internal study are presented below and were used to develop an optimized protocol for sequencing BAC templates.

Member laboratories were asked to sequence a sample of the standard BAC template using the optimized procedure to test its efficacy. These results are also presented and clearly demonstrate the effectiveness of this protocol.

About the Template:

The standard BAC used for this study was selected at random from a library of *Brassica oleracea*. A stock of this clone was used to inoculate and grow cultures from which the BAC DNA was purified using the Midi Prep Kit from Qiagen.

This study was intended to explore the effects of template size on sequencing success. This template does not represent a significant challenge from the point of view of GC content (Avg. about 45%).

Methods:

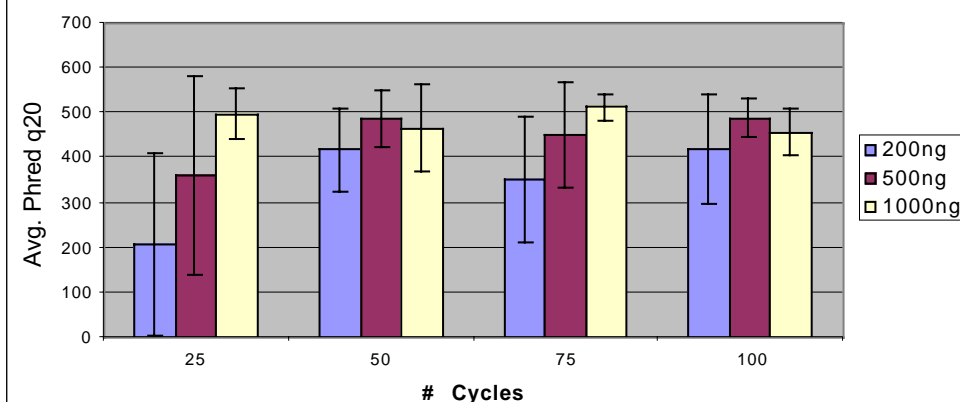
To determine the correlation between amount of template used and length of thermocycling program to sequence quality an experimental program was carried out using the following standard operating procedures:

| Experimental Program | Sequencing Chemistry |
|---------------------------------------|----------------------------|
| (2 X 200 ng) X 25, 50,75, 100 cycles | Reaction Volume: 20 µl |
| (2 X 500 ng) X 25, 50,75, 100 cycles | T7 Primer: 10 pmol |
| (2 X 1000 ng) X 25, 50,75, 100 cycles | Chemistry: Big Dye Term. |
| Total of 96 samples | Premix Volume: 8 µl |
| Thermocycling Conditions | Electrophoretic Conditions |
| Hot Start 95°C 5 min. | 377 Sequencer |
| 96°C 30 sec. | 36 cm WTR |
| 50°C 20 sec. | 0.2mm combs |
| 60°C 4 min. | 64 lanes |
| repeat for N cycles | Module: 36E-1200 |
| 4°C Hold | 10 hours |
| | Temperature 48°C |

The performance level of each submission was judged by sequence quality. To determine sequence quality the unedited sequences were analyzed by phred (1,2). The total number of base calls with a quality value >20 ("phred score") for each sequence was determined with the program "quat" (written by J.V.),

The results of these experiments are summarized in Figure 1:

Figure 1. Effect of # of Cycles on Sequence Quality



Findings:

- For samples containing >1 µg of template a standard 25 cycle thermocycler program is sufficient to produce high quality sequence.
- For samples containing <1 µg of template increasing the number of cycles to 50 will improve sequence quality.
- Increasing the number of cycles in the thermocycling program beyond 50 has little effect on sequencing quality even for very low amounts of template (0.2 µg).

Based on these findings the following protocol is recommended for sequencing BAC templates:

Recommended BAC Sequencing Procedure:

- Template quantity: 1000 ng
- Number of cycles: 50
- Reaction volume: 20 µl
- Primer: 10 pmol total
- Chemistry: Big Dye Terminator
- Volume Big Dyes: 8 µl
- Cycling conditions: Hot start 95° C for 5 min, 96° C for 30 sec, 50° C for 20 sec, 60° C for 4 min x 50 cycles, hold 4 degrees

Samples of the standard BAC template and T7 primer were then sent to participating ABRF laboratories who were asked to sequence them using the recommended procedure. The sequencing results were returned to the DSRG electronically and analyzed for sequence quality as described above. The results of this analysis are shown in Table 1.

Conclusion/ Future Directions:

- The recommended procedure for sequencing BAC templates (see above) is capable of generating high quality sequencing results using the ABRF's standard BAC template.
- However this procedure needs to be tested by sequencing a variety of different BAC templates before it can be considered fully validated.
- Preliminary reports from laboratories that have adopted this procedure suggest that it will prove generally effective.

Table 1. Results of Quality Analysis

| Instrument | Filename | Phred Score | | | WTR | RunTime (hr) | Chemistry |
|------------|-----------|-------------|-----|-----|-------|--------------|-----------------------------|
| | | Q20 | Q30 | Q40 | | | |
| 4200L-2 | 5949A.scf | 805 | 665 | 460 | 66 cm | 12 | IRDye800 Acyclo-Terminators |
| 3100 | 9696a.abi | 537 | 471 | 391 | 50 cm | 2.5 | BigDye Terminators |
| | 9696B.abi | 503 | 399 | 217 | 50 cm | 2.5 | BigDye Terminators |
| 3700 | 7553a.abi | 627 | 553 | 407 | 50 cm | 2.8 | BigDye Terminators |
| | 7553b.abi | 673 | 528 | 333 | 50 cm | 2.8 | BigDye Terminators |
| 373S-36 | 3303C.abi | 21 | 5 | 4 | 36 cm | 14 | Old rhod terminators |
| | 9701A.abi | 472 | 289 | 58 | 36 cm | 12 | BigDye Terminators v2 |
| 373S-48 | 0255A.abi | 639 | 467 | 250 | 48 cm | 15.5 | BigDye Terminators |
| | 0255B.abi | 451 | 292 | 115 | 48 cm | 15.5 | BigDye Terminators |
| | 0583A.abi | 597 | 491 | 421 | 48 cm | 17 | BigDye Terminators |
| | 0583B.abi | 651 | 578 | 474 | 48 cm | 17 | BigDye Terminators |
| | 4666.abi | 5 | 0 | 0 | 48 cm | 18 | BigDye Terminators |
| 377-36-2X | 3358A.abi | 553 | 491 | 363 | 36 cm | 8 | BigDye Terminators |
| | 3358B.abi | 460 | 333 | 127 | 36 cm | 8 | BigDye Terminators |
| | 8488A.abi | 503 | 407 | 217 | 36 cm | 8 | BigDye Terminators |
| | 8488B.abi | 536 | 436 | 264 | 36 cm | 8 | BigDye Terminators |
| | 8980c.abi | 451 | 376 | 274 | 36 cm | 9 | BigDye Terminators |
| | 8980d.abi | 461 | 269 | 126 | 36 cm | 9 | BigDye Terminators v2 |
| | 8980e.abi | 351 | 270 | 171 | 36 cm | 9 | BigDye Terminators |
| | 8980f.abi | 465 | 318 | 187 | 36 cm | 9 | BigDye Terminators v2 |
| 377-48-2x | 2088A.abi | 804 | 699 | 560 | 48 cm | 16.5 | BigDye Terminators |
| | 2088B.abi | 742 | 631 | 520 | 48 cm | 16.5 | BigDye Terminators v2 |
| | 2454A.abi | 652 | 558 | 391 | 48 cm | 14 | BigDye Terminators |
| | 3380A.abi | 613 | 439 | 263 | 48 cm | 18 | BigDye Terminators v2 |
| | 4503.abi | 690 | 592 | 463 | 48 cm | 16 | BigDye Terminators |
| | 5677A.abi | 772 | 686 | 564 | 48 cm | 18 | BigDye Terminators v2 |
| | 5677B.abi | 762 | 650 | 525 | 48 cm | 18 | BigDye Terminators v2 |
| | 5677C.abi | 602 | 395 | 170 | 48 cm | 18 | BigDye Terminators v2 |
| | 5677D.abi | 633 | 401 | 141 | 48 cm | 18 | BigDye Terminators v2 |
| | bac_study | 351 | 195 | 69 | 48 cm | 10 | BigDye Terminators |

Acknowledgements:

The authors would like to thank the following people for their many contributions to this study: Professor June Nasrallah (Cornell University) for providing the BAC clone used in this study; Tom Stelick, Bill Enslow, Tatyana Pyntikova and Farhad Dhabhar (Cornell University) for their excellent technical assistance.

References:

- 1) Ewing, B., Hillier, L., Wendl, M.C. and Green, P. (1998) Base-calling of automated sequencer traces using phred: I. Accuracy assessment. *Genome Research* 8:175-185.
- 2) Ewing, B. and Green, P. (1998) Base-calling of automated sequencer traces using phred: II. Error probabilities. *Genome Research* 8:186-194.