Transcript map and Genome-wide Binding Data for σ^{70}

In a separate study we have determined an unbiased transcript map for *Escherichia coli*, and the genome-wide association of σ^{70}. Some of the data from these experiments is used in the analyses described in Wade *et al.* (2006). This work was performed in collaboration with Nikos Reppas and George Church. Below is a brief description of the methods used and the data generated.

Using a standard transcriptomics approach with high density fully-tiled microarrays (382,177 50-mer oligonucleotide probes spaced every ~12 bp), we identified 1,815 transcripts in *E. coli* grown to mid-exponential phase in M9 minimal media at 30 °C. Comparison of the location of identified transcripts to the gene map of *E. coli* MG1655 and to a 1 Mb rotated gene map confirms a highly significant concordance between the coordinates of identified transcripts and the boundaries of annotated genes (χ^2 p = 1.2e-316), indicating that our transcript map is of high quality.

Using ChIP-chip with the same microarray design and the same cells we identified 1,286 σ^{70} promoters. To validate the ChIP-chip experiment we have performed a ROC (Receiver Operating Characteristic) analysis. We have compared the position of the identified σ^{70} promoters with the positions of 484 experimentally identified σ^{70} target promoters (positive controls; from the EcoCyc and RegulonDB databases) or with the central positions of each of the corresponding genes (negative controls; we reasoned that the central positions of these genes are not σ^{70} target promoters). A σ^{70} ChIP-chip target is defined as being associated with a positive control site or a negative control site if it is within a defined distance (distance discrimination threshold; we set this variable to 180 bp). We have repeated this analysis with a list of 1,286 genomic locations that represent the σ^{70} target sites rotated 1 Mb around the genome (i.e. essentially randomized data).

False discovery rate is the fraction of targets identified by ChIP-chip that are associated with negative controls. Sensitivity is the fraction of positive controls that we identify in our ChIP-chip analysis (note that some of these σ^{70} targets may be repressor-occluded under the conditions tested and hence would not be bound by σ^{70}). Specificity is the fraction of negative controls we identify in our ChIP-chip analysis. Our σ^{70} ChIP-chip data gives a false discovery rate of 1.9%, a specificity of 98.8% and a sensitivity of 62.6%, indicating that it is of high quality.