Appendix B

Alignment of sequencing data of clones with reference sequence and Graphical representation of sequences of clones obtained

Analysis of bisulphite treated nucleic acid samples involved PCR amplification of the modified DNA followed by cloning of the PCR products into TA vectors and sequencing of the individual clones. Plasmid clones obtained from the bisulphite experiments were sequenced with M13F and M13Rev18 primers in the CDFD NGTF facility. The electropherogram output obtained was analyzed using Chromas Software to check for clarity of peaks and authenticity of reads.

A. Alignment of sequences with MG1655 reference sequence

The reads obtained from both primers for each clone were subjected to BLAST analysis on the NCBI web server as follows:

i. Each read was aligned against the original sequence of the region being tested from the published *E. coli* K-12 genome to score for C-to-T or G-to-A changes (depending on the strand being sequenced) and also to identify any sequencing errors.

ii. Both reads obtained for a given clone were aligned against each other to ensure authenticity of the reads obtained.

B. Graphical representation of sequences of clones obtained

The sequences of the inserts after BLAST analysis were aligned against the original sequence from the published *E. coli* K-12 genome using ClustalW multiple sequence alignment software. (URL: http://align.genome.jp).

C-to-T changes were graphically depicted by modifying the sequences as follows:

1. The native (original) and modified sequences (obtained from the clones) were pasted in MS-word in plain text format.

2. Using the “select all” and “replace” commands, all the G’s and A’s were changed to stops (.) in Times New Roman font.

3. All the C’s were then replaced with a red “l” in Wingdings font which appears as a red circle.

4. The sequences were then aligned vertically from start to end and the C-to-T changes were observed as T’s in place of the missing red circles when compared with the native sequences.

5. Each C-to-T transition was individually replaced with a blue “l” in the Wingdings font which appears as a blue circle, by referring to the position of the red circle in the parental reference sequence.

6. Finally all the T’s that remained were T’s in the original sequence, the T’s arising due to bisulphite mutagenesis having been replaced by blue circles. These were replaced with stop (.) in Times New Roman font.
Appendix B

Thus the final graphical output appears like beads on a string where the each of the dots of the string represent the spacial positions of the native A’s, G’s and T’s in the sequences. Native C’s appear as red circles while C-to-T transitions appear as blue circles with their arrangement indicating their positions in the original FASTA sequence output obtained. An example of the graphical output obtained is depicted below:

![Graphical depiction of C-to-T changes in fifteen different lac locus clones obtained from a rho mutant strain, GJ6509. The top stretch represents the parental unmodified sequence followed by the sequences derived from the clones. The large red circles at the ends denote native primers where C’s and G’s have not been modified. The large blue circles at the ends denote primers where C’s or G’s have been modified, as described in the text. The small red circles denote the C’s that were not modified by bisulphite while the small blue circles denote the C’s that were modified to T’s in the sequence analyzed.](image)

**Figure B.1** Graphical depiction of C-to-T changes in fifteen different lac locus clones obtained from a rho mutant strain, GJ6509. The top stretch represents the parental unmodified sequence followed by the sequences derived from the clones. The large red circles at the ends denote native primers where C’s and G’s have not been modified. The large blue circles at the ends denote primers where C’s or G’s have been modified, as described in the text. The small red circles denote the C’s that were not modified by bisulphite while the small blue circles denote the C’s that were modified to T’s in the sequence analyzed.

Similar depictions of modified sequences obtained from various clones have been illustrated in Chapters 4 and 5.