Poster H-34

An in silico and in vivo study on the Caulobacter crescentus SOS regulon

Authors:
Apuã C. M. Paquola (*Depto. de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São P aulo*)
Raquel P. Rocha (*Depto. de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São P aulo*)
Rodrigo S. Galhardo (*Depto. de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São P aulo*)
Carlos F. M. Menck (*Depto. de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São P aulo*)

Short Abstract: We have studied the composition of Caulobacter crescentus SOS regulon by computational prediction of LexA binding sites and posterior real-time PCR evaluation of candidate genes. We observed a remarkable correlation between predicted LexA box distance from start codon and level of SOS induction.

Long Abstract:

Caulobacter crescentus is a free-living alpha-proteopacterium that has become a major model organism for prokaryotic cell cycle studies [1]. The SOS regulon, which is well studied in E. coli [2], comprises a number of genes that are activated to maintain cell viability under extensive DNA damage. In this work we describe the in silico and in vivo approach we used to identify new members of the C. crescentus SOS regulon.

A position weight matrix (PWM) motif model was built for the binding sites of LexA (the known repressor of the SOS regulon) by searching for over-represented motifs in the promoter regions of five C. crescentus genes expected to be SOS regulated: recA, lexA, uvrA, recN and imuA. The first four are orthologous to known E. coli SOS genes. The fifth, imuA, was first observed to have its expression induced by UV radiation and was later confirmed by our group to be part of the SOS regulon [3]. The motif model was generated with the Gibbs Motif Sampling program [4] using as parameters a 19nt motif width and the genomic frequencies of A, C, G and T as a background model. The resulting PWM was used to search for candidate LexA binding sites in the C. crescentus genome. ORFs having a high-scoring candidate site from 250nt upstream to 50nt downstream from its predicted start codon were selected for further experimentation. Among the ORFs downstream of the 10 top scoring candidate LexA sites were: the five ones used to build the model, ssb (which is SOS-regulated in E. coli) and others annotated by TIGR (http://www.tigr.org/cmr) as hypothetical or hypothetical conserved proteins.

We have generated a C. crescentus mutant strain with a deletion that renders the LexA protein unable to bind DNA and thus to repress its target genes. To assess the SOS inducibility of genes having high scoring candidate LexA sites in their promoter regions we performed real-time PCR experiments comparing their levels of expression in lexA and wild type strains. As endogenous controls, we used the genes rho and rpoD, whose expression

levels are not expected to vary between these strains. Genes with measured expression ratios larger than 2 between the two strains were considered SOS induced. As real-time PCR results were obtained, we recalculated the motif model, including the promoter regions of the newly discovered inducible genes into the analysis, and updated the binding site predictions.

We have measured the expression ratios of 32 genes, 23 of which were greater than 2, 8 between 2 and 0.5 and 1 less than 0.5, suggesting this last gene (CC2433) is rather activated than repressed by LexA. These numbers indicate that the presence of a putative LexA binding site in its promoter region is a good predictor of whether or not a gene is SOS regulated.

From these 32 genes, we selected the 27 ones that were firstly positioned in their respective predicted operons. We calculated the Pearson correlation coefficient between the expression ratios of these genes and the scores of the corresponding candidate LexA sites. We obtained a low positive correlation of 0.20 ($p=0.14$). We also calculated the correlation between the expression ratios and the distances from each predicted start codon to the respective predicted LexA site. Surprisingly, we obtained a strong negative correlation of -0.68 ($p<10^{-5}$). If we use instead an exponentially decaying function of the distance, namely $0.94^{distance}$, the correlation becomes even stronger: 0.88 ($p<10^{-5}$).

This correlation suggests that there may be a causal relationship between LexA box distance and level of SOS induction in this bacterium. At present, we can only propose two very speculative hypotheses to explain this: (1) That these genes could be transcribed from multiple promoters and the closer the LexA box is to the start codon, more transcription events would be blocked by DNA-bound LexA. This, however, is in contrast with the fact that most E.coli operons are transcribed from one or a very few well defined start sites; (2) That one-dimensional diffusion of RNA polymerase along the DNA molecule might be playing an important role in promoter search and that DNA-bound LexA might restrict this diffusion. According to the model of [5], DNA at a certain range from a promoter could act as an "antenna", capturing RNA polymerases moving by 3-D diffusion and channelling them to the promoter by 1-D diffusion. This is somewhat consistent with the fact that E. coli regulatory regions are enriched in sigma70 promoter-like signals as compared to coding regions or convergent intergenic regions [6].

Future research that could shed light into this question include:
-Determination of whether LexA box to start codon distances are conserved in bacteria;
-Mapping of transcription initiation sites of some C. crescentus SOS genes.
-Construction of an artificial system to measure the expression of a reporter gene using different spacers between a LexA site and the start codon in both wild type and lexA strains.

References:

[1] Skerker JM, Laub MT. 2004. Cell-cycle progression and the generation of asymmetry in Caulobacter crescentus. Nat Rev Microbiol. 2(4):325-37.

[2] Friedberg, EC, graham, CW, Siede, W, 1995. DNA Repair and Mutagenesis. ASM Press,

Washington, DC.

[3] Galhardo RS, Rocha RP, Marques MV, Menck CF. 2005. An SOS-regulated operon involved in damage-inducible mutagenesis in Caulobacter crescentus. Nucleic Acids Res. 33(8):2603-14.

[4] Thompson W, Rouchka EC, Lawrence CE. 2003. Gibbs Recursive Sampler: finding transcription factor binding sites. Nucleic Acids Res. 31(13):3580-5.

[5] Ricchetti M, et al. 1988. One-dimensional diffusion of Escherichia coli DNA-dependent RNA polymerase: a mechanism to facilitate promoter location. Proc Natl Acad Sci U S A. 85(13):4610-4.

[6] Huerta AM, Collado-Vides J. 2003. Sigma70 promoters in Escherichia coli: specific transcription in dense regions of overlapping promoter-like signals. J Mol Biol. 333(2):261-78.