Dynamics of splicing during the yeast cell cycle

Agnieszka S. Juncker[1*], Rasmus Wernersson[1], Henrik Bjørn Nielsen[1], and Søren Brunak[1]
[1]Center for Biological Sequence Analysis, BioCentrum-DTU, Lyngby, Denmark.
*E-mail: ag@cbs.dtu.dk

Abstract
DNA microarrays have been widely used to investigate the expression of genes during cell cycle progression. We are interested in applying this method for investigating the dynamics of splicing in the eukaryote cell. For that purpose we have designed a Nimblegen microarray targeting all *S. cerevisiae* genes, where probes were designed to cover both exon and intron regions as well as exon-exon and exon-intron junctions. Based on a synchronized *S. cerevisiae* culture*,* we have measured the gene expression for 20 time points, which were estimated to cover two entire cell cycles. In our data analysis we identify the periodically expressed probes and we compare between the intron and exon expression throughout the cell cycle.

Introduction
Genome-wide expression analysis of *S. cerevisiae* genes has previously been used to investigate expression of cell cycle regulated genes[1] as well as the expression of intron containing genes in mutant yeast cells[2]. Here we present an approach to investigate the dynamics of expression of exon and intron regions for 275 intron containing genes throughout the yeast cell cycle.

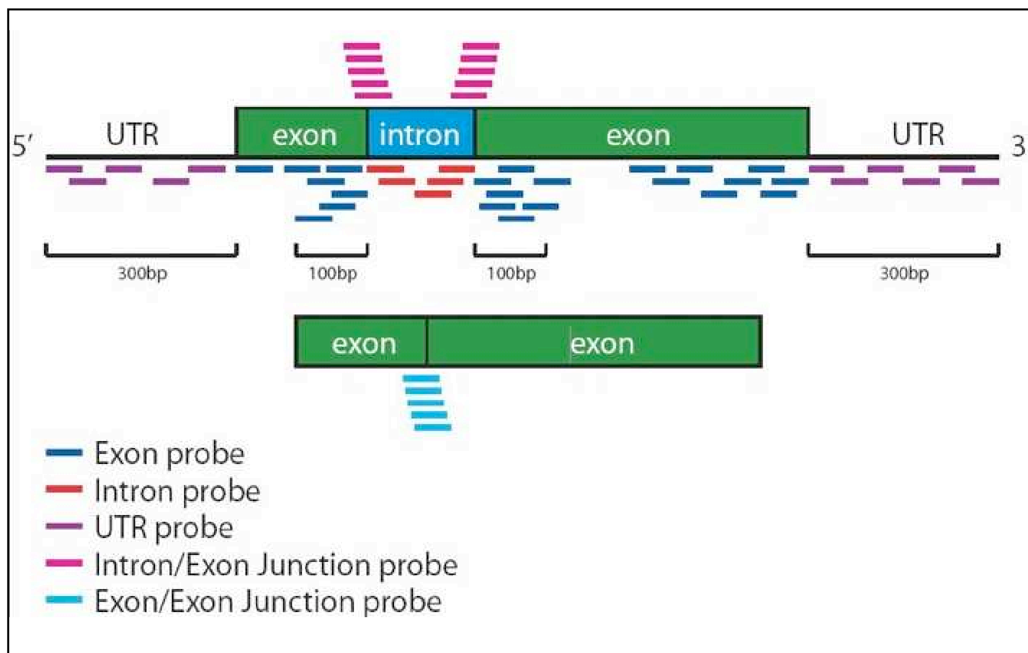

Figure 1. Overview of the probe design for a intron containing gene. Exon and intron regions were densly covered with 10bp as min. distance between probes. Additional exon probes were placed within 100bp from each side of the intron. Intron/exon and exon/exon junctions were covered with min. distance of 2bp, and UTR regions (300bp) were covered with min. distance

Materials and Methods

Probe design. For the subset of intron containing S. *cerevisiae* genes probes were designed to cover both the exon and intron region as well as intron/exon and exon/exon junctions, a total of approx. 64,000 probes.

Experimental procedure. A synchronized culture of the temperature sensitive mutant strain CDC15-2 was grown in a fermentor for 345 minutes, and samples were taken out for hybridization on DNA microarrays in intervals.

Correlation matrices. The statistical analysis was mainly based on correlation measures, where normalized and log-transformed expression values were used to calculate the Pearson's correlation coefficient between each exon and intron probe resulting in a correlation matrix for each intron containing gene.

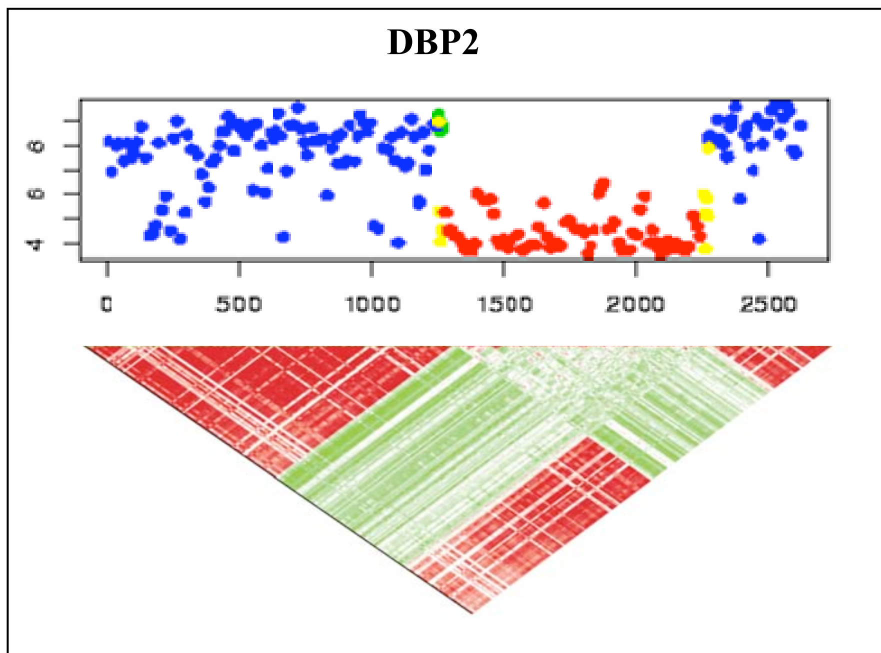

Figure 2. Typical exon-intron-exon signal. Intensity scatterplot and correlation matrix for gene DBP2. The two exons and the intron on both plots are aligned. Intensity scatterplots show the probe expression level as a function of the probe position in the gene sequence. Blue: exon, red: intron, yellow: exon/intron junction, green: exon/exon junction. Correlation matrices visualize the Pearson's correlation coefficient (CC) between probes in time space. Red: CC>0.5, green: CC<0.5, white: CC=0.5.

Results and Discussion

Exon/Intron Expression. For the majority of highly expressed genes with expression levels above detection limit we observed a typical exon-intron-exon signal where intron probes did not measure any activity. For these genes, the absolute exon probe expression levels were significantly higher than the levels for intron probes, and the correlation matrix showed that expression profiles throughout the cell cycle for exon probes correlated with each other but not with the intron probes.

Intron Retention. A small number of genes showed an expression signature where intron probe expression patterns were similar to the ones for exon probes, both when considering the absolute expression levels and the correlation between probes. Among these were two well-known examples of intron retaining genes, a transcription factor HAC1 and a thymidylate synthase CDC21. HAC1 has previously been found to be spliced as a result of induction by ER-stress[4], and for CDC21 it has been reported that the splicing does not occur[5]. Several other intron retention candidates were observed including RPS22B and YHL050C.

Splice Site Reannotation. By expression analysis we were additionally able to identify wrongly annotated splice sites, and one example is gene MRPL15.

References

1. Spellman et al., *Mol Biol Cell* 1998; 9(12):3273-97
2. Clark et al., *Science* 2002; 296(5569):907-10
3. Nielsen et al., *Nucleic Acids Res*. 2003; 31(13):3491-6
4. Kawahara et al., Mol Biol Cell 1997; 8(10):1845-62
5. Taylor et al., J Biol Chem 1987; 262(11): 5298-5307